Survey Methodology, June 2001
Vol. 27, No. 1, pp. 7-31
Statistics Canada, Catalogue No. 12-001

Telesurvey Methodologies for Household Surveys – A Review and Some Thoughts for the Future

Gad Nathan [1]

Abstract

We consider 'telesurveys' as surveys in which the predominant or unique mode of collection is based on some means of electronic telecommunications – including both the telephone and other more advanced technological devices such as e-mail, Internet, videophone or fax. We review, briefly, the early history of telephone surveys, and, in more detail, recent developments in the areas of sample design and estimation, coverage and nonresponse and evaluation of data quality. All these methodological developments have led the telephone survey to become the major mode of collection in the sample survey field in the past quarter of a century. Other modes of advanced telecommunication are fast becoming important supplements and even competitors to the fixed line telephone and are already being used in various ways for sample surveys. We examine their potential for survey work and the possible impact of current and future technological developments of the communications industry on survey practice and their methodological implications.

Key Words: Telephone surveys; Internet surveys; Sample design; Nonresponse; Coverage.

1. Introduction

Electronic telecommunications have become a predominant factor in practically all aspects of modern life at the beginning of the new millennium. Sample surveys are no exception and the widespread use of the telephone as a prime mode of communication for at least the past quarter of a century has had an important influence on survey practice. In fact, the telephone survey has become the major mode of collection in the sample survey field, especially in North America and Western Europe, both for surveys of households and individuals and for surveys of establishments. Other modes of advanced telecommunication, such as e-mail, Internet, videophone, fax and mobile phones are fast becoming important supplements and even competitors to the fixed line telephone. They are already being used in various ways for sample surveys and in this review paper we intend to examine their potential for survey work and the methodological implications of their use. We therefore wish to use the term 'telesurvey' for any survey in which the predominant or unique mode of collection is based on some means of electronic telecommunications – including both the telephone and other more advanced technological devices. Conventional surveys based on face-to-face interviews in the home or (snail-)mail surveys are not included, unless a substantial component of the survey is based on some telecommunications instrument. Although this paper focuses on surveys of individuals and households, much of it is relevant to establishment surveys too. We refer to telesurvey 'methodologies' in the plural, since it seems obvious that no single methodology will be suitable for use with the plethora of possible communication devices available in the future and their combinations.

This paper has been prepared in recognition of Joe Waksberg's unique contributions to survey methodology, in general, and to telephone survey methodology in particular. It is well recognized today that his groundbreaking paper, Waksberg (1978), paved the way for the widespread efficient use of random digit dialing for telephone surveys and serves as a threshold point in the development of telesurvey methodology. Together with many of his subsequent papers, his work has had a profound influence on the theory and practice of telephone survey methodology, some of which will be examined in this paper.

We shall deal primarily with the statistical aspects of telesurvey methodology but recognize that these are not independent of non-statistical aspects, such as the cognitive features of telesurvey interviewing, survey administration and ethical considerations. In the following section we briefly review the early history of telephone surveys, through 1978. Section three reviews in some detail more recent developments in the areas of sample design and estimation, coverage and nonresponse and evaluation of data quality. Finally in section four we consider the possible impact of current and future technological developments of the communications industry on survey practice and their methodological implications.

2. The Early History of Telephone Surveys

In the following we review briefly the overall early development of the use of telephones for survey work, as background for the developments in telesurvey methodologies to be described later. More detailed and comprehensive coverage is provided in several books and survey papers, *e.g.*, Blankenship (1977a), Groves, Biemer, Lyberg, Massey, Nicholls and Waksberg (1988), Frey (1989),

1. Gad Nathan, Departement of Statistics, Hebrew University, 91905 Jerusalem, Israël.

Lavrakas (1993), Casady and Lepkowski (1998, 1999) and Dillman (1978, 2000).

Telephones have been used for survey work since the thirties, though generally as a supplementary mode of collection. Some have erroneously blamed the disastrous failure of the Literary Digest survey's prediction of a landslide victory of Landon over Roosevelt in 1936, at least partially, on telephone undercoverage (Katz and Cantril 1937; Payne 1956; and Perry 1968). In fact the survey was based on mail questionnaires and although telephone lists were used as a sampling frame (in combination with lists of automobile registrations), it seems that the failure was due more to nonresponse than to frame undercoverage (Bryson 1976; Squire 1988; and Cahalan 1989).

Most of the earliest reports on the use of the telephone in survey work were in the areas of public health or in market research applications. Many of them used some combination of telephone interviewing with other modes of collection and in some cases they included empirical comparisons of response rates or outcomes in order to assess mode effects. For instance, Cunningham, Westerman and Fischoff (1956) and Bennet (1961) report on telephone surveys for follow-up studies of patient treatment and Fry and McNaire (1958) on a national follow-up to a mail questionnaire to obtain opinions of hospital staff – all with high response rates. Mitchell and Rogers (1958) used telephone interviewing for a survey of telephone households on the consumption of dairy products and compare the results with those obtained from a control sample of non-telephone households. Cahalan (1960) compares results from telephone interviews with those from personal interviews in measuring newspaper readership with favourable results. Eastlack (1964) in a comparative telephone study of advertising recall and product usage shows that a rigorous call-back protocol provides more accurate results than a method without call-backs. Coombs and Freedman (1964) report on high telephone response (92%) in a longitudinal fertility survey, supplemented by personal interviews. Sudman (1966) describes several supplementary uses of the telephone for survey work, which include making of advance appointments and screening for rare populations, with positive results for cooperation rates and cost reductions.

In the late sixties telephone surveys really came of age, as a result of several different developments. First of all the rapid increase in telephone coverage in Western Europe and North America implied that telephone interviewing could be used as a primary mode of collection. In the US household telephone coverage reached a level of 88%. In 1970 (Massey and Botman 1988) and this level was reached somewhat later in most Western European countries, in Australia and in New Zealand (Trewin and Lee 1988). In parallel to the rapid increase in telephone penetration in many countries a serious decline in response rates and difficulties in contacting respondents by door-to-door collection were experienced in the late sixties. This led to serious consideration of telephone surveys both to reduce costs and to achieve higher cooperation rates. The use of telephone interviewing advanced most rapidly in commercial and academic survey organizations and less so in official government statistics. For instance the Federal Committee on Statistical Methodology (1984) reports that only about 11 percent of US Federal surveys in 1981 involved telephone interview in any form, in most cases in addition to other modes.

At first telephone interviewing was viewed with apprehension, even when used only as a supplementary mode of collection, due to fears of high nonresponse rates and response biases considered inherent when interviewing was not carried out face-to-face. Results of some of the earlier telephone surveys seemed to reinforce these fears. For instance, a study of leaflet receipt by Larson (1952) raises serious doubts on the validity of telephone responses on the basis of a face-to-face interview follow up. Similarly Oakes (1954) reports on suspiciously lower response on improvements to a consumer service via the telephone than obtained in face-to-face interviews. Schmiedeskamp (1962) in an attitude survey on consumer finances finds greater avoidance of taking strong positions when telephone interviewing was used. Wiseman (1972) in a comparison of mail questionnaire, telephone and face-to-face personal interviewing finds mode effects for sensitive issues (abortion and birth control). The main differences, however, are between responses to mail questionnaires and to personal interview (telephone or face-to face).

Many of these fears were allayed at an early stage by the results of a number of more rigorous empirical studies. Thus Hochstim (1967) in a well-designed controlled experiment compares collection by mail, telephone and personal interview as the primary mode of collection. The results demonstrate convincingly that the three strategies of data collection prove to be practically interchangeable when compared with respect to rate of return, completeness of return, comparability of findings and validity of responses. The major difference between modes is with respect to cost, with a clear preference for the mail or telephone strategy. Similarly a small test carried out by Colombotos (1965) on samples of a population of physicians shows no significance differences between responses obtained by telephone and by in-person interviews. Janofsky (1971) reports similarity in willingness to express feelings on health issues between telephone respondents and face-to-face interview respondents. A well designed validation study by Locander, Sudman and Bradburn (1976) of the effects of question threat and mode of collection found no meaningful differences in response bias between telephone and face-to-face interviews. Finally, in a small carefully controlled field experiment, Rogers (1976) tested the effects of alternative interviewing strategies on the quality of responses and on field performance in a survey on a variety of complex attitudinal, knowledge and personal items. The results again indicate that the quality of data obtained by telephone is comparable to that obtained by interviews in person. A major national study comparing telephone and face-to-face interviewing was conducted by Groves and Kahn (1979). It was based on an intensive analysis of the large omnibus surveys carried out under the two modes by the University of Michigan Survey Research Center. It provided important information on data quality which did not indicate any substantial mode effects. These and other early studies, which foreshadowed several systematic studies of mode effects carried out in the eighties and nineties (to be discussed later) contributed to the legitimacy of telephone surveys as a standard mode of collection.

The initial use of telephones for sample surveys was usually based on samples selected from general frameworks, such as telephone directories, or from specific frameworks for small sub-populations. Towards the end of the sixties there was increased awareness of high rates of unlisted telephone numbers and of substantial differences between households with listed and non-listed numbers (see details in section 3.1.1). An important development that overcame this problem was the sampling method of Random Digit Dialing (RDD), first introduced by Cooper (1964) and further improved and developed by Eastlack and Assael (1966) and by Glasser and Metzger (1972). An inherent inefficiency of these basic element RDD methods was the large amount of numbers to be called that did not yield an interview (non working and non residential numbers). A two-stage RDD sampling method was first proposed to deal with this problem by Mitofsky (1970) and subsequently elaborated and put on a firm theoretical basis by Waksberg (1978). The introduction of what was to become known as the Mitofsky-Waksberg scheme contributed greatly to the widespread use of telephone surveys in the eighties and nineties.

Finally the technological advances in telecommunications and automation in the sixties and seventies contributed to the advantages of telephone surveying. Universal direct long distance dialing enhanced the possibilities of carrying out national surveys from a single center or from a small number of interviewing centers with all the advantages of central control and administration. However the greatest impact on the expansion of telephone surveys has undoubtedly been the introduction of Computer Assisted Telephone Interviewing (CATI) in the seventies. This is due both to the simplicity of CATI for conducting telephone interviews and to the possibilities it offers for the use of automation in many important non-interviewing tasks, (*e.g.*, dialing, recall schedules *etc.*).

One of the first uses of the computer for on-line questioning was in the form of a multi-station computer-based laboratory experiment designed to elicit subjective information – Shure and Meeker (1970). A good account of the early history of CATI can be found in the special issue of Sociological Methods and Research (Freeman and Shanks 1983), following the Berkeley Conference on Computer-Assisted Survey Technology held in Spring 1981. Market research organizations were the first to introduce CATI systems for their current operations. Chilton Research Services developed and used the Survey Response Processor on a current basis already in 1972 – Fink (1983). Other commercial survey organizations, applying different systems, realized early on the advantages of CATI – for instance the A&S/CATI™ system (Dutka and Frankel 1980). Academic survey research organizations were quick to follow with the earliest systems developed at UCLA and Berkeley for the large scale CATI-based California Disability Survey – Shanks, Nicholls and Freeman (1981) and Shanks (1983). Another early development of a CATI system at an academic survey organization, using a different approach, based on microcomputers, was that of the University of Wisconsin (Palit 1980; Palit and Sharp 1983). In Europe the first survey research organizations to use CATI were Social and Community Planning Research (SCPR – now the National Centre for Social Research) in the UK (Sykes and Collins 1987) and the State University of Utrecht, Netherlands (Dekker and Dorn 1984). The introduction of CATI systems into official statistics was slower. In the US it started in 1982 at the Census Bureau (Nicholls 1983) and at the National Agricultural Statistics Service (Tortora 1985) and at the same time at Statistics Netherlands (1987). By 1987 practically all organizations surveyed in a (non-probability) sample of 27 survey organizations (eighteen in the US and nine elsewhere) were using CATI for some or all of their telephone surveys – Berry and O'Rourke (1988). A report of the Federal Committee on Statistical Methodology (1990) indicated that the number of CATI installations worldwide at the end of the eighties was estimated to exceed 1,000 and that in 1988, the U.S Government had 51 cooperating CATI centers. It should be noted that the development of CATI quickly became part of a wider movement toward computer assisted interviewing (CAI) or computer assisted information collection (CASIC), which includes also CAPI (Computer Assisted Personal Interviewing) and CASI (Computer Assisted Self Interviewing) – Nicholls (1988). A more complete history of the development of CATI and of CASIC, in general, can be found in Couper and Nicholls (1998).

3. Recent Developments in Telephone Surveys

In the last quarter of a century telephone surveying has definitely come of age. Lyberg and Kasprzyk (1991) claim that it has become the dominant mode of collection in countries with extensive telephone coverage.

Hundreds of scientific papers have been published during this period on a wide range of different aspects of telephone surveys. Several general books on the subject have appeared – Blankenship (1977a), Groves and Kahn (1979), Frey (1989) and Lavrakas (1993). A number of conferences have been devoted to telephone survey methodology or have dealt with specific aspects of the topic. The results have appeared in monographs or special issues of scientific journals. A major conference on telephone survey methodology was held in November 1987 in Charlotte, NC, with the resulting volume edited by Groves, Biemer, Lyberg, Massey, Nicholls and Waksberg (1988) and the special issue of the Journal of Official Statistics, edited by Groves and Lyberg (1988b). The Berkeley Conference on Computer-Assisted Survey Technology held in Spring 1981 (Freeman and Shanks 1983) dealt primarily with telephone surveys and CATI was a major topic at the InterCASIC '96 International Conference on Computer Assisted Survey Information Collection, held in San Antonio, TX in December 1996 (Couper, Bethlehem, Baker, Clark, Martin, Nicholls and O'Reilly 1998) and at the ASC 3rd International Conference at Edinburgh in September 1999 (Banks, Christie, Currall, Francis, Harris, Lee, Martin, Payne and Westlake 1999).

Extensive bibliographies with several hundred entries can be found in the above sources, as well as in Khurshid and Sahai (1995), which covers the period through 1991, and in Survey Research Center (2000), which updates previous bibliographies with respect to sample design for household telephone surveys through 2000.

In the following we review the development of telephone survey methodology for household surveys during the past 25 years in the areas of sample design and estimation, coverage and nonresponse and evaluation of data quality.

3.1 Sample Design and Estimation

Sampling methodology for telephone surveys is based on the general principles of sampling. It is primarily adapted to the special situation of telephone surveys with respect to the sampling framework used. Thus we adopt the classification proposed by Lepkowski (1988) for telephone sampling methods, according to the underlying sampling framework – directory and commercial lists, telephone numbers (RDD) and combined methods (list-assisted and dual frame).

3.1.1 List-based Sampling Procedures

As mentioned above, the earliest telephone surveys were all based on samples selected from lists. In many cases they were mixed-mode surveys where telephone interviewing was used to supplement for non-response in face-to-face interviews or for follow-up. Thus so-called 'warm telephone interviewing' schemes have been used in the US Current Population Survey and in the Canadian Labour Force Survey – Drew, Choudhry and Hunter (1988). In these cases sampling is based on a general list framework to which information on telephone numbers is added and no special features of the use of the telephone are involved in the sample design. The same goes for 'pure' telephone surveys of special populations, such as physicians, for which a complete list of the population is available with telephone numbers and can be used as a sample framework – see, for example, Gunn and Rhodes (1981). Another example is where telephone interviewing is used in follow-up waves of a panel survey with the first contact carried out by a face-to-face interview. For instance in the Israel Labor Force Survey the first contact is by a home visit and the second and third waves are carried out by telephone for households who are willing to respond by telephone – Nathan and Eliav (1988). A related approach, used recently in a pilot study for the US National Study of Health and Activity (Maffeo, Frey and Kalton 2000), is to take an area sample, find telephone numbers where possible, for telephone interviewing, and use face-to-face interviewing for other households and for telephone nonrespondents.

The most easily obtained and low-cost directory that can be used as a framework for a telephone surveys is, of course, the telephone directory itself, or some modification of it. Originally the paper version of the directory was used, while nowadays an electronic version would usually be available. The major deficiencies of the telephone directory as a sampling framework are well documented. They are undercoverage, overcoverage, duplication and lack of auxiliary information. Undercoverage is by far the most serious deficiency and includes both non-telephone households and households with telephones unlisted by choice or those not yet included in the directory. The biases due to non-telephone households are, of course, irrespective of the framework used and will be dealt with in section 3.3.

The extent of unlisted telephones varies considerably by country and type of location, as well as by other household variables. Sykes and Collins (1987) report on an unlisted rate of 4% in the Netherlands and 12% in the UK. Fréjean, Panzani and Tassi (1990) estimate the unlisted rate in France as 14% and national US estimates in the seventies were of over 17-19% (Blankenship 1977b and Glasser and Metzger 1975). Rich (1977) reports on increasing rates of nonpublished telephones (excluding those involuntarily unlisted) in the Pacific Telephone's California serving area from 9% in 1964 to 28% in 1977. In addition some 5% of home telephones in California were estimated to be involuntarily unlisted (assigned after publication of the directory). More recent studies show substantially higher unlisted rates. Thus Genesys (1996) reports unlisted rates of 40% in 1993 and of 37% in 1995, based on national samples of more than 100,000 RDD telephone interviews and Survey Sampling Inc. (1998) estimates the US national unlisted rate for 1997 at 30%. Results of a small-scale study of the Jerusalem area (Nathan and Aframian 1996) indicate an unlisted rate of 27%.

Many studies have shown substantial differences between listed and unlisted telephone household characteristics, indicating disturbing potential coverage biases for directory-based samples. In the US these differences were demonstrated, for instance, in a study by Brunner and Brunner (1971), who found highly significant differences between listed and unlisted telephone households with respect to a wide range of demographic and socio-economic variables. Leuthold and Scheele (1971) found higher rates of nonlisting among blacks, city dwellers, young people, apartment dwellers, divorced and separated and among service workers. Similarly, Roslow and Roslow (1972) found significant differences in audience shares between listed and unlisted telephone households. Glasser and Metzger (1975) showed that nonlisted rates were higher in the West, in major metropolitan areas, among non-whites and the young. Blankenship (1977b) and Rich (1977) found highly significant differences between listed and unlisted households with respect to sex and age of household head, occupation, household size and income. In the UK Sykes and Collins (1987) found more unlisted numbers among the young, the poorest and those living in London. The results of Nathan and Aframian (1996) for the Jerusalem area showed lower rates of TV ownership and of TV viewing (of those with TV) in an RDD sample as compared with a directory listing sample.

Besides the undercoverage resulting from unlisted numbers, as indicated above, directory listings also suffer from problems of overcoverage, duplication and lack of updated auxiliary information. Overcoverage occurs when a unit outside the population is included in the framework. This may be due to the fact that disconnected numbers often remain in the directory, commercial numbers are not always clearly designated as such or other cases of unrecognized ineligibility. Duplication occurs when the same unit is represented in the frame more than once and the duplication is not recognized. Duplication can usually be discovered during sampling if the entries for the same household are listed consecutively but not if they appear separately (*e.g.*, under different surnames). If duplication is ascertained during the interview (*i.e.*, by obtaining information on the number of connected lines available to the household or the number of directory listings) it can be dealt with by appropriate weighting. Although these problems are surmountable, at a cost, that of undercoverage is not and this indicates the need for more representative sample frameworks than provided by directories. A popular alternative to the traditional telephone directory (in general prepared by the company providing telephone service to the area) has been the lists prepared by commercial firms, usually for purposes of marketing. These may be city directories, obtained from municipal address listings with telephone numbers obtained from directories or other sources, subscriber lists of telephone companies or national master address lists, such as that provided by Donelley Marketing, Inc. in the US – Lepkowski (1988). These lists provide important auxiliary data, such as geographic information, from the Census of Population and Housing and from other sources. They do not, in general, overcome the bias due to unlisted numbers and their cost may be high. They can result in some gain in sampling variance, due to the possibility of basing an efficient design on the auxiliary information. Potentially, lists used by emergency services to determine the physical location of callers could be used as frameworks, although access to these lists would be difficult for non-government survey organizations.

3.1.2 Random Digit Dialing – The Mitofsky-Waksberg Scheme

In order to overcome many of the inherent problems of directories and commercial lists, Random Digit Dialing (RDD) methods have become a popular choice for telephone surveys, primarily in the US. These are based on the frame of all possible telephone numbers. The method was originally proposed by Cooper (1964), who added random four digit suffixes to known prefixes in a local survey. This basic element sampling method was further improved and developed by Eastlack and Assael (1966) and by Glasser and Metzger (1972), on a national level, by identifying ‘working banks’ of numbers from telephone company information.

The use of RDD has until recently been confined, by and large, to the US and Canada. Thus Sykes and Collins (1987) report that telephone surveys were still rare in the UK at the end of the eighties, primarily due to low telephone coverage. In particular RDD surveys were rarely used – one of the reasons being the lack of uniformity in the length of telephone numbers at the time. However recently, with the increase of telephone coverage in the UK to some 96% at the end of the nineties and the standardization of telephone numbers to ten digits, RDD surveys have become more popular – see *e.g.*, Collins (1999) and Nicolaas, Lynn and Lound (2000). Similarly, Gabler and Haeder (2000) report that an RDD method, modified in order to deal with varying telephone number lengths (from 6 to 11 digits!), is now standard procedure for telephone surveys in Germany.

Mitofsky (1970) first proposed a two-stage RDD sampling method to deal with the problem of the inherent inefficiency of these basic element RDD methods due to the large amount of numbers to be called that did not yield an interview (non working and non residential numbers). This was subsequently elaborated and put on a firm theoretical basis by Waksberg (1978) and the method became known as the Mitofsky-Waksberg scheme. This scheme or variations of it have become the predominant sampling method for telephone surveys, at least in the US.

The method is based on the fact that household telephone numbers are, in general, clustered in series of consecutive numbers or within banks of numbers with the same first r digits. For the US r is usually set at eight (for ten digit telephone numbers, including area code), so that the banks or clusters (PSU’s) are of size $N = 100$ each. It is assumed that the telephone company can provide a list of all operating prefixes (area code + first three digits), *i.e.*, those to whom residential numbers have been assigned. To the six digit numbers in this list all possible choices of two digits are added, resulting in a sampling frame of eight digit numbers that represent the M PSU’s in the population. Sample PSU’s are selected from this frame at random (with replacement) consecutively and for each PSU selected two final digits are selected at random. The resulting ten digit number is dialed and if the number is not that of a residence (according to the survey definition), the PSU is dropped from the sample. If it is a residence a simple random sample (without replacement) of k additional residential numbers is selected by contacting numbers selected at random (without replacement) from the PSU, until k additional residential numbers are obtained. The procedure of PSU selection continues until a fixed number of PSU's, m, has been selected. It is easily seen that, assuming that the number of residential numbers in each selected PSU, P_i, is at least k, the total sample size of residential telephone households is $m \ (k+1)$ and that the final sample is an equal probability sample from the population of all residential telephone households.

Waksberg (1978) shows that if we designate by: $\pi = (\sum_{i=1}^{M} P_i)/(NM)$ the proportion of residential numbers in the population and by t the proportion of PSU's with no residential numbers (*i.e.*, for which $P_i = 0$), then the expected number of total calls is given by: $m[1+(1-t)k]/\pi$, assuming that $P_i \geq k+1$ for all PSU's with at least one residential number. The last assumption can be dropped if PSU's are grouped so that the restriction holds in each group or if unequal weighting is used. Optimal values of the design parameters are obtained under a simple cost function and the method is extended to deal with repeated surveys. The main advantage of the method is the reduction in the expected number of calls which have to be made in order to attain a given effective sample size, especially if t, the proportion of PSU's with no residential numbers, is larger than 0.5. Groves (1977) provides data for a national study indicating a value of t of about 0.65. This advantage has to be weighed against the increase in variance due to the effect of clustering. However, taking costs into account, illustrative calculations for typical values of the parameters show that reductions in costs run between 20 and 40%.

The major operational drawback of the method is in its sequential nature. This makes it unwieldy to carry out manually. However the sequential operation poses no problem when the process of selection is fully automated. The method as described above has some additional problems, most of which can be overcome by simple modifications. Assuming that prior information on the number of telephone households is not available, selection probabilities are not known, although the value of p can be estimated from the sample. The practical necessity to introduce a stopping rule for the number of calls to numbers which do not answer or to refusals to answer, even whether the number is a residential one, implies that the method cannot be strictly applied as designed, resulting in possible bias. The problem of households with multiple telephone numbers can be overcome if correct information on the number of different lines is obtained but the required re-weighting impinges on the simplicity of equal weighting. In some cases names and addresses can be obtained for RDD numbers by matching with address lists so that advance notice can be sent to at least part of the potential respondents. However this is a complex procedure and the difficulties in sending advance notice to respondents (common to all RDD procedures) has made the procedure difficult to consider for some official statistical agencies.

3.1.3 Modifications of the Mitofsky-Waksberg and Other RDD Methods

Some of the drawbacks of the basic method are overcome by the generalization due to Potthoff (1987a, 1987b). The method is based on the definition of a set of auspicious telephone numbers. This could consist of only residential numbers, as in the Mitofsky-Waksberg method, or a wider set which includes all residential numbers – for instance the set of all numbers which ring (including engaged, recorded messages and operators). The first stage of selection is by simple random sampling of a fixed number, m, of PSU's. From each selected PSU a fixed number of calls, c, are made and for each of them it is determined whether the number is auspicious or not. A PSU is discarded if all c numbers selected are inauspicious. Retained PSU's are defined as Type I if only one number is auspicious and as type II if two or more are auspicious. The second stage consists of selecting and dialing kc numbers from Type I PSU's and $k(c-1)$ numbers from Type II PSU's, where k is an integer. At all dialed numbers the unit is determined as residential or out-of-scope and an interview is attempted for all residential units. A supplementary sequential segment for Type I PSU's selects additional telephone numbers that are dialed until a total of k auspicious numbers are obtained. An interview is attempted at each auspicious numbers dialed in the sequential segment. Potthoff (1987a) shows that, under certain conditions, all residential telephone numbers have the same probability of selection and develops unbiased and ratio estimates and their variances. Cost comparisons and some modifications to overcome practical problems are also given. The method reduces the problem of ambiguity on the status of dialed numbers from which no response is obtained and also the problem of exhaustion of the residential numbers in a PSU.

A large number of additional generalizations and modifications to the basic Mitofsky-Waksberg method have been proposed. Many of these attempt to reduce the burden of interviewing screening and to improve control over the initial contact sample size. Thus Hogue and Chapman (1984) propose determining cutoff points on the basis of an estimation of the probability that a PSU is 'sparse', *i.e.*, has a small proportion of residential numbers, and propose to determine an optimal cutoff procedure on the basis of cost and variance considerations. Alexander (1988) considers two types of cutoff rules to limit interviewing screening for prefixes with low residential densities. An 'increasing rule' stops as soon as a predetermined number of calls, c_i, has been made and less than i residences have been found, where $\{c_i\}$ is an increasing series in i. A 'decreasing rule' stops when i residences have been found if at least c_i calls have been made, where $\{c_i\}$ is a decreasing series in i. The costs for these rules are evaluated under a simple model.

Lepkowski and Groves (1986a) propose a two phase design based on matching prefixes selected in the first stage of the Mitofsky-Waksberg scheme to a commercial directory to obtain counts of listed telephones for each prefix selected. Prefixes are allocated to two strata – a low density stratum where there are no listed telephone numbers, or only a small number of them, and a high density stratum. The Mitofsky-Waksberg design is applied to the low-density stratum and telephone numbers are selected with probability proportional to the number of listed telephone numbers in the high-density stratum.

Brick and Waksberg (1991) propose using a fixed number of telephone numbers in the second stage so as to avoid sequential sampling altogether with a resulting simplicity of operation. The design, originally proposed by Waksberg (1984), is not, however, self-weighting and involves a slight bias and increased variance. Brick and Waksberg (1991) suggest considerations for the choice between the original and modified Mitofsky-Waksberg designs. For an early application of the modified Mitofsky-Waksberg method to the collection of health attitude information, apparently in an erroneous attempt to implement the original method – see Cummings (1979). Smith and Frazier (1993) compare the original and modified schemes, using data collected in the California Behavioral Risk Factor Surveillance System. The results indicate that the modified scheme speeds up the data collection, resulting in a larger sample size for the same cost. This compensates for larger design effects of the modified scheme.

Another alternative to the basic Mitofsky-Waksberg method is the use of stratification and disproportionate allocation to improve 'hit rates', proposed by Palit (1983). An evaluation of alternative treatments of unanswered telephone numbers for the Mitofsky-Waksberg design is carried out by Palit and Blair (1986). The optimal determination of parameters for the Mitofsky-Waksberg method is dealt with by Burke, Morganstein and Schwartz (1981) and the optimal allocation for the stratified version of the method by Casady and Lepkowski (1991, 1993) and by Tucker, Casady and Lepkowski (1992). Further problems relating to minimal cost allocation are treated by Palit (1983) and by Mason and Immerman (1988).

3.1.4 List-Assisted Methods

Although RDD methods overcome the undercoverage of directories due to unlisted numbers, they all still suffer from the basic problem of undercoverage due to non-telephone households (see further detail in section 3.3). In addition the lack of auxiliary information (such as geographical information), which is often available in list frames, leads to inefficiencies, even in the more sophisticated modifications of the basic methods, mentioned above. Thus alternative methods have been sought to combine RDD samples with samples based on list and directory frames. One of the earliest attempts in this direction was that proposed by Stock (1962) and elaborated by Sudman (1973), based on replacing the last two digits of telephone numbers, selected from a directory listing, by random digits. The method was applied by Hauck and Cox (1974) to a methodological study of mode effects in screening for a special sub- population. A simpler version, popularly known as the 'Plus One' method, replaces each telephone number sampled from a directory by the number plus one (or some other digit – known as the 'plus digit method). This supposedly overcomes the bias due to unlisted numbers. Due to its simplicity, the method has gained popularity among market researchers. However several studies – *e.g.*, Landon and Banks (1977); and Mullet (1982) – have indicated that it is not, in fact, bias-free and also suffers from low efficiency.

Forsman and Danielsson (1997) propose a model-based approach for plus digit sampling, based on the assumption of randomly mixed listed and unlisted numbers within prefix. The model, which is tested empirically, provides model unbiased estimates. Ghosh (1984) has proposed an improved method that continues adding one to the last telephone number dialed as long as a household is not reached and stopping once a household is reached. Although still biased, the bias is reduced as compared with the simple 'plus one' method. Other list-assisted methods with RDD components, are discussed by Potter, McNeill, Williams and Waitman (1991), who stratify prefixes according to counts of published telephone numbers, while ensuring inclusion of blocks without any published numbers.

Brick, Waksberg, Kulp and Starer (1995) propose a list-assisted method that overcomes the troublesome problem of the sequential nature of the second stage sampling inherent in the Mitofsky-Waksberg scheme. The method is based on dividing the file of exchanges (100-banks) into two strata. The first consists of all exchanges with at least one listed residential phone and the second those that have none. Sampling only from the first stratum drastically reduces the proportion of nonresidential numbers which have to be dialed, but results in coverage bias. They investigate the bias and conclude that such truncated sampling methods are efficient and have operational advantages, while the resulting coverage bias (about 4%) is not too important. The method has been widely applied to replace the classical Mitofsky-Waksberg method. Similarly Statistics Canada has used the method for their General Social Survey since 1991 for the whole sample, with simple random sampling within banks of numbers identified as having at least one residential number (Norris and Paton 1991). Modifications of this design include a complete stratification of number banks on the basis of list information and using simple RDD for strata with small proportions of banks with no listing and the

Mitofsky-Waksberg method in the remaining strata. A comparison of this design with other stratified designs based on a cost model is carried out by Casady and Lepkowski (1993). Their results show that for low cost ratios (of productive selections to unproductive selections) two and three stratum RDD designs are as efficient as the Mitofsky-Waksberg scheme and that for high cost ratios they are superior.

3.1.5 Multiple Frame Designs

In an attempt to overcome some of the inherent biases of telephone surveys due to directory and telephone under-coverage, the use of dual frame mixed mode surveys, combining telephone with face-to-face interviewing, has received increasing attention. These combine conventional samples for personal interview with RDD or directory samples for telephone interviewing. Biemer (1983) investigated the optimal mix for such designs, via a simulation study, and McCarthy and Bateman (1988) propose the use of mathematical programming for attaining optimal allocation of sample units for a dual frame design, which allows posterior analysis of the effects of variations in design and cost parameters on the optimization. Choudhry (1989) proposes a cost-variable optimization for estimating proportions and Brick (1990) proposes the use of multiplicity sampling for this purpose. In a series of papers, Groves and Lepkowski (1985, 1986); Lepkowski and Groves (1984, 1986b); and Traugott, Groves and Lepkowski (1987) develop error models for these dual frame survey designs. They also report on results of experiments to compare response rates and potential biases of RDD and list samples and of several interviewing methods. The results were applied to the large scale US National Crime Survey.

Whitmore, Mason and Hartwell (1985) report on applications of dual frame dual mode methods in a US Environment Protection Agency sponsored study of personal exposure to carbon monoxide in two metropolitan areas and in a state-wide study of social service needs. In both cases commercially available directory lists were used in association with area household sampling. On the basis of an analysis of their results, they recommend the use of such dual designs in order to both benefit from the relative efficiency of telephone interviewing and to overcome the biases inherent in the use of directories as sampling frames. A combination of RDD and area sampling is reported by Waksberg, Brick, Shapiro, Flores-Cervantes and Bell (1997) for the US National Survey of America's Families in with there was particular focus on the low-income population. The nontelephone households identified in the area screening were given cellular phones for responding to telephone interviewers, thereby avoiding the need to train the area screener interviewers in a non-telephone questionnaire (Cunningham, Berlin, Meader, Molloy, Moore and Pajunen 1997).

3.2 Other Sampling Issues

3.2.1 Sampling for Special Populations

The relative low costs of telephone interviewing have made this survey mode a prime candidate for use in screening large samples in order to locate small special populations. Thus Sudman (1978) discusses the conditions under which the use of a telephone sample for screening a subgroup, to be finally interviewed face-to-face, is more efficient than face-to-face screening. By analyzing cost functions, telephone screening is found to be efficient, unless within-cluster homogeneity is small, interview densities are low and/or location and screening costs are low, relative to interview costs. Blair and Czaja (1982) propose a modification of the Mitofsky-Waksberg procedure to locate special populations that cluster geographically and describe an application to the Black population. As pointed out however by Waksberg (1983), their method requires reweighting when clusters are exhausted, which may result in reduced efficiency. This implies that the method may be efficient for the Black population but not necessarily for other minorities. Another telephone sample design targeting the US black population is proposed by Inglis, Groves and Heeringa (1987). Mohadjer (1988) proposes the stratification of prefix areas in an RDD design for sampling rare populations. The use of the Mitofsky-Waksberg method for selection of households combined with a stratified sample of individuals within household is used for the selection of a population-based control group in four epidemiological studies reported by Hartge, Brinton, Rosenthal, Cahill, Hoover and Waksberg (1984). The effectiveness of the method is studied by Perneger, Myers, Klag and Whelton (1993), on the basis of a simulation of simple random sampling, and found to be effective.

Local area surveys are another example of special populations that can be dealt with efficiently by a telephone survey. Although, in general, telephone exchanges do not define geographical areas exactly, there is a high degree of correspondence and, with some screening for those in the defined area, telephone interviewing can reduce costs considerably. For instance Banks and Hagan (1984) report on the reduction of interviewer screening by a combination of list sampling and RDD for a survey to assess the effectiveness of health programs in specific service areas. Similarly, Campbell and Palit (1988) tested a combination of list sampling and TDD – total digit dialing, using a frame of all numbers in exchanges corresponding to a given census area. They found that this resulted in a substantial saving in enumeration costs, versus face-to-face interviewing.

3.2.2 Sampling Individuals Within Households

Almost all household surveys include questions relating to individuals in the household. In some cases all individuals belonging to the household are included in the sample, but in many cases, for various reasons, a sample of one or more individuals is selected within the household for individual questions. The classic Kish procedure (Kish 1949), predominantly used in face-to-face interview surveys raises particular problems for telephone surveys, because it requires obtaining complete household listings over the telephone. This is more difficult to obtain over the phone than in a face-to-face interview, where some of the persons may be physically present. It should be pointed out however that in many cases the information on household composition is required in any case. In addition the manipulation of the selection rules by the interviewer (*e.g.*, to accomplish high response rates), which has long been suspected in face-to face interviewing is almost impossible in CATI surveys (where selection is invisible to the interviewer).

Troldahl and Carter (1964) proposed a method whereby only the number of persons of each sex is required. Probabilistic rules (*e.g.*, 'oldest man') are then applied to determine the individual selected, ensuring known selection probabilities for each person. However a positive probability of selection for each individual is not ensured (*e.g.*, in households with three males the one of intermediate age is never selected). The method (known as the 'Troldahl-Carter method') has been modified by Bryant (1975), in order to take into account the possibility of households with more than two individuals of the same sex. An alternative method proposed by Salmon and Nichols (1983) and by O'Rourke and Blair (1983) is to select the person with the next (or last) birthday (the 'next-birthday' or 'last-birthday' method), which ensures equal probability of selection for each household member, under the assumption that the date of interview is random. This is of course a reasonable assumption only for surveys carried out over a twelve-month period but not for surveys with shorter interview periods. This and other factors may lead to selection probabilities that are correlated with the individual characteristics. Another selection method proposed by Hagan and Meier (1983), which does not require any preliminary information on household composition, selects a predefined person (*e. g.*, 'eldest man'). The method again fails to ensure a positive probability of selection for each household member.

Several empirical comparisons of the above methods have been carried out. Czaja, Blair and Sebestik (1982) found no significant differences in response rates or in demographic profiles between two versions of the Troldahl-Carter method and the Kish method. Hagan and Meier (1983) compare their method, described above, with the Troldahl-Carter method and find that the method they propose has a significantly lower refusal rate, with no significant differences in demographic profiles. Salmon and Nichols (1983) compare four procedures for selecting respondents within a household unit – Troldahl-Carter, male/female alternation, next-birthday and no-selection methods – in a small telephone survey. They reach the conclusion that the next-birthday method is a relatively efficient procedure for selecting a sample that is representative of all household members. Oldendick, Bishop, Sorenson and Tuchfarber (1988) find no significant differences between the Kish method and the last-birthday method. In a study using the last birthday method, Romuald and Haggard (1994) find that informants self-select to participate at a higher rate than expected. They investigate the effect of using memory cues on respondent self-selection and reach the conclusion that there is no significant effect. Lavrakas, Bauman and Merkle (1993) evaluate the effect of the use of the last-birthday method on within-unit coverage in a national survey and report evidence to suggest that the method leads to incorrect selection in many cases. Forsman (1993) reviews experiences of within-household sampling for 18 private opinion research companies and report on a test to compare the Kish, next/last birthday and the Toldahl-Carter methods. They conclude that the Troldahl-Carter method is somewhat better than the Kish method and that both are superior to the birthday methods. Similarly, Binson, Canchola and Catania (2000) report on a three-way comparison in a national telephone survey between the Kish, next-birthday, and last-birthday methods, and find significant differences between the three methods in the dropout rate, during the initial stages of the screening process. The Kish method had the highest dropout rates and the 'next-birthday' had the lowest rate. They conjecture that interviewers, rather than respondents, are a primary source of the higher rate of refusals when using the Kish method, due to the fact that a full household roster is required.

3.3 Coverage and Nonresponse

3.3.1 Telephone Coverage

The problem of telephone noncoverage was until very recently a major drawback of telephone surveys. Even in the US overall person undercoverage (in nontelephone households) remained at 7.2% by the end of 1986 – Thornberry and Massey (1988). By the mid-eighties household telephone undercoverage was less than 10% in most Western countries, with the highest coverage (99%) in Sweden. But some countries still had high rates of telephone undercoverage, for instance: UK 25%, Italy 29% Ireland 50%, Israel 30% – Trewin and Lee (1988). The situation changed dramatically towards the end of the century, with most Western countries reaching virtual saturation. Telephone coverage reached 94.4% in the US in 1999 (NTIA 2000); 96.6% in Australia in 1996 (St. Clair and Muir 1997); 97.0% in the UK (OFTEL 1999); 97.3% in Israel (Central Bureau of Statistics 2000); 97.9% in Finland (Kuusela and Vikki 1999); 98.2% in Canada (Statistics Canada 1999); and 99% in Germany (Federal Republic of Germany 1999).

Obviously the major problem of telephone undercoverage lies primarily in differential undercoverage rather than in its overall rate and the fact that telephone undercoverage is highly correlated with a wide range of demographic, economic and health variables. This has been demonstrated extensively in a large number of empirical studies in the US and elsewhere – see for instance Groves and Kahn (1979), Collins (1983, 1999), Thornberry and Massey (1983, 1988), Trewin and Lee (1988) and Botman and Allen (1990). The rapid increase in overall telephone coverage over the last decade has not caused any radical change in this situation. Thus in Finland, with an overall telephone undercoverage of 2.1% in 1999, low income households (less than 675 Euros per month) had an undercoverage of 11.3% (*vs.* 0% for high income groups) and those living in rented accommodation 4.9% (Kuusela and Vikki 1999). In Israel telephone undercoverage was 17.9% for the lowest income decile as against 0.8% for the two highest deciles and 24.9% for single adult households with three or more children as against 2.4% for childless households with three or more adults (Central Bureau of Statistics 2000). Similarly in the US large geographical variations are still found and telephone undercoverage is found to correlate with housing deficiencies, race, education income and mobility (Shapiro, Battaglia, Hoaglin, Buckley and Massey 1996; Giesbrecht, Kulp and Starer 1996; Fox and Riley 1996; NTIA 2000). Health- related characteristics were found to differ somewhat between persons in telephone and non-telephone households in the National Health Interview Survey by Anderson, Nelson and Wilson (1998) and in the National Health and Nutrition Examination Survey by Ford (1998). However telephone coverage effects were considered to be minor in both studies.

However the main problem of telephone coverage foreseen for the near future relates to the introduction and rapid proliferation of mobile telephones. In the late nineties the proportion of households with access to at least one mobile telephone reached 76% in Finland, 59% in Denmark, 35% in Italy (Rouquette 2000) and 52% in Israel (Central Bureau of Statistics 2000). If all these mobile telephones were additional to fixed line telephones no problem would arise. However there are already strong indications of a tendency in several countries to consider the mobile telephone as an alternative to a fixed line telephone, rather than a supplement. Kuusela and Vikki (1999) report that 20% of Finnish households now have exclusively one or more mobile telephones and no fixed line and predict that within a year the number of mobile phones will exceed the number of fixed lines. Similar figures for the UK are 3% (OFTEL 2000) and for Israel 2.9% (Central Bureau of Statistics 2000). This implies that fixed line telephone coverage is down to 77% in Finland and to 94% in the UK and in Israel. In Germany it is estimated that the percentage of households with fixed line telephones will decrease to 92% by 2004 (Gabler and Haeder 2000). Furthermore the characteristics of persons with only mobile telephones are quite different from those with fixed telephone lines. In Finland, according to Kussela and Vikki (1999), they tend to be young, often living alone in rented apartments in urban areas. It should be noted that the transfer from fixed phone lines to mobile telephones is apparently not occurring to any large extent in North America, due to differences in pricing strategies.

Theoretically RDD sampling could be extended to mobile telephones. In practice, this may be quite difficult due to the fact that mobile telephones are by nature a personal appliance, rather than a household one. Sampling persons within a household, via a mobile telephone contact with one of the members, is well nigh impossible. Interviewing via a mobile telephone of individuals who may be anywhere is also extremely difficult. Even the determination of the total number of telephone numbers (mobile and fixed line) available to a household (required for weighting) may be daunting. We consider some possible approaches to these and other problems of the move to mobile telephones in section four.

Undercoverage of persons within covered households relates primarily to the method of selection for individuals within the household – see section 3.2.2 – and to the undercoverage due to the failure to obtain complete listings of individuals in the households. The latter effect is investigated by Maklan and Waksberg (1988), by comparing data on individuals obtained from an RDD survey with those obtained from the US Current Population Survey and from the population census. They find that while mean household sizes are comparable, the RDD results are skewed towards two-person households and away from one-person households. Some of the difference could be attributed to different residence rules, but the results do not indicate undercoverage of persons in the RDD survey. They also report on an experiment in which more detailed questions were asked on household composition and found practically no improvement in accuracy of reporting. In a similar experiment, carried out by Bercini and Massey (1979), the effects of the use of names in the household roster and the position of the question on the household roster (before or after the first interview) were tested in a survey on smoking. They found that both the use of names and the position of the household roster had an effect on response and that obtaining the roster after the interview without names is optimal.

3.3.2 Nonresponse

The problem of nonresponse and the biases associated with nonresponse is basic to all survey research, but there are some specific issues of nonresponse associated with telephone interviewing. One of the major problems is the ambiguity of the results of many attempts at dialing – *e.g.*, continually engaged or no reply, numbers connected to fax machines, computer modems or answering machines. Recently automated screening devices have been developed to identify telephone numbers connected to recordings indicating whether they are not in service (Casady and Lepkowski 1999). Thus proprietary hardware and software have been developed to detect "tri-tone" recording which indicates "not-in-service" and these numbers when dialed can be removed from the sample. Prior removal of many business phones can be carried out by matching with "Yellow Page" files. These and other methods reduce the costs of screening and the ambiguity of calls that continually receive no reply.

Technological advances, such as "call forwarding" and caller identification enhance the possibilities for non-response. In addition refusals are easier over the phone than in face-to-face interviews and breaking off the interview in its midst is also easier. These and other problems of nonresponse for 'cold' telephone interviewing and the US experience in dealing with them are reviewed extensively by Groves and Lyberg (1988a). In particular they follow CASRO (1982) and White (1983) in recommending a definition of nonresponse rates which includes in the denominator an estimate of the number of unanswered numbers that are working numbers in addition to the complete and incomplete interviews, refused eligible numbers and other noninterviewed units. The estimate of the proportion of unanswered numbers that are eligible is obtained as the proportion of answered numbers that are eligible. However this may be a biased estimator. For instance the intensive use of answering technology by businesses implies that practically all businesses will respond and can be identified as businesses. Also, as pointed out by Massey (1995), this measure has to be modified in the case of screening by defining a household screening response rate as the estimated proportion of eligible households identified as such by the screening, rather than the proportion of all households screened for eligibility. Cunningham, Brick and Meader (2000) present several detailed measures of response rates and eligibility rates for each stage of a survey with screening, as well as overall rates, in reporting on the methodology of the National Survey of America's Families.

Telephone nonresponse rates are, in general, higher than those obtained from face-to-face interviews, due to the reasons mentioned above – see Hochstim (1967), Groves and Kahn (1979), Fitti (1979), Groves and Lyberg (1988a) for US experience; Wilson, Blackshaw and Norris (1988), and Collins, Sykes, Wilson and Blackshaw (1988) for experience in UK surveys; and Drew, Choudry and Hunter (1988) for the experience of Canadian government surveys. The latter includes also comparisons of 'cold' and 'warm' telephone interviews, which show only small differences in nonresponse rates. More recently an analysis of the experience in 39 US telephone surveys carried out in the nineties (Massey, O'Connor and Krotki 1997) indicates a slight further reduction in response rates to an average of 62% and a range from 42% to 79% (though it seems that Canadian response rates have not decreased over recent years). Among the factors to which this increase in nonresponse can be attributed are the increase in the use of technological devices (answering machines, call forwarding, multi -purpose telephone lines) and the increased prevalence of telephone solicitation, already identified as a potential problem for telephone surveys by Biel (1967). The American Statistical Assocation (1999) considers the effect of near saturation calling conducted by telemarketers on lowering survey cooperation rates as a serious challenge not fully addressed by survey researchers. It concludes that unless the trend can be reversed, "telephone surveys, as we know them, could disappear within the next five years". A similar view is expressed by Kalton (2000).

As is the case for telephone noncoverage, the effect of nonresponse on biases in survey estimates is made more severe by the correlation between nonresponse and many socio-economic characteristics. Groves and Lyberg (1988a) on the basis of a review of previous work identify the main correlates of telephone nonresponse. They are age (elderly persons have higher refusal rates – see also Collins *et al.* 1988) and education (higher nonresponse among lower education groups - see, *e.g.*, Cannel, Groves, Magilavy, Mathiowetz, Miller and Thornberry 1987). On the other hand, there is evidence showing that urban-rural differences in nonresponse are diminished in telephone surveys, as compared with face-to-face surveys – Groves and Kahn (1979). More recent papers on the effects of nonresponse concentrate on specific issues. Thus Diehr, Koepsell, Cheadle and Psaty (1992) investigate the relationship of response rate and other summary variables at the prefix and at the person level. They find relationships between nonresponse and age, race and family size and type. Merkle, Bauman and Lavrakas (1993) in an investigation of the impact of callbacks on the quality of survey estimates show that age and employment status are the major correlates with the number of callbacks required. Kalsbeek and Durham (1994) investigate the effect of nonresponse in a follow-up telephone survey on breastfeeding among low-income women and find that the main correlates with nonresponse are age and degree of urbanization. Finally, multilevel modeling is applied to an extensive meta-analysis of reports on inter-mode comparisons of nonresponse by Hox, DeLeeuw and Kreft (1991). The results, based on the analysis by multi-level modeling of a total of 45 studies (35 of which included a telephone component), indicate significantly lower response for telephone studies than for face-to-face studies when models with fixed slopes are used. However when random-slope models are used the difference is no longer significant.

In attempts to reduce nonresponse in telephone surveys the effect of survey operational variables on nonresponse has been investigated. Thus Sebold (1988) finds that doubling the survey period (from two to four weeks) increased the response rate by 3 percentage points in an experiment for the US National Crime Survey. Brick and Collins (1997) investigated the effect of advance letters and screening questions on response in the US National Household Education Survey. They found that a screen-out question approach increased response rates considerably but that the advance letter did not add to the effect of screening. Other survey variables that have been found to affect response rates are interview length (Collins, *et al.* 1988) and interviewer vocal characteristics (Oksenberg and Cannel 1988). The effect of the method of selection of sample individuals on nonresponse (in particular the requirement for household rosters) has already been mentioned in section 3.2.2.

Finally, in recent years there has been a significant increase in the use of answering machines and caller ID devices for screening unwanted calls, with obvious increased potential for nonresponse. For instance, the proportion of households with answering machines in France increased from 21% in 1995 to 40% in 1999 (Rouquette 2000), the same as in Germany (Federal Republic of Germany 1999), while in the US the proportion increased from about 25% in 1988 (Tuckel and Feinberg 1991) to over 73% by 1997 (Decision Analyst 1997). However, based on a nationwide telephone survey, Tuckel and Feinberg (1991) reach the conclusion that, in comparison to other initial non-contact groups (*e.g.*, 'no answer' or 'busy'), those with answering machines are more likely to respond and less likely to refuse, resulting in a contact rate which is definitely not smaller than that of other non-contacts. In fact, it seems, according to a study by Oldendick and Link (1994), that the use of answering machines to screen out survey calls is limited to some 2-3 percent. However screeners tend to be in higher income groups, urban and with higher education. Similarly, Piazza (1993) finds on the basis of extensive data from the California Disability Survey, a telephone survey with a high number of callbacks, that although answering machine owners are more difficult to contact initially, once contacted they are at least as likely to respond as those without answering machines. They point out also that reaching an answering machine ensures that a household has been reached and that its residents do not want to miss important calls. In a study by Xu, Bates and Schweitzer (1993), designed to investigate the effect of leaving messages on answering machines, households with answering machines were found to be more likely to be contacted and to complete the interview than those without answering machines. Furthermore leaving a message on the answering machine led to a significant increase in response rate and reduction in refusals. Similarly, Harlow, Crea, East, Oleson, Fraer and Cramer (1993), based on results of a controlled experiment, found that leaving a message on the answering machine led to an increase of 15% in response, after adjusting for age, interviewer and town of residence. Koepsell, McGuire, Longstreth, Nelson and van Belle (1996) carried out a randomized trial of leaving messages on answering machines and found an overall increase of 20% in response rate. Although in a similar study Tuckel and Shukers (1997) found no significant effect, the overall findings in a range of studies indicate that the increase in the use of answering machines has a beneficial effect on survey response, probably due to their providing the possibility of leaving positive messages and thereby enabling the screening out of telemarketing calls.

Tuckel and O'Neill (1996) estimate that the percentage of US households with caller ID increased from 3% in 1992 to 10% in 1996. Based on a national study, in which the profiles of both caller ID subscribers and answering machine owners are analyzed, they reach the conclusion that these technological devices do not yet present major obstacles for telephone survey research, since their owners tend to use the screening devices primarily to screen out recognized undesirable numbers of acquaintances rather than unrecognized numbers. However, they point out that the possibility of screening will probably lead to increases in answering machine response to repeated callbacks.

3.3.3 Weighting and adjustment

Telephone surveys often require special attention to weighting and adjustment. Although sampling designs are usually based on equal probabilities of selection, in practice these are not always achieved. For instance RDD sample designs are theoretically self-weighting but in fact unequal selection probabilities may result due to the multiplicity of telephone lines (numbers) for the same household. In this case, if information is collected on the number of telephone lines to which the household is connected, the required adjustment is straightforward. Similarly reweighting is required to take into account PSU's for which the number of in-scope numbers is less that the required cluster sample size. An additional problem arises due to the fact that it is often difficult to determine whether a telephone, from which no answer can be obtained after repeated attempts, is indeed a case of in-scope nonresponse or is, in fact, out-of-scope. Other problems requiring reweighting are nonresponse, the inherent undercoverage due to non- telephone households and the obvious necessity to use some form of multiplicity estimator for multiple-frame sample designs, based on information on the frames on which the unit is represented.

These problems are dealt with for national RDD samples carried out by the US National Center for Health Statistics in a series of papers by Thornberry and Massey (1978); Botman, Massey and Shimizu (1982); and Massey and Botman (1988). They describe the weighting adjustments carried out for the RDD US National Health Interview Survey (NHIS) and for a smoking survey to account for multiple telephones per household, for telephone coverage and for nonresponse. The adjustments were based on external data for race and geographic region and on survey information on nonresponse and on multiple telephones. Several alternative adjustment and weighting procedures are compared and evaluated. Chapman and Roman (1985) compare substitution with nonresponse adjustment in a feasibility study for the RDD NHIS and report that the results with respect to bias and variance are similar. Drew and Groves (1989) compare alternative adjustment procedures for unit nonresponse based on external administrative data, on an explicit response prediction model and on response probabilities estimated on the basis of callback data. Casady and Sirken (1980) propose a multiplicity estimator for a multiple-frame sampling design applied to data from the US National Health Interview Survey. Brick (1990) compares the multiplicity estimator with the traditional multiple frame estimator for an educational RDD survey.

Goksel, Judkins and Mosher (1991) report on adjustments, based on modeling nonresponse propensities, for a telephone follow-up of a face-to face interview in the US National Survey of Family Growth. Adjustment based on response propensities by intensity of follow-up effort and by smoking status are proposed for a Canadian survey of attitudes to smoking restrictive legislation by Bull, Pederson and Ashley (1988).

Following a comparison by Keeter (1995) of non-telephone households with 'transient'households (those who recently gained or lost telephone service), Brick, Waksberg and Keeter (1996) propose the use of data on interruptions in telephone service in order to adjust for the undercoverage due to non-telephone households. Their results indicate that such adjustment can lead to a reduction of mean square error. Hoaglin and Battaglia (1996) compare a modified poststratification method and a model-based estimation with simple poststratification for adjusting for noncoverage in an RDD survey of vaccination coverage. The modified poststratification uses national data on vaccination rates for telephone and non-telephone children in addition to demographic and socioeconomic data used for simple poststratification, while the model- ased adjustment is based on a logit model to estimate the probability of residing in a telephone household. The results show gains from the use of the modified poststratification but only slight differences between the modified poststratification and the model based adjustment. A similar adjustment based on telephone interruption data is applied by Frankel, Srinath, Battaglia, Hoaglin, Wright and Smith (1999) to NHIS data and shows conclusively a substantial reduction in bias.

3.4 Data Quality – Response Errors and Mode Effects

The quality of information obtained over the telephone has always been a controversial issue. As mentioned in section 2, apprehensions on the supposed inferiority of the quality of data from telephone interviewing were allayed at an early stage, to a large degree by some of the extensive empirical appraisals carried out in the sixties and seventies. However there was still some conflicting evidence from different studies on the relative quality of telephone and face-to-face interviewing. Although the intensive analysis of large omnibus surveys carried out under the two modes by the University of Michigan Survey Research Center (Groves and Kahn 1979), provided important information on data quality and other issues, the mode comparisons and a comparison with external data were not conclusive. In an attempt to resolve the issue, de Leeuw and van der Zoowen (1988) carried out an extensive meta-analysis of 28 major empirical studies in which comparisons of face-to-face and telephone interviewing were investigated. The studies, carried out between 1952 and 1986 on a variety of topics, were primarily from the US but some European studies were also covered. Data quality indicators used were response validity (based on validation studies), absence of social desirability bias, item response, amount of information (for open questions or check-lists) and similarity of response. The overall finding is that if there are any differences in quality between the two modes, they are definitely very minor and that other considerations, such as costs and convenience, should be used in decisions on the use of the telephone for survey work. Similar conclusions are reached for the UK by Sykes and Collins (1988), on the basis of four comparative studies; for income data in Denmark by Körmendi (1988), in a validation study, based on administrative data; and in a comparison of financial data in a Canadian Farm Financial Survey (Caron and Lavallée 1998).

Other recent studies on mode effects concentrate on specific issues and topics but reach similar conclusions. Thus Herzog and Rodgers (1988) report on a mode comparison in a study of older adults and find only small differences. Similar results are reported by Foley and Brook (1990) for a survey on the last days of life. In a study of the sensitive topic of drug use Aquilino and Lo Sciuto (1990) find almost identical results for whites, but some significant differences for blacks, even after controlling for variables possibly related to telephone undercoverage. This may be explained by results reported by Johnson, Fendrich, Shaligram and Garey (1997) for a telephone survey of drug use, which supports a social distance model of interviewer effects.

There is little doubt that interviewers have a great effect on quality, both in face to face and in telephone surveys. The use of central telephone interviewing facilities provides more opportunities to control and monitor interviewer effects than in field interviewing. Some of the issues involved are treated by Stokes and Yeh (1988), who propose a Bayesian model for interviewer effects and methods for estimating the model parameters. A beta-binomial model for the interviewer variance component and methods of estimation of its parameters are proposed by Pannekoek (1988).

An effective way of reducing response errors in face-to-face interview surveys has been the use of records provided by the respondent to verify and recall information on income, insurance, health events *etc.* Obviously, the extension of this method to telephone interviewing involves some problems, since the interviewer cannot see the documents and even asking the respondent to get them may involve a disruptive break in the telephone interview more frequently than in a face-to-face interview. However the use of records by respondents in telephone surveys can help to reduce response bias. Battaglia, Shapiro and Zell (1996) report on an attempt to ask respondents to use vaccination records in one of the rounds of the US National Immunization Survey and to compare the information obtained with provider records. Some 47% of the respondents did in fact use vaccination records but substantial underreporting bias was still found, possibly due to the fact that the vaccination reports were not always up to date. Similar effects are found in face-to-face surveys – see Brick, Kalton, Nixon, Givens and Ezzati-Rice (2000).

4. Current and Future Technological Developments

Together with almost complete telephone coverage, the very intensive technological development and the diversity of communications possibilities are continuously opening up new opportunities and potentials for using novel communication options for survey work. On the other hand, some of these developments may cause difficulties for telesurveys under the conventional methodology of today. Thus the increased sophistication of filtering devices and algorithms (as a development of the simple answering machines and caller ID devices mentioned in section 3.3) may make it easier than ever for respondents not to cooperate. In the following we examine present applications and conjectured future developments and comment on the methodological problems involved in their use.

4.1 E-Mail and Web Surveys

Internet access for households has experienced a very rapid increase in recent years. For instance in the US the proportion of households with access to the Internet has risen from 26% in December 1998 to 42% in August 2000 – NTIA (2000). Other countries have reached somewhat lower levels – the UK 28% (in August, 2000 – OFTEL 2000), Canada 25%, Finland 22%, France 7% and Belgium 5% in 1999, according to Rouquette (2000), Israel 12% (in 1999 – Central Bureau of Statistics 2000) and Germany 11% (Federal Republic of Germany 1999). This rapid increase in coverage, is still far off from attaining completeness. Furthermore, there are also some indications that, together with the increase in total use, there is also a growing category of ex-users. Katz and Aspden (1998) report that the proportion of former users of the Internet increased from 8% to 11% between 1995 and 1996. However the overall increase in access has encouraged the use of e-mail and the Internet for survey work. While coverage for an e-mail survey (EMS) is comparable to that of a Web (or Internet) survey and both are based on the use of a computer self-administered questionnaire (CSAQ), there is a basic difference between these two types of telesurveys. The e-mail survey is very similar to a mail survey, in that it is based on sending out a text questionnaire and asking the respondent to send back the completed questionnaire. The advantage over the mail survey is in cost and in the ease and simplicity of transmission and receipt. The Web survey is, in general, based on interaction between the respondent and the survey instrument, via the use of Java, XML, or a similar instrument. It allows multiple enhancements, such as colour and animation, and extensive possibilities for sophisticated skip patterns and real-time editing. The exciting potential for innovative collection systems based on ever-developing Web tools cannot yet overcome the basic problem inherent in both e-mail and Web surveys that current coverage is completely inadequate for most human populations of interest (Dillman 2000).

Nonetheless, e-mail and Internet surveys can and are being used, with varying degrees of success, for certain populations where coverage is virtually complete or in conjunction with other modes of collection. Thus Couper, Blair, and Triplett (1999) report on an experimental study comparing e-mail and regular mail for a survey of employees in several U.S. government statistical agencies. The sampled employees were randomly assigned to a mail or e-mail mode of data collection and comparable procedures were used for advance contact and follow-up of subjects across modes. The results indicated somewhat higher response rates for mail than for e-mail, but data quality (item missing data) was similar across the two modes. In field tests for the 1999 US National Study of Postsecondary Faculty both administrators and faculty were offered the choice between completing and mailing a conventional paper questionnaire or completing a CSAQ via the Web (Abraham, Steiger and Sullivan 1998). Although it may be assumed that practically all respondents had access to the Web, only 8% of responding faculty and 17% of the institution administrators opted for the CSAQ mode. The US National Science Foundation is planning to use a Web-based option in its 1999 National Survey of Recent College Graduates, under the hypothesis that most of the survey population would be relatively computer literate and have access to the Web (Meeks, Lanier, Fecso and Collins 1998). For a review of the use of CSAQ by government agencies and private survey organizations and the problems involved, see Ramos, Sedivi and Sweet (1998).

However, most current Web surveys of general populations are based on non-probability sampling – mostly by some form of self-selection. Fischbacher, Chappel, Edwards and Summerton (1999) report on a meta-analysis of 28 surveys in the health field using e-mail and the Internet. Many of these were epidemiological studies aimed at patients of specific diseases and the problem of selection bias meant that most of the results could not be generalized. One of the largest Web surveys is the WWW User Survey carried out by the Graphics Visualization and Usability

Center at Georgia Institute of Technology (Kehoe, Petkow, Sutton, Aggarwal and Rogers 1999). Although the survey population is defined as Internet users, the lack of any sample framework for this population implies that respondents had to be solicited by various methods (Web and other media announcements, advertising banners, incentive cash prizes *etc.*), rather than sampled with known probabilities. Although some 20,000 users participated, the survey report points out that the data is biased towards experienced and more frequent users and recommends the augmentation of their data with random sample surveys. In an attempt to overcome the bias inherent in basing surveys on samples of those with internet access only, some commercial survey organizations distribute devices, which let users access the Internet through television sets, to all of its panelists on an RDD sample, to ensure consistent results (Felson 2001). However Poynter (2000) predicts that by the year 2005 95% of market research surveys will be conducted via the internet but that 80% will be based on respondents who have 'opted in', rather than on probability sampling.

On the other hand, there is evidence that Web-based data collection can be applied with relative success for establishment surveys. Nusser and Thompson(1998) report on its use for the US Department of Agriculture's National Resources Inventory Surveys; Rosen, Manning and Harrel (1998) on Web-based collection from establishments for the US Current Employment Statistics Survey and Meeks *et al.* (1998) on its use for data collection from academic institutions, federal agencies and private corporations for US National Science Foundation surveys. Assuming that the problem of coverage and sampling will eventually be resolved for households and individuals, this holds hope for Web-based collection for household surveys at some point in the future.

4.2 Other Computer Self Administered Questionnaire (CSAQ) and Computer Assisted Self Interviewing (CASI) Methods

Couper and Nichols (1998) differentiate between computer self administered questionnaire (CSAQ) collection, in which an interviewer is not present, and computer assisted self interviewing (CASI), in which an interviewer is present or delivers the survey instrument. Thus both e-mail and Internet surveys are based on CSAQ with the assistance of telecommunications technology. Other CSAQ methods are touchtone data entry (TDE), whereby respondents enter data using their touchtone telephones, and interactive voice recognition (IVR) or voice recognition entry (VRE). Both are based on respondents initiating calls to report at their convenience, after initial contact has been established, and have been extensively tested and successfully used by the US Bureau of Labor Statistics for data collection from establishments for its Current Employment Statistics program – Werking, Tupek and Clayton (1988), Winter and Clayton (1990) and Clayton and Winter (1992). Phipps and Tupek (1991) report on a study of the quality of TDE collection, by means of a record check. Their results show that there are few problems with the method and that response errors diminish with experience. More recently US statistical agencies have initiated tests of the possibility of applying these CSAQ methods to household surveys. McKay, Robison and Malik (1994) report on initial laboratory testing of TDE for the Current Population Survey. Malakhoff and Appel (1997) report on the development of an IVR prototype at the US Bureau of Census, albeit for a listing operation by field staff. It should be noted that while TDE is obviously unique to telephone surveys, IVR could be used for other modes of collection.

Computer assisted self interviewing (CASI) methods include audio (ACASI) and video (VCASI) modes of collection and have long been regarded as the natural extensions of mail surveys that benefit from modern day technology (Dillman 2000). Their usefulness has been especially emphasized for surveys of sensitive and embarrassing topics, where the presence of the interviewer during the interview may make respondents reluctant to answer in a face-to-face interview. For a review of recent advances in these methods see Baker (1998), O'Reilly, Hubbard, Lessler, Biemer and Turner (1994), Rogers, Miller, Forsyth, Smith and Turner (1996) and Tourangeau and Smith (1998). Practically all the reported applications are of surveys in which the survey instrument is brought to the respondent's home by field staff. The use of the telephone for ACASI (T-ACASI) collection has already been tried – Turner, Forsyth, O'Reilly, Cooley, Smith, Rogers and Miller (1998). The long-expected development of videotelephony to become a widespread common form of telephone service for households has not yet materialized. If and when it occurs it should make telephone VCASI (T-VCASI) possible in the future, with important implications for telesurvey work. The addition of a visual element will help to overcome many of the problems of present day telephone surveys that are not present in face-to-face interviews (eye contact with the interviewer, use of cue cards and other visual aids). The use of videotelephony will probably not be universal for a very long time, so that at least for the time being, T-VCASI will only be able to serve as a supplementary mode of collection.

4.3 Mobile Telephones

The problems envisaged for coverage of fixed line RDD surveys due to the rapid proliferation of mobile telephones have been mentioned in section 3.3.1. In the future it is obvious that mobile telephones will have to be used to reach the ever-increasing numbers of households without fixed telephone lines. Present levels of mobile telephone coverage imply that mobile telephone surveys can, in general, only be used for specific populations or for supplementing fixed line RDD surveys. For instance Perone, Matrundola and Soverini (1999) report on a mobile telephone survey for a naturally accessible population - that of mobile telephone subscribers in order to assess customer satisfaction. Refusal rates were found not to exceed those found in fixed line telephone surveys. However, non-contact rates were high, primarily due to subscribers being outside the signal range or shutting down their telephones. An additional problem associated with mobile phone surveys is that in many cases in North America the subscriber has to pay for received calls – Casady and Lepkowski (1999).

As mentioned above, Cunningham, *et al.* (1997) report on the use of mobile telephones to interview nontelephone households (primarily in rural areas), with the mobile telephone brought to the respondent by field interviewers. This was designed to minimize mode effects by having telephone interviews conducted by the same interviewers as those conducted for telephone households. The response rates were high, even though in some cases the interviews had to be conducted outdoors in order to obtain reasonable reception. The most intensive use of mobile phones for household surveys is no doubt for the Finish Labour Force Survey – Kuusela and Notkola (1999). Out of some 97% of interviews completed by telephone, over 20% are carried out by mobile telephone. Although the average duration of mobile telephone interviews is somewhat longer than those of conventional telephone interviews, this is probably due to socio-demographic differences between the respondent groups.

4.4 Future Technological Developments and their Effect on Telesurvey Methodology

The rapid advances in technological developments in the areas of telecommunications and information systems make it very difficult to forecast their influence on survey work. Not all these technological changes will necessarily increase the potential for using advanced telecommunications technology for survey work. The problems raised by persons who have opted to ‘drop-out’ from the Internet (Katz and Aspden 1998) or from fixed line telephone service (see *e.g.*, Gabler and Haeder 2000; and Kuusela and Vikki 1999) have already been mentioned. Furthermore, in some areas, such as market research and official statistics, technological developments may lead to a reduced reliance on surveys to gather information for decision-making. Thus Baker (1998) and Poynter (2000) predict that techniques such as data mining of existing data resources may become predominant for market research. Similarly, Scheuren and Petska (1993) discuss the possibilities for the use of administrative record systems for official statistics. However, there still remain important areas (for instance for opinions and unobservable behaviour) in which surveys will remain the predominant source of data. The technological advances will open new possibilities for telesurvey work, though the required methodology might become more complex than that used today.

One of the expected developments forecast for the near future is the integration of multiple communication devices and methods – telephony (fixed line and wireless), fax, internet, e-mail, videotelephony, data transmission, television transmissions *etc.* – Baker (1998). This implies that each individual will have access to a variety of telecommunication services possibly via the same physical instrument, which could be a mobile phone (*e.g.*, via WAP technology), a PC or a TV set or some combination of these. Similarly, the survey taker may be able to gain access to respondents via several different modes. See Ranta-aho and Leppinen (1997) for some of the issues involved in this plethora of possible avenues of access. It is envisaged that the recipient will have a large degree of control over whether to receive communications at all and, if, so by which mode. This is already now ensured for many users by means of sophisticated devices for screening, forwarding, message transfer, multiple message transmission *etc*. On the other hand, the degree of control of mode of transmission by the sender will probably decrease as a result.

The implications of these developments for survey work are that mixed mode surveys and possibly multi-frame methodology will have to become predominant. Although we consider that overall telecommunications coverage will increase to some saturation point that is close to universal coverage, it seems unlikely that any given mode of telecommunication will by itself provide virtual complete coverage. Furthermore, even when a single mode may provide practically complete coverage, it is not clear that a mixed mode approach, taking into account respondents’ mode preferences, is not preferable. The increased reliance of survey work on the voluntary cooperation of respondents practically dictates that we should offer the respondent the choice of mode. However it should be pointed out that mixed mode surveys are very expensive and that the present technology does not allow the simple transfer of questionnaires developed for one mode (*e.g.*, the CAI Blaise questionnaire) to another mode – *e.g.*, to a paper form.

The major problem that the new developments in telecommunications pose for survey design will probably be the choice of relevant frameworks and the allocation of sample units to modes of collection. Eventually it is envisaged that each individual will have a unique, permanent, personal communication number (or ID) through which he/she can be reached by a multiplicity of modes (written, oral or visual), via a variety of fixed line or wireless devices which could be at home, in the office or mobile. The choice of mode will be ultimately controlled by the joint decision of recipient and sender. While the idea of such a universal number (which would basically be an identity number) is no doubt anathema to libertarians, there is little doubt that it will eventually become acceptable, even if small activist groups may attempt to evade its use and even disrupt its proliferation. In fact standard universal identity number systems have been operating and are well accepted for several decades in many countries in Northern Europe and in Israel. The identity number in these countries is not regarded as confidential information and is widely used for many administrative and commercial purposes. For example, in Israel personal cheques are required by law to include the person's ID number, name, address and telephone number.

Once such a system of unique communication numbers is operable, standard methods of sampling can be used. It may well be that complete lists of these numbers will be generally available – possibly with only limited geographical or other information. This is the situation with respect to ID's in many national registration systems. There are reasons to expect that a similar situation may prevail for communication numbers – initially at least in Europe rather than in North America. This could come about since the need for unlisted status might well be made redundant because of sophisticated screening techniques. Although screening may enhance the ease of non-response, the possibility of transmitting prior written messages by e-mail or voice mail could reduce the problem.

Sampling from such lists would be simple but in most cases might be inefficient, since it could benefit only marginally from auxiliary information. While differentiation between personal and business contacts might be ensured by the listings, it is doubtful that any household information would be available. This dictates that the sampling and reporting unit would be the individual rather than a household. This is in any case the aim of many surveys and the usefulness of the household as a sampling unit for telesurveys is definitely doubtful, even under current practice. Household information, if required, would have to be obtained from the individual and include information on household size to ensure proper weighting for household characteristics. If the communications numbering system ensures the allocation of a single number to each individual, no information is required on the modes of communication or their multiplicity.

If listings of communication numbers are not available or if the problem of unlisted numbers does persist, some form of RDD will have to be used. This should not differ much from the RDD techniques currently employed. Assuming that the communication numbering system is indeed unique and universal and also arranged by some logic, efficient methods for sampling could easily be developed. Hopefully the numbering system will still bear some relationship to geography, via the individual's permanent address. Otherwise local or even national RDD surveys will become extremely difficult to design efficiently. If sufficient information on the numbering system is available, the extent of out-of-scope numbers could be minimized.

Since it is likely that choice of the mode of communications will be largely under the control of the recipient, the question of allocation of sample units to mode of communication will probably hardly arise. The survey taker will have to prepare a whole range of collection instruments suitable for the different modes of communication. These would have to include written instruments, such as faxed, e-mail and Internet versions of questionnaires, oral instruments, such as traditional voice interviews and automated interviewing, and combinations of these. The integration of the data obtained from these modes of collection into a uniform data set would be a formidable but surmountable technological challenge.

The almost utopian situation described above will probably take a long time to reach and in the interim suitable methodologies will have to be developed to deal with the problems arising from the short-term developments in communications technology and their application. The necessity to move from telephone surveys based uniquely on fixed line telephones to some combination of mobile and fixed-line telephone situation will have to be dealt with very shortly, as pointed out in section 4.3. Basically multiple frame methodology developed to cover both telephone households and non-telephone households can easily be extended to deal with this. The development of suitable frames and/or RDD sampling methods for mobile telephones still has to be carried out, but the necessary principles are available. The problem of combining data obtained from mobile phones which are basically personal devices with that obtained from fixed-line telephones, which are still fundamentally household devices, will have to be worked out to ensure proper weighting. To ensure this, sufficiently complete information on all the communication devices available to the household is required.

In conclusion, the advances in telesurvey methodology over the past few decades, which have made telephone surveys a viable and predominant survey instrument, will have to be continually updated to deal with the ever-changing developments in telecommunications technology and it usage. However the basic elements for these new developments are available and will continue to allow the use of advanced options to obtain high quality survey data

References

Abraham, S.Y., Steiger, D.M. and Sullivan, C. (1998). Electronic and mail self-administered questionnaires: A comparative assessment of use among elite populations. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 833-841.

Alexander, C.H. (1988). Cutoff rules for secondary calling in a random digit dialing survey. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al*.). New York: John Wiley & Sons, Inc., 113-126.

American Statistical Association (1999). *More About Telephone Surveys*. ASA series: What is a survey? Section on Survey Research Methods [http://www.amstat.org/sections/srms/brochures/telephone.pdf].

Anderson, J.E., Nelson, D.E. and Wilson, R.W. (1998). Telephone coverage and measurement of health risk indicators: Data from the National Health Interview Survey. *American Journal of Public Health*, 88, 1392-1395.

Aquilino, W.S., and Lo Sciuto, L.A. (1990). Effects of interview mode on self-reported drug use. *Public Opinion Quarterly*, 54, 362-395.

Baker, R.P. (1998). The CASIC future. *Computer Assisted Survey Information Collection*, (Eds., M.P. Couper, *et al.*). NewYork: John Wiley & Sons, Inc., 583-604.

Banks, M.J., and Hagan, D.E. (1984). Reducing interviewer screening and controlling sample size in a local-area telephone survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 271-273.

Banks, R., Christie, C., Currall, J., Francis, J., Harris, P., Lee, B., Martin, J., Payne, C. and Westlake, A. (Eds.) (1999). ASC'99 - Leading Survey & Statistical Computing into the New Millennium. *Proceedings of the ASC International Conference*. Association for Survey Computing Chesham, Bucks, UK.

Battaglia, M.P., Shapiro, G. and Zell, E.R. (1996). Substantial response bias may remain when records are used in a telephone survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 452-455.

Bennet, C.T. (1961). A telephone interview: A method for conducting a follow-up study. *Mental Hygiene*, 45, 216-220.

Bercini, D.H., and Massey, J.T. (1979). Obtaining the household roster in a telephone survey: The impact of names and placement on response rates. *Proceedings of the Social Statistics Section*, American Statistical Association, 136-140.

Berry, S.H., and O'Rourke, D. (1998). Administrative designs for centralized telephone survey centers: Implications of the transition to CATI. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 457-474.

Biel, A.L. (1967). Abuses of survey research techniques: The phony interview. *Public Opinion Quarterly*, 31, 298.

Biemer, P.P. (1983). Optimal dual frame sample design: Results of a simulation study. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 630-635.

Binson, D., Canchola, J.A. and Catania, J.A. (2000). Random selection in a national telephone survey: A comparison of the Kish, next-birthday, and last-birthday methods. *Journal of Official Statistics*, 16, 53-59.

Blair, J., and Czaja, R. (1982). Locating a special population using random digit dialing. *Public Opinion Quarterly*, 46, 585-590.

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

Blankenship, A.B. (1977b). Listed versus unlisted numbers in telephone-survey samples. *Journal of Advertising Research*, 39-42.

Botman, S.L., and Allen, K. (1990). Some effects of undercoverage in a telephone survey of teenagers. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 396-400.

Botman, S.L., Massey, J.T. and Shimizu, I.M. (1982). Effect of weighting adjustments on estimates from a random-digit-dialed telephone survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 139-144.

Brick, J.M. (1990). Multiplicity sampling in an RDD telephone survey. *Proceedings of the Section on Survey Research Methodology*, American Statistical Association, 296-301.

Brick, J.M., and Collins, M.A. (1997). A response rate experiment for RDD surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1052-1057.

Brick, J.M., Kalton, G., Nixon, M., Givens, J. and Ezzati-Rice, T. (2000). Statistical issues in a record check study of childhood immunizations. *Proceedings of the 1999 Federal Committee on Statistical Methodology Research Conference*: *Statistical Policy Working Paper*, 30, 625-634.

Brick, J.M., and Waksberg, J. (1991). Avoiding sequential sampling with random digit dialing. *Survey Methodology*, 17, 27-42.

Brick, J.M., Waksberg, J. and Keeter, S. (1996). Using data on interruptions in telephone service as coverage adjustments. *Survey Methodology*, 22, 185-197.

Brick, J.M., Waksberg, J., Kulp, D. and Starer, A. (1995). Bias in list-assisted telephone samples. *Public Opinion Quarterly*, 59, 218-235.

Brunner, J.A., and Brunner, G.A. (1971). Are voluntary unlisted telephone subscribers really different? *Journal of Marketing Research*, 8, 121-124.

Bryant, B.E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research*, 12, 129-135.

Bryson, M.C. (1976). The literary digest poll: Making of a statistical myth. *The American Statistician*, 30, 184-185.

Bull, S.B., Pederson, L.L. and Ashley, M.J. (1988). Intensity of follow up; effects on estimates in a population telephone survey with an extension of Kish's (1965) approach. *American Journal of Epidemiology*, 127, 552-561.

Burke, J., Morganstein, D. and Schwartz, S. (1981). Toward the design of an optimal telephone sample. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 448-453.

Cahalan, D. (1960). Measuring newspaper readership by telephone: Two comparisons with face to face interviews. *Journal of Advertising Research*, 1, 2, 1-6.

Cahalan, D. (1989). Comment: The digest poll rides again! *Public Opinion Quarterly*, 53, 129-133.

Campbell, J., and Palit, C.D. (1988). Total digit dialing for a small area census by phone. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 549-551.

Cannell, C.F., Groves, R.M., Magilavy, L.J., Mathiowetz, N.A., Miller, P.V. and Thornberry, O.T. (1987). An experimental comparison of telephone and personal health interview surveys. *Vital and Health Statistics*, Series 2, 106, Public Health service.

Caron, P., and Lavallée, P. (1998). Comparison study on the quality of financial data collected through personal and telephone interviews. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 208-213.

Casady, R.J., and Lepkowski, J.M. (1991). Optimal allocation for stratified telephone survey design. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 111-116.

Casady, R.J., and Lepkowski, J.M. (1993). Stratified telephone survey designs. *Survey Methodology*, 19, 103-113.

Casady, R.J., and Lepkowski, J.M. (1998). Telephone Sampling. *Encyclopedia of Biostatistics*. New York: John Wiley & Sons, Inc., 4498-4511.

Casady, R.J., and Lepkowski, J.M. (1999). Telephone Sampling. *Sampling of Populations*: *Methods and Applications - third edition*, (Eds., P.S. Levy and S. Lemeshow). New York: John Wiley & Sons, Inc., 455-479.

Casady, R.J., and Sirken, M.G. (1980). A multiplicity estimator for multiple frame sampling. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 601-609.

Casro (1982). *Report of the Council of American Survey Research Organization Completion Rate Task Force*. New York: Audits and Surveys Inc. (Unpublished report).

Central Bureau of Statistics (2000). *The Household Expenditure Survey 1999*. Special Publication 1147. Jerusalem.

Chapman, D.W., and Roman, A.M. (1985). An investigation of substitution for an RDD survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 269-274.

Choudhry, G.H. (1989). Cost-variable optimization of dual frame design for estimating proportions. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 566-571.

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, 377-388.

Collins, M. (1983). Computer assisted telephone interviewing in the UK. *American Statistical Association*, *Proceedings of the Section on Survey Research Methods*, 636-639.

Collins, M. (1999). Editorial: Sampling for UK telephone surveys. *Journal of the Royal Statistical Society*, A 162, 1-4.

Collins, M., Sykes, W., Wilson, P. and Blackshaw, N. (1988). Nonresponse: The UK experience. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 213-232.

Colombotos, J. (1965). The effects of personal *vs*. telephone interviews on socially acceptable responses. *Public Opinion Quarterly*, 29, 457-458.

Coombs, L., and Freedman, R. (1964). Use of telephone interviews in a longitudinal fertility study. *Public Opinion Quarterly*, 28, 112-117.

Cooper, S.L. (1964). Random sampling by telephone: An improved method. *Journal of Marketing Research*, 1, 45-48.

Couper, M.P., Baker, R.P., Bethlehem, J., Clark, C.Z F., Martin, J., Nicholls, W.L., II and O'Reilly, J.M. - (Eds.) (1998). *Computer Assisted Survey Information Collection*. New York: John Wiley & Sons, Inc.

Couper, M.P., Blair, J. and Triplett, T. (1999). A comparison of mail and e-mail for a survey of employees in US statistical agencies. *Journal of Official Statistics*, 15, 39-56.

Couper, M.P., and Nicholls, W.L., II (1998). The history and development of computer assisted survey information collection methods. *Computer Assisted Survey Information Collection*, (Eds., M.P. Couper, *et al.*). NewYork: John Wiley & Sons, Inc., 1-22.

Cunningham, J.M., Westerman, H.H. and Fischoff, J. (1956). A follow-up study of patients seen in a psychiatric clinic for children. *American Journal of Orthopsychiatry*, 26, 602-610.

Cunningham, P., Berlin, M., Meader, J., Molloy, K., Moore, D. and Pajunen, S. (1997). Using cellular telephones to interview nontelephone households. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 250-254.

Cunningham, P., Brick, J.M. and Meader, J. (2000). 1999 *NSAF In-Person Survey Methods Report No. 5*. Washington, DC: Urban Institute. [http://newfederalism.urban.org/nsaf/ methodology_rpts/ 1999_Methodology_5.pdf].

Cummings, M.K. (1979). Random digit dialing: A sampling technique for telephone surveys. *Public Opinion Quarterly*, 43, 233-244.

Czaja, R., Blair, J. and Sebestik, J.P. (1982). Respondent selection in a telephone survey: A comparison of three techniques, *Journal of Marketing Research*, 19, 381-385.

Decision Analyst (1997). More households using answering machines. *News Release*, *October 15*, *1997*. [http://www. decisionanalyst.com/publ_data/1997/ansmachi.htm].

Dekker, F., and Dorn, P.K. (1984). Computer Assisted Telephone Interviewing: A research project in the Netherlands. Paper presented at: *Conference of the Institute of British Geographers*.

De Leeuw, E.D., and Van Der Zoowen, J. (1988). Data quality in telephone and face to face surveys: A comparative meta-analysis. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 283-299.

Diehr, P., Koepsell, T.D., Cheadle, A. and Psaty, B.M. (1992). Assessing response bias in random-digit dialing surveys: The telephone-prefix method. *Statistics in Medicine*, 11, 1009-1021.

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 Total Design Method* (2^{nd} edition). New York: John Wiley & Sons, Inc.

Drew, J.D., Choudhry, G.H. and Hunter, L.A. (1988). Nonresponse issues in government telephone surveys. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 233-246.

Drew, J.H., and Groves, R.M. (1989). Adjusting for nonresponse in a telephone subscriber survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 452-456.

Dutka, S., and Frankel, L.R. (1980). Sequential survey design through the use of computer assisted telephone interviewing. *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 73-76.

Eastlack, J.O., Jr. (1964). Recall of advertising by two telephone samples. *Journal of Advertising Research*, 4, 25-29.

Eastlack, J.O., Jr., and Assael, H. (1966). Better telephone surveys through centralized interviewing. *Journal of Advertising Research*, 6, 1, 2-7.

Federal Committee on Statistical Methodology (1984). *The Role of Telephone Data Collection in Federal Statistics*. Statistical Policy Working Paper 12, Washington, D.C.

Federal Committee on Statistical Methodology (1990). *Computer Assisted Survey Information Collection*, Statistical Policy Working Paper 19, Washington, D.C.

Federal Republic of Germany (1999). Continuous family budget surveys for January 1999. Statistiches Bundesamt: *Press release*, 20 December, 1999. [http://www.statistik-bund.de/presse/englisch/pm1999/p4350024.htm].

Felson, L. (2001). Netting limitations: Online researchers' new tactics for tough audiences. Marketing News (*American Marketing Association*), 35, 5 [http://www.ama.org/pubs/article.asp?id=4881].

Fink, J.C. (1983). CATI's first decade: The Chilton experience. *Sociological Methods and Research*, 12, 153-168.

Fischbacher, C., Chappel, D., Edwards, R. and Summerton, N. (1999). The use and abuse of the Internet for survey research. *Proceedings of the Association for Survey Computing 3rd International Conference*, Edinburgh, 501-507.

Fitti, J.E. (1979). Some results from the telephone health interview system. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 244-249.

Foley, D.J., and Brock, D.B. (1990). Comparison of in-person and telephone responses in a survey of the last days of life. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 382-386.

Ford, E.S. (1998). Characteristics of survey participants with and without a telephone; findings from the third National Health and Nutrition Examination Survey. *Journal of Clinical Epidemiology*, 51, 55-60.

Forsman, G. (1993). Sampling individuals within households in telephone surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1113-1118.

Forsman, G., and Danielsson, S. (1997). Can plus digit sampling generate a probability sample? *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 958-963.

Fox, A., and Riley, J.P. (1996). Telephone coverage, housing quality and rents: RDD survey biases. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 515-519.

Frankel, M.R., Srinath, K.P., Battaglia, M.P., Hoaglin, D.C., Wright, R.A. and Smith, P.J. (1999). Reducing nontelephone bias in RDD surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 934-937.

Freeman, H.E., and Shanks, J.M. - Eds. (1983). The emergence of computer-assisted survey research. *Sociological Methods and Research*, 12 (special issue), 115-230.

Fréjean, M., Panzani, J.-P. and Tassi, P. (1990). Les ménages inscrits en liste rouge et les enquêtes par téléphone. *Journal de la Société de Statistique de Paris*, 131, Nos. 3-4, 86-102.

Frey, J.H. (1989). *Survey Research by Telephone* (2nd edition). Beverly Hills, CA: Sage Publications.

Fry, H.G., and McNaire, S. (1958). Data gathering by long distance telephone. *Public Health Records*, 73, 831-835.

Gabler, S., and Haeder, S. (2000). Telephone sampling in Germany. *Paper presented at Fifth International Conference of Social Science Methodology*, Koln.

Genesys (1996). Unlisted numbers: What's really important. *Genesys News* (Genesys Sampling Systems, Fort Washington, PA), 1-2.

Ghosh, D. (1984). Improving the plus 1 method of random digit dialing. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 285-288.

Giesbrecht, L.H., Kulp, D.W. and Starer, A.W. (1996). Estimating coverage bias in RDD samples with current population survey data. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 503-508.

Glasser, G.J., and Metzger, G.D. (1972). Random digit dialing as a method of telephone sampling. *Journal of Marketing Research*, 9, 59-64.

Glasser, G.J., and Metzger, G.D. (1975). National estimates of nonlisted telephone households. *Journal of Marketing Research*, 12, 359-361.

Goksel, H., Judkins, D.R. and Mosher, W.D. (1991). Nonresponse adjustments for a telephone follow-up to a national in-person survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 581-586.

Groves, R.M. (1977). *An Empirical Comparison of Two Telephone Sample Designs*. Unpublished report of the Survey Research Center, the University of Michigan, Ann Arbor, MI.

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., II and Waksberg, J. - Eds. (1988). *Telephone Survey Methodology*. New York: John Wiley & Sons, Inc.

Groves, R.M., and Kahn, R.L. (1979). *Surveys by Telephone*: *A National Comparison With Personal Interview*. New York: Academic Press.

Groves, R.M., and Lepkowski, J.M. (1985). Dual frame, mixed mode survey designs. *Journal of Official Statistics*, 1, 264-286.

Groves, R.M., and Lepkowski, J.M. (1986). An experimental implementation of a dual frame telephone sample design. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 340-345.

Groves, R.M., and Lyberg, L.E. (1988a). An overview of nonresponse issues in telephone surveys. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 191-211.

Groves, R.M., and Lyberg, L.E. - Eds. (1988b). Telephone survey methodology. *Journal of Official Statistics* (special issue), 4, 283-416.

Gunn, W.J., and Rhodes, I.N. (1981). Physician response rates to a telephone survey: Effects of monetary incentive level. *Public Opinion Quarterly*, 45, 109-115.

Hagan, D.E., and Meier, C.C. (1983). Must respondent selection procedures for telephone surveys be invasive? *Public Opinion Quarterly*, 47, 547-556.

Harlow, B.L., Crea, E.C., East, M.A., Oleson, B., Fraer, C.J. and Cramer, D.W. (1993). Telephone answering machines: The influence of leaving messages on telephone interviewing response rates. *Journal of Epidemiology*, 4, 380-383.

Hartge, P., Brinton, L.A., Rosenthal, J.F., Cahill, J.I., Hoover, R.N. and Waksberg, J. (1984). Random digit dialing in selecting a population-based control group. *American Journal of Epidemiology*, 120, 825-833.

Hauck, M., and Cox, M. (1974). Locating a sample by random digit dialing. *Public Opinion Quarterly*, 38, 253-260.

Herzog, A.R., and Rodgers, W.L. (1988). Interviewing older adults: Mode comparison using data from a face-to-face survey and a telephone resurvey. *Public Opinion Quarterly*, 52, 84-99.

Hoaglin, D.C., and Battaglia, M.P. (1996). A comparison of two methods of adjusting for noncoverage of nontelephone households in a telephone survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 497-502.

Hochstim, J.R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62, 976-989.

Hogue, C.R., and Chapman, D.W. (1984). An investigation of PSU cutoff points for a random digit dialing survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 286-291.

Hox, J.J., De Leeuw, E.D. and Kreft, I.G.G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: A multilevel model. *Measurement errors in surveys* (Eds., P.P. Biemer, L.E. Lyberg, N.A. Mathiowetz and S. Sudman). New York: John Wiley & Sons, Inc., 439-461.

Inglis, K.M., Groves, R.M. and Heeringa, S.G. (1987). Telephone sample designs for the U.S. Black household population. *Survey Methodology*, 13, 1-14.

Janofsky, A.I. (1971). Affective self-disclosure in telephone versus face-toface interviews. *Journal of Humanistic Psychology*, 11, 93-103.

Johnson, T., Fendrich, M., Shaligram, C. and Garey, A. (1997). A comparison of interviewer effects models in an RDD telephone survey of drug use. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 887-892.

Kalsbeek, W.D., and Durham, T.A. (1994). Nonresponse and its effects in a followup telephone survey of low-income women. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 943-948.

Kalton, G. (2000). Developments in survey research in the past 25 years. *Survey Methodology*, 26, 3-10.

Katz, D., and Cantril, H. (1937). Public opinion polls. *Sociometry*, 1, 155-179.

Katz, J.E., and Aspden, P. (1998). Internet dropouts in the USA. *Telecommunications Policy*, 22, 4/5, 327-339.

Keeter, S. (1995). Estimating telephone noncoverage bias with a telephone survey. *Public Opinion Quarterly*, 59, 196-217.

Kehoe, C., Pitkow, J., Sutton, K., Aggarwal, G. and Rogers, J.D. (1999). *Results of GVU's Tenth World Wide Web User Survey*. Atlanta, GA: Graphics Visualization and Usability Center, College of Computing, Georgia Institute of Technology. [http://www.gvu.gatech.edu/user_surveys].

Khurshid, A., and Sahai, H. (1995). A bibliography on telephone survey methodology. *Journal of Official Statistics*, 11, 325-367.

Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American Statistical Association*, 44, 380-387.

Koepsell, T.D., McGuire, V., Longstreth, Jr., W.T., Nelson, L.M. and Van Belle, G. (1996). Randomized trial of leaving messages on telephone answering machines for control recruitment in an epidemiological study. *American Journal of Epidemiology*, 144, 704-706.

Körmendi, E. (1988). The quality of income information in telephone and face to face surveys. *Telephone Survey Methodology* (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 341-356.

Kuusela, V., and Notkola, V. (1999). Survey quality and mobile phones. Paper presented at *International Conference on Survey Nonresponse*, Portland OR.

Kuusela, V., and Vikki, K. (1999). Change of telephone coverage due to mobile phones. Paper presented at *International Conference on Survey Nonresponse*, Portland OR.

Landon, E.L., Jr., and Banks, S.K. (1977). Relative efficiency and bias of plus-one telephone sampling. *Journal of Marketing Research*, 14, 294-299.

Larson, O.N. (1952). The comparative validity of telephone and face-to-face interviews in the measurement of message diffusion from leaflets. *American Sociological Review*, 17, 471-476.

Lavrakas, P.J. (1993). *Telephone Survey Methods*: *Sampling, Selection and Supervision* (2^{nd} edition). Newbury Park, CA: Sage Publications.

Lavrakas, P.J., Bauman, S.L. and Merkle, D.M. (1993). The last-birthday method and within-unit coverage problems. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1107-1112.

Lepkowski, J.M. (1988). Telephone sampling methods in the United States. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 73-98.

Lepkowski, J.M., and Groves, R.M. (1984). The impact of bias on dual frame survey design. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 265-270.

Lepkowski, J.M., and Groves, R.M. (1986a). A two phase probability proportional to size design for telephone sampling. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 357-362.

Lepkowski, J.M., and Groves, R.M. (1986b). A mean square error model for dual frame, mixed mode survey design. *Journal of the American Statistical Association*, 81, 930-937.

Leuthold, D.A., and Scheele, R. (1971). Patterns of bias in samples based on telephone directories. *Public Opinion Quarterly*, 35, 249-257.

Locander, W., Sudman, S. and Bradburn, N. (1976). An investigation of interview method, threat and response distortion. *Journal of the American Statistical Association*, 71, 269-275.

Lyberg, L., and Kasprzk, D. (1991). Data collection methods and measurement error: An overview. *Measurement Errors in Surveys* (Eds., P.P. Biemer, L.E. Lyberg, N.A. Mathiowetz and S. Sudman). New York: John Wiley & Sons, Inc., 237-258.

McCarthy, W.F., and Bateman, D.V. (1988). The use of mathematical programming for designing dual frame surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 652-653.

McKay, R.B., Robison, E.L. and Malik, A.B. (1994). Touch-tone data entry for household surveys: Research findings and possible applications. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 509-511.

Maffeo, C., Frey, W. and Kalton, G. (2000). Survey design and data collection issues in the Disability Evaluation Study. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, forthcoming.

Maklan, D., and Waksberg, J. (1988). Within household coverage in RDD surveys. *Telephone Survey Methodology* (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 51-69.

Malakhoff, L.A., and Appel, M.V. (1997). The development of a voice recognition prototype for field listing. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 234-238.

Mason, R.E., and Immerman, F.W. (1988). Minimum cost sample allocation for Mitofsky-Waksberg random digit dialing. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 127-141.

Massey, J.T. (1995). Estimating the response rate in a telephone survey with screening. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 673-677.

Massey, J T., and Botman, S.L. (1988). Weighting adjustments for random digit dialed surveys. *Telephone Survey Methodology*, (Eds., R.M. Groves *et al.*). New York: John Wiley & Sons, Inc., 143-160.

Massey, J.T., O'Connor, D. and Krotki, K. (1997). Response rates in random digit dialing (RDD) telephone surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 707-712.

Meeks, R.L., Lanier, A.T., Fecso, R.S. and Collins, M.A. (1998). Web-based data collection in national science foundation surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 349-353.

Merkle, D.M., Bauman, S.L. and Lavrakas, P.J. (1993). The impact of callbacks on survey estimates in an annual RDD survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1070-1075.

Mitchell, G.H., and Rogers, E.M. (1958). Telephone interviewing. *Journal of Farm Economics*, 40, 743-747.

Mitofsky, W. (1970). Sampling of Telephone Households. Unpublished CBS memorandum.

Mohadjer, L. (1988). Stratification of prefix areas for sampling rare populations. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 161-173.

Mullet, G.M. (1982). The efficacy of plus-one dialing: Self- reported status. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 575-576.

Nathan, G., and Aframian, N. (1996). An experiment with CATI in Israel. Paper presented at *InterCasic '96 Conference*, San Antonio, TX.

Nathan, G., and Eliav, T. (1988). Comparison of measurement errors for telephone interviewing and home visits by misclassification models. *Journal of Official Statistics*, 4, 363-374.

Nicholls, W.L., II (1983). CATI research and development at the Census Bureau. *Sociological Methods and Research*, 12, 191-198.

Nicholls, W.L., II (1988). Computer-assisted telephone interviewing: A general introduction. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 377-386.

Nicolaas, G., Lynn, P. and Lound, C. (2000). Random digit dialling in the UK: Viability of the sampling method revisited. Paper presented at Fifth International Conference of Social Science Methodology, Koln.

Norris, D.A., and Paton, D.G. (1991). Canada's General Social Survey: Five years of experience. *Survey Methodology*, 17, 227-240.

NTIA (2000). *Falling Through the Net*, *Toward Digital Inclusion*. Washington DC: National Telecommunications and Information Administration.

Nusser, S., and Thompson, D. (1998). Web-based survey tools. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 951-956.

Oakes, R.H. (1954). Differences in responsiveness in telephone versus personal interviews. *Journal of Marketing*, 19, 169.

Oksenberg, L., and Cannel, C. (1988). Effects of interviewer vocal characteristics on nonresponse. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 257-269.

Oldendick, R.W., Bishop, G.F., Sorenson, S.B. and Tuchfarber, A.J. (1988). A comparison of the Kish and last birthday methods of respondent selection in telephone surveys. *Journal of Official Statistics*, 4, 307-318.

Oldendick R.W., and Link, M.W. (1994). The answering machine generation: Who are they and what problem do they pose for survey research? *Public Opinion Quarterly*, 58, 264-273.

Oftel (1999). *Homes Without a Fixed Line Phone - Who Are They*? [http://www.oftel.gov.uk/publications/research/unph0400.htm].

Oftel (2000). *Consumers' use of Internet*. [http://www.oftel.gov.uk/publications/research/int1000.htm]

O'Reilly, J.M., Hubbard, M.L., Lessler, J.T., 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, 197-214.

O'Rourke, D., and Blair, J. (1983). Improving random respondent selection in telephone surveys. *Journal of Marketing Research*, 20, 428-432.

Palit, C.D. (1980). A microcomputer based computer assisted interviewing system. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 243-244.

Palit, C.D. (1983). Design strategies in RDD sampling. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 627-629.

Palit, C.D., and Blair, J. (1986). Some alternatives for the treatment of first phase telephone numbers in a Waksberg-Mitofsky RDD sample. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 363-369.

Palit, C.D., and Sharp, H. (1983). Microcomputer-assisted telephone interviewing. *Sociological Methods and Research*, 12, 169-189.

Pannekoek, J. (1988), Interviewer variance in a telephone survey. *Journal of Official Statistics*, 4, 375-384.

Payne, S.L. (1956). Some advantages of telephone surveys. *Journal of Marketing*, 20, 278-281.

Perneger, T.V., Myers, T.L., Klag, M.J. and Whelton, P.K. (1993). Effectiveness of the Waksberg telephone sampling method for the selection of population controls. *American Journal of Epidemiology*, 138, 574-584.

Perone, C., Matrundola, G. and Soverini, M. (1999). A quality control approach to mobile phone surveys; the experience of Telecom Italia Mobile. *Proceedings of the Association for Survey Computing 3rd International Conference*, Edinburgh, 180-187.

Perry, J.B. (1968). A note on the use of telephone directories as a sample source. *Public Opinion Quarterly*, 32, 691-695.

Phipps, P.A., and Tupek, A.R. (1991). Assessing measurement errors in a touchtone recognition survey. *Survey Methodology*, 17, 15-26.

Piazza, T. (1993). Meeting the challenge of answering machines. *Public Opinion Quarterly*, 57, 219-231.

Potter, F.J., McNeill, J.J., Williams, S.R. and Waitman, M.A. (1991). List-assisted RDD telephone surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 117-122.

Potthoff, R.F. (1987a). Some generalizations of the Mitofsky-Waksberg technique of random digit dialing. *Journal of the American Statistical Association*, 82, 409-418.

Potthoff, R.F. (1987b). Generalizations of the Mitofsky-Waksberg technique for random digit dialing: Some added topics. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 615-620.

Poynter, R. (2000). *We've Got Five Years*. London: Association for Survey Computing meeting on Survey Research on the Internet.

Ramos, M., Sedivi, B.M. and Sweet, E.M. (1998). Computerized self-administered questionnaires. *Computer Assisted Survey Information Collection* (Eds., M.P. Couper, *et al.*). NewYork: John Wiley & Sons, Inc., 389-408.

Ranta-Aho, M., and Leppinen, A. (1997). Matching telecommunication services with user communication needs. *Proceedings of the International Symposium on Human Factors in Telecommunications*, (Eds., K. Nordby and L. Grafisk). Oslo, Norway, 401-408. [http://www.comlab.hut.fi/hft/publications/matcharticle.pdf].

Rich, C.L. (1977). Is random digit dialing really necessary? *Journal of Marketing Research*, 14, 300-305.

Rogers, T.F. (1976). Interviews by telephone and in person: Quality of responses and field performance. *Public Opinion Quarterly*, 40, 51-65.

Rogers, S.M., Miller, H.G., Forsyth, B.H., Smith, T.K. and Turner, C.F. (1996). Audio-CASI: The impact of operational characteristics on data quality. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1042-1047.

Romuald, K.S., and Haggard, L.M. (1994). The effect of varying the respondent selection script on respondent self-selection in RDD telephone surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1299-1304.

Rosen, R.J., Manning, C.D. and Harrell, L.J., Jr. (1998). Web-based data collection in the current employment statistics survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 354-359.

Roslow, S., and Roslow, L. (1972). Unlisted phone subscribers are different. *Journal of Advertising Research*, 7, 4, 35-38.

Rouquette, C. (2000). La percée du telephone portable et d'Internet. *INSEE Première No. 200. INSEE*, Paris. [http://www.insee.fr/fr/ffc/docs_ffc/ip700.pdf].

St. Clair, J., and Muir, J. (1997). Household adoption of digital technologies. *Year Book Australia 1997.* Canberra: Australian Bureau of Statistics.

Salmon, C.T., and Nichols, J.S. (1983). The next birthday method of respondent selection. *Public Opinion Quarterly*, 47, 270- 276.

Scheuren, F., and Petska, T. (1993). Turning administrative systems into information systems. *Journal of Official Statistics*, 9, 109-119.

Schmiedeskamp, J.W. (1962). Reinterviews by telephone. *Journal of Marketing*, 26, 28-34.

Sebold, J. (1988). Survey period length, unanswered numbers, and nonresponse in telephone surveys. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 247-256.

Shanks, J.M. (1983). The current status of computer assisted telephone interviewing: Recent progress and future prospects. *Sociological Methods and Research*, 12, 119-142.

Shanks, J.M., Nicholls, W.L., II and Freeman, H.E. (1981). The California Disability Survey: Design and execution of a computer-assisted telephone study. *Sociological Methods and Research*, 10, 123-140.

Shapiro, G.M., Battaglia, M.P., Hoaglin, D.C., Buckley, P. and Massey, J.T. (1996). Geographical variation in within-household coverage of households with telephones in an RDD survey. *Proceedings of the Section on Survey Research Methods*, 491-496.

Shure, G.E., and Meeker, R.J. (1978). A mini-computer system for multi-person computer-assisted telephone interviewing. *Behavior Methods and Instrumentation*, (April) 196-202.

Smith, C., and Frazier, E.L. (1993). Comparison of traditional and modified Waksberg. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 926-931.

Squire, P. (1988). Why the 1936 Literary digest poll failed. *Public Opinion Quarterly*, 52, 125-133.

Statistics Canada (2000). *Selected Dwelling Characteristics and Household Equipment*. Income Statistics Division. [http://www.statcan.ca/english/Pgdb/People/Families/famil09b.htm].

Stock, J.S. (1962). How to improve samples based on telephone listings. *Journal of Advertising Research*, 2, 3, 50-51.

Stokes, L., and Yeh, M.-Y. (1988). Searching for causes of interviewer effects in telephone surveys. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 357-376.

Sudman, S. (1966). New uses of telephone methods in survey research. *Journal of Marketing Research*, 3, 107-120.

Sudman, S. (1973). The uses of telephone directories for survey sampling. *Journal of Marketing Research*, 10, 204-207.

Sudman, S. (1978). Optimum cluster designs within a primary unit using combined telephone screening and face-to-face interviewing. *Journal of the American Statistical Association*, 73, 300-304.

Survey Research Center (2000). *Sample Design for Household Telephone Surveys*: *A Bibliography* 1949-1996. College Park, MD: University of Maryland. [http://www.bsos.umd.edu/src/sampbib.html].

Survey Sampling Inc. (1998). Random digit samples – part 1. [http://www.ssisamples.com/ssi.x2o$ssi_gen.search_item?id=119].

Statistics Netherlands (1987). *Automation in Survey Processing*. Voorburg/Heerlen: Netherlands Central Bureau of Statistics (CBS Select 4).

Sykes, W.M., and Collins, M. (1987). Comparing telephone and face-to-face interviewing in the United Kingdom. *Survey Methodology*, 13,15-28.

Sykes, W.M., and Collins, M. (1988). Effects of mode of interview: Experiments in the UK. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 301-320.

Thornberry, O.T. Jr., and Massey, J.T. (1978). Correcting for undercoverage bias in random digit dialed National Health Surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 224-229.

Thornberry, O.T. Jr., and Massey, J.T. (1983). Coverage and response in random digit dialed national surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 654-659.

Thornberry, O.T. Jr., and Massey, J.T. (1988). Trends in United States telephone coverage across time and subgroups. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 25-49.

Tortora, R.D. (1985). CATI in an agricultural statistics agency. *Journal of Official Statistics*, 1, 301-314.

Tourangeau, R., and Smith, T.W. (1998). Collecting sensitive information with different modes of data collection. *Computer Assisted Survey Information Collection*, (Eds., M.P. Cooper, *et al.*). NewYork: John Wiley & Sons, Inc., 431-453.

Traugott, M.W., Groves, R.M. and Lepkowski, J.M. (1987). Using dual frame designs to reduce nonresponse in telephone surveys. *Public Opinion Quarterly*, 51, 522-539.

Trewin, D., and Lee, G. (1988). International comparisons of telephone coverage. *Telephone Survey Methodology*, (Eds., R.M. Groves, *et al.*). New York: John Wiley & Sons, Inc., 3-24.

Troldahl, V.C., and Carter, R.E. (1964). Random selection of respondents within households in phone surveys. *Journal of Marketing Research*, 1, 71-76.

Tuckel, P.S., and Feinberg, B.M. (1991) The answering machine poses many questions for telephone survey researchers. *Public Opinion Quarterly*, 55, 200-217.

Tuckel, P.S., and O'Neill, H. (1996), New technology and nonresponse bias in RDD surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 889-894.

Tuckel, P., and Shukers, T. (1997). The effect of different introductions and answering machine messages on response rates. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1047-1051.

Tucker, C., Casady, R. and Lepkowski, J. (1992). Sample allocation for stratified telephone sample designs. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 291-296.

Turner, C.F., Forsyth, B.H., O'Reilly, J.M., Cooley, P.C., Smith, T.K., Rogers, S.M. and Miller, H.G. (1998). Automated self-interviewing and the survey measurement of sensitive behaviors. *Computer Assisted Survey Information Collection*, (Eds., M.P. Cooper, *et al.*). NewYork: John Wiley & Sons, Inc., 455-473.

Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, 73, 40-46.

Waksberg, J. (1983). A note on 'Locating a special population using random digit dialing'. *Public Opinion Quarterly*, 47, 576-578.

Waksberg, J. (1984). *Efficiency of Alternative Methods of Establishing Cluster Sizes in RDD Sampling*. Unpublished Westat memorandum.

Waksberg, J., Brick, J.M., Shapiro, G., Flores-Cervantes, I. and Bell, B. (1997). Dual-frame RDD and area sample for household survey with particular focus on low-income population. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 713-718.

Werking, G., Tupek, A.R. and Clayton, R.L. (1988). CATI and touchtone self-response applications for establishment surveys. *Journal of Official Statistics*, 4, 349-362.

White, A.A. (1983). Response rate calculation in RDD telephone health surveys: Current practices. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 277-282.

Whitmore, R.W., Mason, R.E. and Hartwell, T.D. (1985). Use of geographically classified telephone directory lists in multi-mode surveys. *Journal of the American Statistical Association*, 80, 842-844.

Wilson, P., Blackshaw, N. and Norris, P. (1988). An evaluation of telephone interviewing on the British Labour Force Survey. *Journal of Official Statistics*, 4, 385-400.

Winter, D.L.S., and Clayton, R.L. (1990). Speech data entry: Results of the first test of voice recognition for data collection. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 387-392.

Wiseman, F. (1972). Methodological bias in public opinion surveys. *Public Opinion Quarterly*, 36, 105-108.

Xu, M., Bates, B.J. and Schweitzer, J.C. (1993). The impact of messages on survey participation in answering machine households. *Public Opinion Quarterly*, 57, 232-237.