New Technologies and Survey Data Collection: Challenges and Opportunities[1]

Mick P. Couper
University of Michigan

> "It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of light, it was the season of darkness, it was the spring of hope, it was the winter of despair …" (Charles Dickens (1859), *A Tale of Two Cities*, part 1, page 1)

1. Introduction

How do we view the impact of technology on survey data collection? For some, each new wave of technological innovation brings new opportunity, offering new ways to enhance and extend survey capabilities. For others, each such innovation is viewed as a portend of the end of surveys as we know them, as a threat to all we hold near and dear as survey researchers. More often than not, the voices of the proponents of technological innovation and change drown out those of the detractors in the rush to adopt the latest technology.

To illustrate, Gordon Black (CEO of Harris Interactive) proclaimed: "Internet research is a 'replacement technology' — by this I mean any breakthrough invention where the advantages of the new technology are so dramatic as to all but eliminate the traditional technologies it replaces: like the automobile did to the horse and buggy" (Press release August 1, 1999, Harris Interactive web site). In responding to critics of the Harris Online Poll, Black stated further, "It's a funny thing about scientific revolutions. People who are defenders of the old paradigm generally don't change. They are just replaced by people who embrace new ideas." (*Wall Street Journal*, April 13, 1999). On the other hand, in his AAPOR[2] presidential address, James Beniger, noting that the "era of survey clutter" has already begun on the Web, opined: "Good luck to any serious survey firms which pin much of their futures on the hope of being heard for long above the mounting background noise and confusion of the swelling tide of amateur and slapdash pseudopolls" (1998, p. 446).

The concern about technology's impact on survey research is not new. Similar concerns and hopes have been expressed with each successive innovation, beginning with CATI (computer assisted telephone interviewing) in the 1970, and including CAPI (computer assisted person interviewing) in the late 1980s and Internet surveys in the 1990s. The rhetoric has not changed much over the years — each new technological development is hailed by some as the

---

[1]Invited paper presented at the International Conference on Improving Surveys, Copenhagen, August 2002.

[2]American Association for Public Opinion Research.

next big thing. For example, in 1972 Nelson, Peyton and Bortner[3] claimed that CATI reduced costs, increased timeliness, and improved data quality, but they did not provide detailed comparative data to support these claims (see Couper and Nicholls, 1998, p. 8). In summarizing the research evidence on computer assisted interviewing (CAI), Nicholls, Baker and Martin (1997, p. 241) concluded that "Although moving a survey from P&P to CAI generally improves its data quality, these improvements have not been as large, broad, or well documented as the early proponents of CAI methods anticipated." But, as they also note, "None of the dire predictions made about CATI and CAPI when they were first introduced proved true after they passed from feasibility testing to production" (Nicholls, Baker, and Martin, 1997, p. 241).

In other words, the reality is often neither as wonderful as the proponents of the technologies argue, nor as dire as the major detractors fear. Each new technology enhances and extends the range of possibilities and opportunities for survey research, but also often introduces new challenges and issues for further research. Technology by itself is not inherently good or bad. It is how technology is harnessed in the service of human endeavor that determines its effect. To the extent that we effectively exploit the advantages that technology may bring, we can potentially improve the quality of survey data, or at least reduce the costs or time involved in producing such data. An example of this is audio-CASI (computer assisted self-interviewing) — a true technological innovation that has shown demonstrable gains in the quality of data collected on a variety of sensitive topics (see, e.g., Turner et al., 1998).

On the other hand, we as survey researchers are not the only ones exploiting each new technological breakthrough. To the extent that we or others misuse the technology, we can harm not only our own survey endeavors, but also make life harder for others too. One example is the rapid rise of telemarketing in the U.S. and elsewhere that is commonly blamed for the recent declines in telephone survey response rates. The survey profession is not immune from the nefarious uses of the technology for related purposes — that is reaching large numbers of potential customers/respondents cheaply and quickly. A similar scenario with spam could have a damaging effect on the nascent Web survey sector (as noted by Beniger above).

With this perspective as point of departure, this paper sets out to review several recent technological developments and their potential impact on survey research. This review is by no means exhaustive, and is of necessity speculative, as many of the developments are not yet fully matured. One thing is sure in the area of computing and technology, and that is change will continue apace, making it dangerous to offer predictions beyond a few years hence.

2. Technology-Related Trends in Survey Research

This paper focuses on the following interrelated and overlapping trends in survey research, inspired by or made possible by the introduction of new technology: (1) the move from interviewer-administration to self-administration, (2) the move from verbal (written or spoken) inputs and outputs to visual and haptic[4]/sensorimotor inputs and outputs, (3) the move from fixed

[3] Given the title of their paper, "Use of an On-Line Interactive System," they could well have been writing about Web surveys rather than CATI.

[4] Haptics refer to the sense of touch.

to mobile computing, both for data collectors and for respondents, and (4) the move from discrete surveys to continuous measurement. Each of these technology-driven trends is discussed in turn below.

2.1. Interviewer-administration to self-administration

There appear to be two key drivers behind the move from mostly interviewer-administered survey data collection, to an increase in self-administered components or completely self-administered surveys. The first is the steadily increasing cost of interviewer-administered data collection, due in no small measure to the increased time interviewers need to locate, screen, persuade, and interview respondents. The second is a body of research evidence — particularly from self-administered components of face-to-face interviews (computer assisted self-interviewing or CASI), but also from interactive voice response (IVR) or telephone audio-CASI — that self-administration improves the reporting of socially sensitive information relative to interviewer administration.

Given this, we have seen the rapid adoption of technologies to facilitate the automation of self-administered surveys, including CASI and its variants (e.g., audio-CASI), IVR, and the World Wide Web. These technologies offer the control and complexity of computer assisted interviewing (CAI) such as branching, edits, randomization, etc., together with the benefits of self-administration (reduction of interviewer effects, reduced costs).

This trend is likely to continue, with efforts to further reduce the costs associated with developing and deploying such self-administered surveys, and to counteract problems of coverage and nonresponse.

2.2. Audio-visual advancements

The development of graphic user interfaces (such as Windows and the World Wide Web) and the subsequent rise in multimedia computing, has led to several major innovations in survey data collection, and this trend is also likely to continue in the near future. Given the importance of these developments, we discuss auditory (primarily voice) and visual communication in turn.

Auditory communication. The use of digital sound (and in particular, recorded voice) as *output* — i.e., for the presentation of survey questions — has already had a marked effect on survey data collection. Even before the widespread availability of multimedia-capable computers, audio-CASI was being explored as a tool to enhance computer self-administration while reducing the literacy requirements for responding to survey questions. The first audio-CASI systems were developed on a Macintosh computer (Johnston and Walton, 1995) and using a separate sound device attached to a DOS-based laptop computer (O'Reilly et al., 1994).

In addition to audio-CASI, the use of digitally recorded voice for the delivery of survey questions is widely used over the telephone, in systems variously called interactive voice response (IVR; Tourangeau, Steiger and Wilson, 2002), telephone audio-CASI (Cooley et al., 2000; Gribble et al., 2000), or even voice recognition entry (VRE; Clayton and Winter, 1992) or automatic speech recognition (ASR; Blyth, 1997). Despite what is implied by their names, the latter two approaches seem to be dominated by voice output rather than input. We are even seeing survey applications using voice output on the Web (e.g., www.talksaudio.com).

To date, almost all of these applications have involved the (digital or analog) recording of a live human's voice. However, such recording is time-consuming and expensive. Variations in voice quality over different days or times of day, or variations in the recording setup, associated with the need to re-record every change or correction to the instrument and every variation of customized text, makes this an inefficient process. One consequence of this is that a single voice is typically used for audio-CASI or IVR applications — meaning that one has to choose a gender and other characteristics for that voice, and all respondents hear the same voice.

Recent advances in text-to-speech systems (TTS) have the potential to change all this, making the production of sound files for audio-CASI or IVR much less of a drudgery. Such systems are already making major inroads in a variety of customer service applications, and customer acceptance will further encourage their use. But thus far, there has been little use of text-to-speech for survey data collection applications such as audio-CASI or IVR.

With funding from the National Institutes of Health, we (Couper, Singer, and Tourangeau, in preparation) have just completed a study exploring the feasibility of TTS for survey applications. Specifically, we compared two different versions of text to speech systems to digital recording of real interviews in an IVR system. We contrasted a more human-sounding TTS system from AT&T (http://www.naturalvoices.att.com/demos/) with a more mechanical-sounding system from Bell Labs (http://www.bell-labs.com/project/tts/voices.html), using both male and female voices, and contrasting these in turn with live interviewers (CATI) and recorded interviewers. We are looking both at respondents' answers to a variety of socially sensitive questions and at their reactions to each interviewing approach. Preliminary results suggest that while respondents can clearly distinguish between human and computer-generated voices, neither the break-off rates nor the answers to sensitive questions appear to be affected.

As TTS systems are becoming less "mechanical" and sounding more like natural speech, we are even seeing the introduction of "personality" to such systems. For example, General Magic (http://www.genmagic.com/demos) offers several different TTS personalities, ranging from Greg (confident, outgoing, friendly), to Jim (authoritative, formal, self-assured and friendly) Mary (helpful, articulate, confident), and Zach (selfish, sarcastic and sometimes a troublemaker), among others.

If respondents react positively to such automated speech generation for surveys, this approach can have major cost implications for systems using voice output (such as audio-CASI and IVR). Furthermore, this will allow easy customization of voices to match respondent characteristics — such as gender or age — were this to be found desirable from a measurement perspective. In addition, other voice qualities such as pace, pitch, or tone, can be better controlled using TTS.

The use of voice as *input* for automated or computerized surveys has lagged behind that of voice output, but again this is an area of rapid technological advance that has many promising applications for survey research.

One of the drawbacks of large-scale surveys is their reliance on quantitative data (i.e., questions with closed-ended response options). In translating respondent's verbal responses into check marks on a questionnaire or having the interviewer paraphrase their comments into a fixed-length field, we have lost the rich nuances of respondents' answers. We capture their choice of a particular response option without the accompanying clarifications or qualifications that give meaning to the choices made. We have sacrificed this rich detail in exchange for efficiency in large-scale sample surveys, relegating such information to the domain of small-scale qualitative studies. There have, as always, been a few exceptions — notably the traditions of taping structured interviews for the purposes of behavior coding (Cannell, Lawson, and Hausser, 1975), conversational analysis (Suchman and Jordan, 1990; Maynard et al., 2002) or other methodological inquiries into the quality of the data.

However, the ability to digitally record a respondent's answers as a routine aspect of a computer assisted interview (whether CATI or CAPI, or their self-administered variants) produces potential efficiency gains that may make the use of such an approach a more routine part of structured interviews (c.f., Biemer et al., 2001). Digital recording offers several advantages over analog. First, there is no need for a separate device and accessories (tape recorder, cassette tapes, etc.). The sound is recorded by the computer directly onto the hard drive. Second, the system can be controlled by the interviewing software — there is no need for the interviewer to remember to switch the recording on or off. Third, the system is less obtrusive than an external tape recorder, potentially eliciting more candid answers from the respondent. Fourth, given that the system can be switched on and off at will, it is no longer necessary to record the entire interview (although this is certainly possible). In this way, one could restrict the recording to open-ended questions or a subset of key items. The technology now allows us to implement what Howard Schuman (1966) suggested many years ago, namely the random probe. Not only can CAI software be used to identify items for further probing among a random subset of respondents, but their responses can also be digitally recorded for later in-depth analysis. Fifth, because the recordings are digital, the storage and manipulation of a large number of sound files is a much less daunting task than when the information was stored on cassette tapes. Finding and analyzing the appropriate segments becomes a trivial task. The interviewing software can be used to stamp the audio files with the question identifier and any other relevant information. And, finally, speech recognition software can be used to transcribe the information, permitting the use of qualitative analysis software to code or extract themes from the information.

One can imagine a day when the capture of such verbal inputs would allow one to analyze not only the selected response to a particular question, but also the certainty with which the respondent holds that view — not only from response latency measures, as people are doing now (e.g., Bassili, 2000), but also from the verbal qualifiers used in responding, or even extracted from other nonverbal qualities of the vocal response (e.g., verbal disfluencies). In fact, Conrad and Schober (1999) are already exploring the use of verbal disfluencies to detect the need for help and offer appropriate intervention in a text-based self-administered survey setting, and one can easily imagine this being extended to spoken input.

With the ease with which digital recording can be implemented, we are likely to see a revival and extension of techniques such as behavior coding and conversation analysis, both for pretesting and for the in-depth analysis of responses in large-scale surveys.

The use of voice input may also be extended to self-administered methods. One drawback of computerized methods is that they force a respondent to pick from among a limited set of response options, and typically provide little opportunity for a respondent to clarify or qualify a response. We are thus led to believe that the responses provided are held with much more certainty than is justified. In a debriefing questionnaire following interviewer-administered and self-administered (using video-CASI) surveys on a variety of race-related questions, Krysan and Couper (2002) found that respondents complained of their inability to explain their choice of a particular response option in the CASI version. This constraint could counter the possible benefits of self-administration in reducing socially desirable responding. This is supported by evidence from laboratory studies (e.g., Moon, 1998) that voice input may lead to less socially desirable responding than keyed responses.

Visual communication. For the first century of their existence, surveys have primarily relied on words to elicit information from respondents. Words — whether presented visually (e.g., as written words in a mail or self-administered questionnaire) or aurally (e.g., as written words read aloud by an interviewer) — were the primary medium of survey questioning — and responding, for that matter.

As always, there are exceptions. These include show cards for ad testing or readership surveys, pill cards, as well as visual scales such as the Faces scale (Wong and Baker, 1988) or Oucher scale (Beyer, Denyes, and Villaruel, 1992) used in clinical settings for children or low literacy populations. There are several other examples of the use of images in survey instruments, but for the most part these were limited to specialized tasks or used for only a subset of items. In part this was because of the high cost of developing and reproducing the materials.

With digital imaging technology, we have the power to change all that. As we enter the second century of surveys, technology gives us the wherewithal to extend survey measurement well beyond the use of words to include a wide range of visual stimulus material. The inclusion of full-color images and photographs is a technically trivial and relatively inexpensive undertaking. This is true both in graphical user interfaces (GUIs) for CASI and in surveys on the Web.

We are already seeing many examples of this, and they are beginning to make their way into mainstream survey research. For example, a variety of scenic preference surveys have been conducted online (see e.g., Wherrett, 1999, 2000), using unretouched photographs, digitally manipulated images (to alter key elements of the landscape) or computer-generated representations of various scenic attributes. Joan Nassauer (see http://www.snre.umich.edu/nassauer/) has made similar use of digital images for a survey of home landscape design. David Harris (2002) recently completed a Web survey of University of Michigan students in which he looked how people classified mixed-race persons. To do this he created a series of digital pictures that were "morphed" composites of real people. The Implicit Attitudes Test (Nosek, Banaji, and Greenwald, 2002) also uses digitally generated human faces. Digital images are already widely used in a variety of other Web surveys on topics ranging from ad testing to packaging and new product or concept testing, magazine readership, and so on.

Along with this increasing interest in the use of images and other visual materials in surveys, comes the increasing appreciation of the potential effects of visual presentation on survey responses. Despite decades of survey research using self-administered questionnaires, relatively little research has focused on the nonverbal elements of the questionnaires. Recently, Redline and Dillman (2002; see also Jenkins and Dillman, 1997) have been examining the effects of a variety of visual representations (especially for skip patterns) in paper-based surveys. We have been doing similar work on the web (e.g., Couper, 2001c; Couper, Traugott, and Lamias, 2001). The lesson of this research is that visual design is a powerful tool for facilitating the task of survey completion, but can also affect the answers provided if used inappropriately.

These developments have the potential for dramatic change to survey measurement. Survey measurement has in large part been constrained by our ability to craft well-worded questions to convey increasingly complex ideas and elicit responses on a wide range of topics.

By extending measurement to pictures (and drawings, and other visual representations), we can greatly expand the range of measurement possibilities in surveys. On the other hand, there is much we still need to learn, for example about what kind of information images convey and what effect they have on measurement (e.g., Couper, Kenyon, and Tourangeau, 2001).

Audio-visual (multimedia) communication. Thus far we have discussed audio and visual modalities separately. In multimedia applications these come together, and here again we are already seeing several interesting developments in survey measurement. Two brief examples — both using video-CASI applications — will suffice. In a survey in the Detroit area conducted prior to the 2000 U.S. presidential election, Hutchings, Valentino and colleagues (Hutchings, Valentino, and Rusch, 2001; Valentino, Traugott, and Hutchings, 2001) created several versions of a television campaign commercial for a candidate, varying the race priming across versions. These 30-second videos were randomly assigned to respondents in a self-administered part of the survey, embedded within two other filler commercials. In similar fashion, Maria Krysan and I (Krysan and Couper, 2001) are exploring race of interviewer effects using what we called "virtual interviewers" — digital videotapes of real interviewers reading survey questions, displayed on a laptop computer screen.

We see video-CASI as a natural extension of audio-CASI, permitting a wide range of stimulus material to be presented to respondents in a relatively controlled setting. Use of such approaches is likely to increase in the future. Multimedia applications are currently limited on the Web but, as bandwidth increases, we are also likely to see increasing use of multimedia in self-administered surveys on the Web.

Here I briefly mention another related development — the increasing interactivity of self-administered surveys. Computerized survey instruments, whether on the Web or on a standalone computer, are not passive like paper-based instruments. They can react to user input. This feature is being exploited in a number of ways, both to develop new methods of *input* (answering survey questions) but also in providing customized feedback to respondents. This involves not only the presentation of a visual stimulus, but also allowing the respondent to interact directly with the stimulus, for example by pointing to, clicking on, or otherwise manipulating the image. For example, an experiment on facial beauty by Victor Johnson (http://www-psych.nmsu.edu/~vic/faceprints/) has respondents gradually change the gender of the facial image displayed until the desired level of attractiveness is reached. Direct manipulation input devices also include slider bars, drag-and-drop, or other tools for indicating a response. In addition, the interactive nature of computer administration can be used to provide a variety of visual and auditory feedback (e.g., online lookup, running totals) to aide respondents in surveys involving complex decision-making or other difficult tasks. Such use of haptic (touch) interfaces is already possible on the Web and other graphical environments — and we're not even talking about virtual reality, immersion systems, and other exotic interfaces already under development (Briggs, 2002).

2.3. Mobility

The third broad technology-related trend is the move from fixed to mobile computing, both for data collectors/interviewers, and for respondents. Again, each of these are discussed in turn.

Mobility for interviewers/data collectors. Some dozen years after the introduction of

CAPI, interviewers conducting face-to-face surveys still lug 5-7 pound (about 2-3 kg) laptops around. Several researchers have explored the use of handheld computers (e.g., Bosley, Conrad, and Uglow, 1998; Nusser, Thompson, and Delozier, 1996) but their power, storage capacity, and screen size have limited them to niche applications, such as in-store pricing surveys or short doorstep interviews (see also Gray, 1999; De Heer, 1991; Nusser and Fox, 2002). Personal digital assistants (PDAs) have been used for household screening in several large-scale surveys in the U.S. (most notable the National Household Survey on Drug Abuse and the National Survey of Family Growth in the U.S.), but the interviewers also carry laptops for conducting the actual interviews.

Despite major advances in the field of mobile computing, the hardware used for CAPI surveys has not changed much since when first introduced in the late-1980s. To be sure, we have seen major gains in computing power and storage capacity, and we have seen the evolution from DOS to Windows, but the basic form-factor (i.e., the look and feel) of the CAPI device is largely unchanged.

For many years we have been waiting for a machine that is powerful enough to conduct long, complex interviews, yet portable enough that it can be used for screening, short doorstep interviews, exit surveys, and so on. As long ago as the mid-1980s (before the widespread introduction of laptop computers), Statistics Sweden identified the ideal characteristics of such a machine (see Lyberg, 1985). Robert Groves and I did similar research in the early 1990s (Couper and Groves, 1992). After many promising developments and false starts, we are still waiting for the ideal field interviewing computer. Several machines have been introduced on the market, but have not lasted long. One reason has been the quality of the handwriting recognition on keyboardless machines. Recent developments in this area are encouraging.

Several factors — including improvements in handwriting recognition, storage capacity (obviating the need to recognize the handwriting at the time of entry; the digital image can simply be stored for later manipulation), the development of alternative input modes for open-ended responses (e.g., digital audio recording), and the development of touchscreen audio-CASI systems (Schneider and Edwards, 2000; Cooley et al., 2001) — are all converging to point to the likely success of the new wave of tablet machines for survey interviewing applications. The fact that the software juggernaut, Microsoft, is fully behind these new developments with the recent introduction of its tablet-based version of Windows XP doesn't hurt. Several manufacturers are already on board to produce the new machines.

Nonetheless, for the same reasons that Statistics Sweden's requirements for a handheld machine were not taken up by any manufacturer, it is not the survey industry that will determine the success or survival of this new class of machines. Without a "killer app" that results in millions of such machines being purchased for other uses, the new tablets will go the way of other promising but commercially unsuccessful machines such as the GridPad and Apple Newton.

Field data collectors are becoming more mobile in other respects too. With the advent of what is called ubiquitous computing or wearable computers, the devices that an interviewer can carry, and the kinds of data they can collect, are both expanding. One example is the use of portable global positioning systems (GPS), permitting the collection of location and spatial data as part of the survey process. Sarah Nusser and colleagues (see http://dg.statlab.iastate.edu/dg/) are exploring a range of survey applications, including both field measurement in land-use studies and more traditional survey data collection such as in-store pricing or doorstep surveys.

A variety of other portable devices and add-ons to handheld computers — digital cameras (video and still), bar code readers, and so on — are further expanding the view of what constitutes survey "data."

Mobility for respondents. The most obvious of these developments is the rise of the mobile or cellular phone as a communication device, and its potential impact on survey research. There are others much more qualified to address this topic. However, I offer a few brief observations on issues relating to the increase in mobile phone use and its implications for survey research.

Mobile phones present a number of challenges with respect to sampling and coverage. One key issue is the move from the household to the person as a sampling unit. Fixed lines are more likely to be associated with a household, while mobile phones are mostly linked to individuals. This will remain an issue during the transition from fixed to mobile telephony, and has important implications for sample design and coverage. Already in some countries this is less of a problem. For example, mobile phone coverage has already exceeded fixed-line coverage in Finland (Kuusela, 2002), and not only is this trend likely to continue in Finland, but also several other European countries are not far beyond. However, mobile phone coverage is still at a level in most countries that — for the immediate future at least — telephone surveys will be conducted with both fixed and mobile phones. This will require dual frame designs (where separate area codes are assigned to each system) or the need for unduplication (where separate frames are not available). Another coverage concern is that the definition of access or use is trickier with mobile phones (as with Internet access/use) than fixed lines, given the relative impermanence of mobile telephone numbers (as with e-mail addresses).

On the other hand, the cost of these devices is reaching a point where coverage problems could be overcome by providing every sample person with the requisite equipment (e.g., disposable phones). This obviously requires some other mode for initial selection, but for ongoing panels, this investment may pay off. Knowledge Networks in the U.S. and CentERdata in the Netherlands are already doing this for Web surveys. But one could imagine other applications such as tracking studies following an initial face-to-face survey, where the respondent is provided with a mobile Internet device (e.g., Blackberry), personal digital assistant (PDA) or mobile phone for follow-up data collection.

The change in personal telephony has many potential implications for nonresponse too, and this may affect all types of telephone surveys. Given the wide variety of uses to which fixed lines are put (modem, fax machine, etc.) and the increase in screening devices (caller ID, answering machines), along with the fact that mobile phones are not always switched on, the odds of reaching a live human being on any particular call attempt are likely to continue to decrease. This means that making contact with a sample person will be an ever-increasing component of the data collection costs. There are technologies available to assist in this process — predictive dialers, auto-dialers, and the like — but these same technologies are used by telemarketers, and potential respondents have become savvy at recognizing the delay after picking up the phone as a likely marketing call, and quickly hanging up. This is an example of how potentially useful technology can be usurped by others, making it harder to use for survey data collection.

With regard to nonresponse, the very mobility of cellular phones is a double-edged sword for survey researchers. While their portability may increase the chances of making contact with someone (relative to a fixed line), the likelihood of them being willing to do a telephone survey at that time may decrease. How respondents react when called while in transit, or in the middle of a wide variety of other activities, is an important area of research. This also has important implications for measurement error (see Fuchs, 2002).

Another nonresponse issue relates to the cost structure of mobile phone systems. For example, in the U.S., users pay for incoming as well as outgoing calls on cell phones. This is likely to affect the willingness of respondents to do lengthy surveys on a mobile phone. In several other countries, there are highly variable cost structures, given the variety of different providers. The cost of a call varies depending on whether one is calling within a calling area or outside, calling from a fixed line to a mobile phone or vice versa, calling to the same provider or to a different provider, and so on (Vehovar, 2002). This may affect not only the likelihood of successfully reaching a potential respondent and gaining their cooperation, but also the efficiency of telephone surveys. On the other hand, there are potential benefits in terms of delivery of incentives — for example, electronic transfer of money, or adding minutes or money to a cell phone account in exchange for participation, especially for pay-as-you-go mobile phone plans.

Finally, a mobile phone is a very different device from a traditional landline telephone. Indeed, the very notion of what constitutes a mobile phone is becoming murky. With the ability to send text messages, exchange pictures, connect to the Internet, and so on, the mobile phone is becoming much more than a human-to-human speech communication device. In addition, other mobile internet devices (e.g., PDAs) are adding voice (telephone) capabilities, further blurring the line between devices. With the widespread use of text messaging (e.g., short message service or SMS), the telephone is moving from a synchronous to asynchronous communication[5]. SMS and voice-mail may have more in common with e-mail than with traditional telephone communication. While SMS is currently limited to only 160 characters, it is device independent and asynchronous (i.e., not dependent on making direct two-way contact with the respondent). This may make it useful for short frequent data collection activities, or for inviting respondents to complete surveys at their convenience (Callegaro, 2002). This is an area ripe for methodological research.

One thing is clear in the world of mobile phones, and that is there is a wide variety of different plans, systems, allocation of area codes and numbers, cost structures, coverage, cultures of use, etc., within and between countries. This is thus an area that will greatly benefit from cross-national research, as we explore how different models and structures affect survey errors and data collection strategies. In sum, then, the growth of mobile phones represent several challenges for survey research, but also opportunities for exploring and developing alternative methods.

2.4. Continuous measurement

As already noted, many of these technological developments are related. The growth of

---

[5]Synchronous (or real-time) communication takes place simultaneously. Examples include chat rooms, instant messaging (IM) or telephone conversations. Asynchronous (turn-based or delayed) communication takes place at different times. Examples include e-mail and SMS.

portable Internet devices and mobile computing, like that of mobile phones, permits the development and extension of continuous measurement in surveys.

In traditional interviewer-administered survey methods, the costs of sampling, contacting, persuading, and interviewing respondents is so high that it makes sense to maximize the opportunity to collect large amounts of data (often several hours of questions) at a single point in time, or, if a panel is used, at relatively lengthy intervals (often 1-2 years apart). Continuous measurement using self-administered methods — beepers, mobile phones, Blackberries, etc. — changes the cost equation.

A relatively large initial investment may be required to enroll sample persons in a panel and provide them with the required equipment. Thereafter, using self-administered methods, with automated prompts, survey instruments, and reminders, small amounts of data can be collected at much more frequent intervals. This approach may reduce respondent burden, by spreading the load over a large number of shorter interactions. It may also improve data quality, by reducing the length of the recall period, and collecting information much closer to the time of occurrence. These methods hold much promise for the study of relatively frequent and recurring behaviors (e.g., alcohol consumption, diet and exercise, mood, interpersonal interaction, and so on). But these benefits may come at a cost in terms of other sources of error, including coverage, nonresponse and panel effects.

The idea of frequent or continuous measurement is not new. Diary surveys have a long history. These methods are already used in audience measurement (TV and radio), often using passive detection methods to further reduce burden on respondents. Inbound interactive voice response (IVR) surveys are often used for the regular reporting of a wide range of behaviors. But one weakness shared by all these approaches is that the initiative to call in or complete the diary usually rests with the respondent. The advent of affordable wireless Internet devices means that the machine itself can prompt the user to report the information. There is an immediate feedback loop that permits tracking respondent progress, checking compliance and completeness on a flow basis, and automatically delivering reminders, motivational messages, and even incentives as and when needed.

Another aspect of continuous measurement involves the use of transaction data. Computers are becoming less objects we interact with at certain times to complete certain tasks, and much more part of most everyday activities, producing digital traces of everyday life. Examples include any type of digital transaction, such as electronic payment (credit or debit cards), affinity cards (frequent shopper cards, frequent flyer cards), e-commerce, library card use, video rental, prescription cards, and so on. These are becoming so much part of everyday life that we're often unaware of the presence of these computing devices, and that a digital record of the transaction exists. Technology is thus giving us much greater access into people's lives. This offers unprecedented opportunities for researching aspects of activities that were heretofore inaccessible or prohibitively expensive to study on a large scale.

Transaction-based surveys have enabled us to study *what* people are doing and *when*. But now, in combination with other mobile devices (cellular telephone or Internet, voice-based or text-based) we can ask questions of *who* and *why* too, potentially giving us much greater insight into people's behaviors at or close to the time they occur. But herein also lies the potential danger of such technology. The more we (and others, whether for surveillance or marketing or other purposes) pry into the lives of our fellow human beings, the greater the potential backlash against such intrusions. The more we can potentially learn about people, the fewer people there will be who may be willing to give us such access into their everyday activities. This raises a number of challenging methodological issues. For example, what kinds of information are people willing to share, under what circumstances? How can data quality be assured when combining imperfect transaction-based data with the often-imperfect responses to survey questions?

3. Implications and Research Challenges

In this section I review several implications or consequences of the technological changes discussed above, along with some thoughts on how the field of survey research is changing and what we can do as a profession in response to these changes. I group these loosely into two broad areas, first focusing on the survey process and next on issues of data quality.

3.1. Implications for survey research and the survey profession

One consequence of the recent spate of technological innovations has been a proliferation of modes of survey data collection. In the early days it was simple — there were surveys and questionnaires. The former were administered face-to-face by interviewers, while the latter were self-administered. Then came the telephone, a major innovation that vastly expanded the realm of surveys. The introduction of telephone surveys was largely responsible for the growth in small-scale university-based surveys centers in the U.S. and (to a lesser extent) elsewhere. For example, O'Rourke, Sudman, and Ryan (1996), reported a fivefold increase in university and nonprofit survey organizations in the U.S. between the early 1970s and mid-1990s.

Now we have an ever-expanding array of methods — and associated acronyms. It is no longer sufficient to talk about a face-to-face survey. Were computers used or was it paper-based? Were there any self-administered components? Were these text-CASI or audio-CASI, or even video-CASI? How was the initial screening and sample selection carried out — using the same mode and device or different ones? Was the same mode used for recruitment and interviewing? What about follow-up? Even within a relatively new method, like the Web, there are numerous ways to implement the survey (see Couper, 2001a), with implications for various surveys errors and costs. This requires us to be much more precise in how we describe the methods we use to collect survey data, especially given how important these details may be more understanding the error structure of our estimates.

A related outcome of the more recent developments, particularly the Web, has been the increasing democratization of the survey data collection process. In the early days of computerization, the cost of investing in the technology restricted its use to a few large survey organizations. The first CATI systems were run on mainframe computers. The introduction of PCs saw an increase in the use of CATI by smaller survey organizations and academic research centers, often conducting local or regional studies. But there are still relatively large infrastructure costs associated with starting and maintaining a CATI facility. CAPI was similar — an early report on CAPI (Rudolph and Greenberg, 1994) estimated that an initial investment of around $1 million was needed to make the conversion from paper to CAPI. Even now, the costs associated with equipping a large field force with laptops, along with the associated software and support systems, means that CAPI remains the domain of a few large survey organizations.

The Web, on the other hand, has the potential to fully democratize the survey process, giving the lone researcher the power to survey large numbers of potential respondents cheaply and quickly. But in doing so, the profession is losing control over the quality of the work being done.

Furthermore, with the expanding range of measurements and the ability to collect data anywhere at any time comes an increasing concern about over-surveying or stiff competition from those using the same technologies for other purposes. Telemarketing is having a serious effect on telephone surveys. The rapid proliferation of spam on the Internet similarly threatens to engulf Web surveys. But even if marketing didn't do it, we might be doing it to ourselves. In the days of face-to-face interviewing, the chances of being included in a survey were relatively rare. With the widespread use of the telephone dramatically decreasing the cost of surveys, the number of surveys have gone up. The Internet further expands this trend. *Anyone* with an Internet connection can attempt to survey thousands or even millions of potential respondents at the cost of a little effort and a few dollars. And some days it seems that *everyone* is doing so. A Web search for the terms "survey" or "poll" will reveal the staggering number of such instruments available on the Web. Virtually every large media website and most Internet service providers (ISPs) or portals have a question of the day poll on their site. For example, the question of the day on CNN's website (www.cnn.com) routinely gets 50,000 or more responses a day. A visit to money4surveys.com will reveal the list of dozens of Web survey panels one can join to participate in countless other surveys.

To be sure, good research still takes time and effort (and therefore money), but how many potential respondents know the difference between good and bad, or are willing to invest the time and effort to make the distinction? Not only are the sheer number of surveys increasing, we are also seeing increasing diversity in the quality of surveys being conducted. In this way, bad could drive out good. Now more than ever we need an educated public and clients, who understand the difference between shoddy work cheaply and quickly done and high quality surveys that may take extra time and effort to conduct, and consequently cost more.

Along with the proliferation of new methods comes an increased specialization. In the early days of surveys, one person could know (and do) it all. Now many surveys (except for the simplest of ones) require a range of different specialized skills. It has become even more critical that people in each specialism communicate with each other and understand enough about the entire survey process to make such communication effective.

In a related way, the nature of survey statistics is also changing. For many decades, the basic object of analysis was one or more rectangular data files containing a set of variables for each of several cases. This is now expanding to include not only complex hierarchies (e.g., doctor visits, persons, households) but also spatial data, transaction data and less-structured observational data. The data survey researchers have to analyze are becoming much richer, but also correspondingly more complex. This presents new challenges and opportunities for survey analysis, as well as for the design and conduct of surveys.

Yet another consequence of the technological trends discussed above is that the distinction between surveys and experiments is becoming increasingly blurred. Given the ability to randomize both within and across questions, to control the presentation of stimulus material, to deliver rich multimedia content, and to record a range of possible reactions (e.g., verbal responses, response latencies), experiments are increasingly being conducted using large and diverse samples of respondents (often via the Web). Similarly, surveys are increasingly including experiments as an integral part of data collection.

When an online experiment such as the Implicit Attitudes Test (Nosek, Banaji, and Grreenwald, 2002) can attract over 1.5 million participants over the course of a few years, experiments are no longer constrained to small samples of homogenous populations (college sophomores). While it is not a probability-based survey permitting generalization to a known population, the sheer size of such an undertaking has the potential to transform the research into something new and different.

Similarly, David Spiegel and colleagues at Stanford University posted an unrestricted survey on post-traumatic stress disorder (PTSD) and related issues on the Web shortly after the events of September 11, 2001. They garnered over 10,000 responses in the space of several weeks, and approximately 70% of these respondents completed a follow-up survey six months later. Again, while not a probability sample, the data from such a study are uniquely valuable, and understanding the strengths and weaknesses of these large-scale survey-like efforts is worth of research attention. On the other hand, we should not be blinded by the number or variety of respondents from all around the world that can be obtained via the Web, but we should strive to learn how these new technologies can be used to complement both survey and experimental methods.

Given all of the above, the notion of what constitutes a survey — and even more importantly, what constitutes a *good* survey — is becoming increasingly important. The total survey error framework (see, e.g., Groves, 1989) is a valuable tool for understanding the error properties of surveys, and applies to newer technologies and approaches as well as to old.

3.2. Implications for survey errors and data quality

I've already mentioned the progression from voice and visual extensions of survey data collection to multimedia applications to virtual reality. With the rise in "natural" interfaces and smart computers, the division between man and machine is breaking down — it's becoming less obvious what is human and what is computer. The well-knowing Turing test for artificial intelligence[6], by which one can distinguish a human from a computer actor, is under threat. For example, with developments in speech systems (both input and output) for telephony, it's sometimes unclear whether we are interacting with a live human or a computer. Similarly, using digital imagery, it is possible to create or morph "people" that are not real, and imbue them with rich personalities in artificial worlds or virtual reality. Already we are seeing software where, with a few photos of a person's head from different angles and the recording of a few sentences of spoken word, one can produce a fully animated three-dimensional talking head with

---

[6] In 1950 Turing wrote: "It is proposed that a machine may be deemed intelligent, if it can act in such a manner, that a human cannot distinguish the machine from another human merely by asking questions in a mechanical link." If we replace "asking questions" with "answering questions" and consider mechanical link (remember, this was written in the era of mainframes and vacuum tubes) to include a variety of digital input or output devices, this applies fully to automated surveys.

computer-generated speech that can look and sound remarkably like the real thing. Such avatars[7] are becoming more and more human, further blurring the boundary between what is real and what is computer-generated.

Given this development, one of the challenges is figuring out which survey tasks are best for human interviewers to perform (e.g., motivating, persuading, locating, etc.) and which are best for computers (e.g., processing, calculating, standardizing, recording, transmitting, etc.). Asking questions in a standard way may be the ideal domain for computers, as the trend towards CASI is already showing. If we want to remove — or control for — interviewer influences on measurement, automation may be the way to go. But we need to be careful about going too far. Artificial intelligence has its limits. For one, it is just that — artificial. We also need to remember that an interviewer's task involves much more than the standardized delivery of survey questions and recording of responses. To what extent will new survey technologies replace interviewers, versus supplementing or supporting interviewers in their tasks?

Another issue arises with the increasing trend towards rich interfaces for self-administered surveys, namely the possible reintroduction of social desirability effects. While computerized self-administration of surveys is argued to reduce (or eliminate) social desirability effects, other research suggests that the introduction of humanizing cues on a computer (such as voice, images, etc.) may in fact *increase* socially desirable responding (e.g., Nass, Moon, and Green, 1997; Nass, Moon, and Carney, 1999; Kiesler, Sieff, and Geary, 1992). We have been engaged in a program of research to explore this issue for some time, because it has important implications for survey research. Thus far, our research on the Web (Tourangeau, Couper, and Steiger, 2002), interactive voice response (IVR) and audio-CASI (Couper, Singer, and Tourangeau, 2002) suggests that social presence — the addition of humanizing features to an automated survey interface — has little effect on the answers people give to sensitive questions. We are exploring this issue further, but if these findings hold, it suggest a range of additional cues that can be added to computerized self-administered instruments (whether aurally or visually delivered) to support, encourage, motivate and otherwise assist respondents in providing accurate and honest responses in a private setting.

On the other hand, using digital videos of interviewers in a video-CASI application, we have been able to replicate some race-of-interviewer effects found in surveys conducted by live interviewers (Krysan and Couper, 2001). This suggests that we need to remain vigilant that, as we move towards replacing live interviewers with rich virtual equivalents, we are not reintroducing some of the very effects we hoped to eliminate with the move to self-administration.

Another issue that will increasingly face the survey industry relates to privacy and confidentiality. Technology is giving us much greater access into people's lives. This offers unprecedented opportunities for researching aspects of people's lives that were heretofore inaccessible or prohibitively expensive to study on a large scale. In 1998, Baker wrote "We are about to enter an age in which the transactions of our daily lives will be entirely digital" (p. 597). Beniger (1998) further suggested that a time would come where people would willingly make

---

[7]Avatars are computer-generated characters. They may be two- or three-dimensional, they may be still or animated, they may be real or fantasy characters. MTV's Max Headroom is an early example of an avatar. A more recent Web-based example is Ananova (www.ananova.com).

such digital information based on everyday transactions available to researchers, possibly at a price. For many purposes, especially in marketing, advertizing, product testing, and other fields, such self-selected samples are quite sufficient. But there are many attitudes, behaviors, intentions, attributes, etc., which people would be unwilling to have as a matter of public record. These are many of the things we as survey researchers or social scientists are interested in studying, precisely because this information is hard to obtain in indirect ways. People may be willing to put their "normal" or everyday lives on display for the researcher, but it is the deeper, often hidden selves, that are of most interest to us — sexual behavior, drug use, other risky or illicit behaviors, prejudice, welfare cheating, purchase intentions (as opposed to past purchasing behavior), and the like.

The increasing ability to pry into the intimate corners of people's lives may produce a backlash that limits such opportunity. As Rawlins (cited in Baker, 1998, p. 601) noted: "Today's encryption technology could, if used widely enough, make us the last generation ever to fear for our privacy ... if misused, it could make us the last generation with any notion of personal privacy at all." In a recent (July 25th, 2002) *New York Times* article on the amount of personal information available online, Jennifer Lee wrote: "The Internet, which was supposed to usher in an era of limitless information, is leading some people to restrict the information that they make available about themselves."

While we're making great strides in extending the measurement capabilities of surveys, allowing us to explore new realms of inquiry and improve the quality of existing measures, the new technologies have thus far brought few discernible benefits in terms of reducing errors of nonobservation.

In terms of *sampling*, Web surveys are at the vanguard of an apparent trend away from probability-based samples. A large number of Web surveys are already based on self-selected samples. First, there is a select group (among those with Internet access) who have expressed a willingness to participate in such surveys and have signed up for one of the numerous opt-in Web panels developed in recent years. Second, even among these volunteers, the proportion who respond to any one survey is very low (typically less than 20%; see Couper, 2001b). Furthermore, this decision is often made on the basis on the topic or content of the survey, as revealed in the e-mail invitation. Despite the highly selective nature of such participation, researchers are making inference from estimates derived from these surveys to the Internet population, and even to the full population.

In the area of *nonresponse*, we can expect to gain little from the introduction of new technology. With some exceptions (and with certain populations, e.g., college students), mail surveys get higher response rates than equivalent Web surveys (Couper, 2001b), and, when given a choice of mode, respondents overwhelmingly choose mail. The early claims that CATI and CAPI would result in improvements in response rates, or at least stem the decline, have proved unfounded. But fortunately, so too have the fears that the introduction of these technologies would hasten the decline in willingness to participate in surveys. The evidence from CATI and CAPI suggests that the technology itself has played little part in declining response rates in interviewer-administered surveys. While CASI has had demonstrable effects on reducing measurement error, its impact on nonresponse has been negligible.

Given the trend toward opt-in surveys — to use the language of the Web — more and more sample surveys are studies of cooperative volunteers, rather than a full cross-section of the population of interest. Poynter (2000) predicts that by 2005, 80% of research in Europe will be conducted with opt-in lists rather than probability samples, and 95% of this work will be conducted via the Internet. But how representative are these samples? Sure, they are large and diverse, but can we assume that those who are accessible and willing to respond are identical to the balance of the population of interest on the key statistics of concern? Despite some evidence that propensity-adjusted samples of Web volunteers produce some estimates that resemble those from other modes (e.g., RDD telephone surveys; Taylor et al., 2001), there are also examples where the differences are large (Schonlau et al., 2000; Robinson, Neustadtl, and Kestnbaum, 2002). Similarly, while recent evidence suggests that the impact of decreasing response rates on survey estimates is not large (Keeter et al, 2000; Curtin, Presser, and Singer, 2000), this remains an area of great concern — and one in need of much further research — to those using surveys for inference to broader populations.

In concluding their review of the extant literature on survey technology and data quality, Nicholls, Baker, and Martin (1997, p. 242) noted that "...computer assisted methods do not provide a panacea (or even a general palliative) for survey noncoverage, nonresponse, and measurement error." The same is likely true of the newer methods of survey data collection. With the countervailing forces at work, each potential gain in terms of cost reductions, timeliness, or improved measurement may be offset by possible losses with respect to sampling, coverage, or nonresponse.

4. Conclusions

I began this paper with a question on how one views technology and it's impact on survey data collection. I return to this issue in concluding the paper. The key issue, as I have noted, is not whether technology is inherently positive or negative, but how it is used.

There are many different views on the uses of technology in the service of improved survey data collection quality and reduced costs. How successfully new technologies are adopted and applied to survey data collection may well depend on how we view the role of technology. Here are just some of the ways technology can be viewed:

- Technology to *replace*: How can we emulate the valuable features of interviewers in an automated environment? How can we save money by using technology instead of interviewers?
- Technology to *support*: How can we use technology as a support tool to facilitate or enhance the work of interviewers?
- Technology to *enable*: How can we use the technology to do things that weren't previously possible? What new forms of data collection, new types of measurement, and new ways to combine and utilize different types of data can be developed using technology? How can we extend the existing capabilities of surveys?
- Technology for *efficiency*: With the increasing complexity of automated interviewing systems and instruments, surveys (with few exceptions) are becoming more expensive, not less so — how can we harness the power of technology to make survey data collection more cost-effective?
- Technology to *control*: How can we use technology to exercise greater control over the process, e.g., through enforced standardization, or through closer monitoring of

interviewers and respondents?

- Technology to *inform*: How can we use the technology to learn more about the process of data collection? Can we harness the paradata[8] generated by automated survey processes to improve our methods?
- Technology for the sake of *technology*. How can we find a way to use the latest technologies, without any other clear purpose or goal in mind?

Any or all of these views could be held at various points during the introduction and maturation of new data collection technologies. These different views will likely lead to the development of different tools for survey data collection. In evaluating both present and future innovations, we could ask what technology has brought in terms of each of these different views.

The more things change, the more they stay the same. The rhetoric surrounding the introduction of each new technological innovation is remarkably similar to that for previous such inventions. For example, Web surveys were hailed with much fanfare as the better, cheaper, faster way of collecting survey data. Similarly arguments were made some 30 years ago by the proponents of CATI, and some 10-15 years ago upon the introduction of CAPI. Neither of these previous innovations can be deemed truly revolutionary, and it is likely that the Web will not be too. The contributions of technology to survey research, although certainly not to be underestimated, are often more evolutionary than revolutionary.

What is also remarkably constant over this time is that the criteria we use to evaluate survey quality have not lost their relevance. To be sure, we have made gains in terms of the timeliness of data delivery, in terms of the measurement of sensitive behaviors, and in other areas. But the complexity of surveys has also increased, and the challenges we face with respect to sampling, coverage and nonresponse have not diminished.

Given this, how do we prepare for a future in which we are only likely to see the pace of change increase? Unfortunately, we are not fully in control of our own fate. The survey profession is not large or powerful enough to dictate the course of technological development, nor can we hope to claim monopoly use on any advancement. One key to our future as a profession lies in our success in distinguishing ourselves in the minds of the potential respondents from other similar-appearing activities using the same technology for other purposes.

With the rapid development of new technologies and the accompanying proliferation of ways to design, conduct and analyze surveys, it is more important than ever to have an educated profession and an educated public, one that can distinguish good from bad surveys, and can evaluate the quality of different survey products.

Quality is not an absolute. Quality must be weighed against other considerations, such as costs and timeliness. Good survey design always involves trade-offs, of maximizing quality for a given level of effort. The total survey error framework is useful now more than ever in evaluating the host of new methods that are being introduced. With the increasing complexity of the survey enterprise, and increasing specialization within the profession, a common language and a common understanding of the importance of quality is vital.

---

[8] Paradata are auxiliary data about the *process* of data collection, and are often generated as an automatic byproduct of automated systems. Paradata include keystroke files, time stamps, call record information, and the like.

Two important ways we can prepare for an increasingly technological future are as follows:

- Education and training. Both specialty skills and cross-cutting knowledge are needed. Surveys are not simply a sum of a number of disparate parts. Decisions made in one area (e.g., sample design, interviewer training) affect many other areas (e.g., analysis, nonresponse reduction, mode choice). Now more than ever we need trained professionals designing, conducting and evaluating surveys, and analyzing the data they produce.
- Good theory and good research. The rate of technological change is likely to outpace our ability to do sufficient research on every new method prior to adoption. We need to strike a balance between a headlong rush to adopt each new technology at once, and waiting for all the research evidence to be accumulate before making the transition.

Research and practice will of necessity go hand in hand. The era of large-scale methodological inquiries are probably behind us. On the other hand, every survey can be seen as an opportunity, not only to produce data of substantive relevance, but also to advance our knowledge of surveys. These two points portend the rise of survey methodology as a profession and an academic discipline. With a solid foundation of theory and practice, research and experience, survey research is ready to face the challenges and opportunities of the technological future.

References

Baker, R.P. (1998), "The CASIC Future." In M.P. Couper, R.P. Baker, J. Bethlehem, C.Z.F. Clark, J. Martin, W.L. Nicholls II, and J. O'Reilly (eds.), *Computer Assisted Survey Information Collection*. New York: Wiley.

Bassili, J.N. (2000), "Reflections on Response Latency Measurement in Telephone Surveys." *Political Psychology*, 21 (1): 1-6.

Beniger, J. (1998), "Survey and Market Research Confront Their Futures on the World Wide Web." *Public Opinion Quarterly*, 62 (3): 442-452.

Beyer, J., Denyes, M., and Villaruel, A. (1992), "The Creation, Validation, and Continuing Development of the Oucher: A Measure of Pain Intensity in Children." *Journal of Pediatric Nursing*, 7: 335-346.

Biemer, P., Morton, J., Herget, D., and Sand, K. (2001), "Computer Audio Recorded Interviewing (CARI): Results of Feasibility Efforts." Paper present at the annual meeting of the American Association for Public Opinion Research, Montreal, Quebec, May.

Blyth, W.G. (1997), "Developing a Speech Recognition Application for Survey Research." In L. Lyberg, P. Biemer, M. Collins, E. deLeeuw, C. Dippo, N. Schwarz, D. Trewin (eds.), *Survey Measurement and Process Quality*, New York: Wiley.

Bosley, J., Conrad, F.G., and Uglow, D. (1998), "Pen CASIC: Design and Usability." In M.P. Couper, R.P. Baker, J. Bethlehem, C.Z.F. Clark, J. Martin, W.L. Nicholls II, and J. O'Reilly (eds.), *Computer Assisted Survey Information Collection*. New York: Wiley.

Briggs, J.C. (2002), "Virtual Reality Is Getting Real: Prepare to Meet Your Clone." The Futurist, May/June: 34-41.

Callegaro, M. (2002), Remarks made at Roundtable on "Mobile Phones in Telephone Surveys." Annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.

Cannell, C.F., Lawson, S.A., and Hausser, D.L. (1975), *A Technique for Evaluating Interviewer Performance*. Ann Arbor: Institute for Social Research, University of Michigan.

Clayton, R.L. and Winter, D.L.S. (1992), "Speech Data Entry: Results of a Test of Voice Recognition for Survey Data Collection." *Journal of Official Statistics*, 8 (3): 377-388.

Conrad, F.G. and Schober, M.F. (1999). "A Conversational Approach to Text-Based Computer-administered Questionnaires." In Proceedings of the Third International ASC conference. Chesham, UK: Association for Survey Computing, pp. 91-101.

Cooley, P.C., Miller, H.G., Gribble, J.N., and Turner, C.F. (2000), "Automating Telephone Surveys: Using T-ACASI to Obtain Data on Sensitive Topics." *Computers in Human Behavior*, 16: 1-11.

Cooley, P.C., Rogers, S.M., Turner, C.F., Al-Tayyib, A.A., Willis, G., and Ganapathi, L. (2001), "Using Touch Screen Audio-CASI to Obtain Data on Sensitive Topics." *Computers in Human Behavior*, 17: 285-293.

Couper, M.P. (2001a), "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly*, 64 (4), 464-494.

Couper, M.P. (2001b), "The Promises and Perils of Web Surveys." In A. Westlake et al. (eds.), *The Challenge of the Internet*. London: Association for Survey Computing, pp. 35-56.

Couper, M.P. (2001c), "Web Surveys: The Questionnaire Design Challenge." Paper presented at

the International Statistical Institute, Seoul, Korea, August.
Couper, M.P. and Groves, R.M. (1992), "Interviewer Reactions to Alternative Hardware for Computer-Assisted Personal Interviewing." *Journal of Official Statistics*, 8 (2): 201-210.
Couper, M.P. and Nicholls II, W.L. (1998), "The History and Development of Computer Assisted Survey Information Collection." In M.P. Couper, R.P. Baker, J. Bethlehem, C.Z.F. Clark, J. Martin, W.L. Nicholls II, and J. O'Reilly (eds.), *Computer Assisted Survey Information Collection*. New York: Wiley.
Couper, M.P., Kenyon, K., and Tourangeau, R. (2001), "Picture This! An Analysis of Visual Effects in Web Surveys." Paper presented at the annual conference of the American Association for Public Opinion Research, Montreal, Canada, May.
Couper, M.P., Singer, E., and Tourangeau, R. (2002), "Understanding the Effects of Audio-CASI on Self-Reports of Sensitive Behavior." Paper presented at the annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.
Couper, M.P., Traugott, M., and Lamias, M. (2001), "Web Survey Design and Administration." *Public Opinion Quarterly,* 65 (2): 230-253.
Curtin, R., Presser, S., and Singer, E. (2000), "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly*, 64 (4): 413-428.
De Heer, W.F. (1991), "The Use of Handheld Computers in Social Surveys of the Netherlands Central Bureau of Statistics." *The Statistician*, 40: 125-138.
Fuchs, M. (2002), Remarks made at Roundtable on "Mobile Phones in Telephone Surveys." Annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.
Gray, J. (1999), "Handhelds in a Laptop World," *Survey Methodology Bulletin*, No .44. London: Social Survey Division, Office for National Statistics, pp. 7-13.
Gribble, J.N., Miller, H.G., Cooley, P.C., Catania, J.A., Pollack, L. and Turner, C.F. (2000), "The Impact of T-ACASI Interviewing on Reported Drug Use among Men Who Have Sex with Men." *Substance Use & Misuse*, 35 (6-8): 869-890.
Groves, R.M. (1989), *Survey Errors and Survey Costs*. New York: Wiley.
Harris, D.R. (2002), "In the Eye of the Beholder: Observed Race and Observer Characteristics." Paper presented at the annual meeting of the American Sociological Association.
Hutchings, V.L., Valentino, N.A., and Rusch, L. (2001), "Priming In-Group Identity: The Effects of Subtle Racial Cues in Campaign Ads." Unpublished paper.
Jenkins, C.R. and Dillman, D.A. (1997), "Towards a Theory of Self-Administered Questionnaire Design." In L. Lyberg, P. Biemer, M. Collins, E. deLeeuw, C. Dippo, N. Schwarz, D. Trewin (eds.), *Survey Measurement and Process Quality*, New York: Wiley.
Johnston, J. and Walton, C. (1995), "Reducing Response Effects for Sensitive Questions: A Computer-Assisted Self Interview with Audio." *Social Science Computer Review*, 13 (3): 304-309.
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.
Kiesler, S., Sieff, E., and Geary, C. (1992), "The Illusion of Privacy in Human-Computer Interaction." Carnegie Mellon University: Unpublished paper.

Krysan, M. and Couper, M.P. (2001), “Using Virtual Interviewers to Explore Race of Interviewer Effects.” Paper presented at the annual conference of the American Association for Public Opinion Research, Montreal, Canada, May.

Krysan, M. and Couper, M.P. (2002), “Measuring Racial Attitudes Virtually: Respondent Reactions, Racial Differences, and Race of Interviewer Effects.” Paper presented at the annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.

Kuusela, V. (2002), Remarks made at Roundtable on “Mobile Phones in Telephone Surveys.” Annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.

Lee, J.S. (2002), “Net Users Try to Elude the Google Grasp.” *New York Times,* July 25th, Circuits Section, p. 1.

Lyberg, L.E. (1985), “Plans for Computer Assisted Data Collection at Statistics Sweden.” *Proceedings of the 45th session, International Statistical Institute*, Book III, Topic 18.2.

Maynard, D.W., Houtkoop-Steenstra, H., Schaeffer, N.C., and van der Zouwen, J. (eds.)(2002), *Standardization and Tacit Knowledge; Interaction and Practice in the Survey Interview*. New York: Wiley.

Moon, Y. (1998), “Impression Management in Computer-Based Interviews: The Effects of Input Modality, Output Modality, and Distance.” *Public Opinion Quarterly*, 62 (4): 610-622.

Nass, C., Moon, Y., and Carney, P. (1999), “Are People Polite to Computers? Responses to Computer-Based Interviewing Systems.” *Journal of Applied Social Psychology*, 29 (5): 1093-1110.

Nass, C., Moon, Y., and Green, N. (1997), “Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers with Voices.” *Journal of Applied Social Psychology*, 27 (10): 864-876.

Nelson, R.O., Peyton, B.L., and Bortner, B.Z. (1972), “Use of an On-Line Interactive System: Its Effects on Speed Accuracy, and Cost of Survey Results.” Paper presented at the 18th ARF Conference, New York, November.

Nicholls II, W.L., Baker, R.P., and Martin, J. (1997), “The Effect of New Data Collection Technologies on Survey Data Quality.” In L. Lyberg, P. Biemer, M. Collins, E. deLeeuw, C. Dippo, N. Schwarz, and D. Trewin (eds.), *Survey Measurement and Process Quality*, New York: Wiley.

Nosek, B.A., Banaji, M. and Greenwald, A.G. (2002), “Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Web Site.” *Group Dynamics*, 6 (1): 101-115.

Nusser, S.M., Thompson, D.M., and Delozier, G.S. (1996), “Conducting Surveys with Personal Digital Assistants.” Paper presented at the International Conference on Computer Assisted Survey Information Collection, San Antonio, TX, December.

Nusser, S. and Fox, J. (2002), “Using Digital Geospatial Information to Locate Sample Units in the Field.” Paper prepared for the Committee on National Statistics’ Workshop on Survey Automation, Washington, D.C., April.

O'Reilly, J.M, Hubbard, M., Lessler, J., Biemer, P.P., and Turner, C.F. (1994), "Audio and Video Computer Assisted Self-Interviewing: Preliminary Tests of New Technologies for Data Collection." *Journal of Official Statistics*, 10 (2): 197-214.

O’Rourke, D., Sudman, S., and Ryan, M. (1996), “The Growth of Academic and Not-for-Profit

Survey Research Organizations." *Survey Research*, 27 (1-2): 1-5.
Poynter, R. (2000), "We've Got Five Years." Paper presented at the Association for Survey Computing's Meeting on Survey Research on the Internet, London, September.
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 R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (eds.), Survey Nonresponse. New York: Wiley.
Robinson, J.P., Neustadtl, A., and Kestnbaum, M. (2002), "Why Public Opinion Polls are Inherently Biased: Public Opinion Differences among Internet Users and Non-Users." Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL, May.
Rudolph, B.A., and Greenberg, A.G. (1994), *Surveying of Public Opinion: The Changing Shape of an Industry*, Chicago: National Opinion Research Center: Report to the Office of Technology Assessment.
Schonlau, M., Zapert, K., Simon, L.P., Sanstad, K., Marcus, S., Adams, J., Spranca, M., Kan, H.-J., and Berry, S. (2000), "Comparing Random Digit Dial with Web Surveys: The Case of Health Care Consumers in California." Santa Monica, CA: RAND, unpublished paper.
Schneider, S., and Edwards, B. (2000), "Developing Usability Guidelines for AudioCasi Respondents with Limited Literacy Skills." *Journal of Official Statistics*, 16 (3): 255-271.
Schuman, H. (1966), "The Random Probe: A Technique for Evaluating the Validity of Closed Questions." *American Sociological Review*, 31 (2): 218-222.
Suchman, L. and Jordan, B. (1990), "Interactional Troubles in Face-to-Face Survey Interviews." *Journal of the American Statistical Association*, 85: 232-241.
Taylor, Bremer, J., Overmeyer, C., Siegel, J.W., and Terhanian, G. (2001), "The Record of Internet-Based Opinion Polls in Predicting the Results of 72 Races in the November 2000 US Election." *International Journal of Market Research*, 43 (2): 127-135.
Tourangeau, R., Couper, M.P., and Steiger, D.M. (2002), "Humanizing Self-Administered Surveys: Experiments on Social Presence in Web and IVR Surveys." *Computers in Human Behavior*, forthcoming.
Tourangeau, R., Steiger, D.M., and Wilson, D. (2002), "Self-Administered Questionnaires by Telephone: Evaluating Interactive Voice Response." *Public Opinion Quarterly*, 66 (2): 265-278.
Turing, A.M. (1950), "Computing Machinery and Intelligence." *Mind*, 59 (236): 433-460.
Turner, C.F., Ku, L., Rogers, S.M., Lindberg, L.D., Pleck, J.H., and Sonenstein, F.L. (1998), "Adolescent Sexual Behavior, Drug Use and Violence: Increased Reporting with Computer Survey Technology." *Science*, 280 (May 8): 867-873.
Valentino, N.A., Traugott, M.W., and Hutchings, V.L. (2001), "Group Cues and Ideological Constraint: A Replication of Political Advertizing Effects Studies in the Lab and in the Field." Unpublished paper.
Vehovar, V. (2002), Remarks made at Roundtable on "Mobile Phones in Telephone Surveys." Annual conference of the American Association for Public Opinion Research, St. Petersburg Beach, FL, May.
Wherrett, J.R. (1999), "Issues in Using the Internet as a Medium for Landscape Preference

Research." *Landscape and Urban Planning*, 45: 209-217.
Wherrett, J.R. (2000), "Creating Landscape Preference Models Using Internet Survey Techniques." *Landscape Research*, 25 (1): 79-96.
Wong, D., and Baker, C. (1988), "Pain in Children: A Comparison of Assessment Scales." *Pediatric Nursing*, 14 (1): 9-17.