Media Rating Council, Inc.

Within-Household Respondent Selection

An MRC Staff Point of View

In 2008, the MRC retained noted survey researcher Dr. Paul J. Lavrakas to review the state of knowledge concerning within-household respondent selection methods, and to make recommendations concerning current and possible enhanced practices related to this subject. We did so because we had a growing concern that some widely used methods (e.g., "Last Birthday") might not be working as intended nor yielding samples with satisfactory demographic compositions, at least in some cases.

Dr. Lavrakas' full report for the MRC follows this cover note. Because of the depth and quality of his work, we thought it worth sharing in its entirety. Here we summarize the conclusions the MRC Staff has drawn from this work.

First, we think the following findings are evident:

- For both telephone and mail survey methods, the research is clear that pure birthday-based methods for selecting random persons within households "have been found to lead to a non-ignorable amount of selection errors in which the 'wrong' person becomes the designated respondent for her/his household. These selection errors are not all random and lead to an overrepresentation of females and those most willing to participate in surveys."

- Furthermore, it also seems clear that creative alternative approaches and enhancements are available and worthy of consideration and testing by ratings services that interact with the MRC. These alternatives and enhancements may reduce the sampling biases which have been observed with the traditional birthday-only selection methods.

Therefore, the MRC Staff sees the following implications for ratings services which utilize within-household respondent selection methods as a part of Accredited services, or services seeking Accreditation:

1. **A need for short-term reviews of such procedures for potential enhancements.** As outlined in the attached report, some aspects of within-household respondent selection procedures can be enhanced with relatively easy modifications. The MRC will be expecting all relevant Services to conduct a timely review of their within-household selection procedures for near-term enhancement opportunities.

2. **A need for contemporary, Service-appropriate methodological research on within-household respondent selection methods.** As the Lavrakas report makes clear, changes in selection methods involve trade-offs among randomness purity, sample composition, and cooperation. However, we see real potential for improvement in at least one new telephone-based technique (e.g., the "Westat Method"), and in certain mail-based en-

hancements, and the MRC will be expecting meaningful and timely commitments from relevant Services to the testing of new ideas in this arena.

It is important to note that the MRC itself is not presently advocating a *particular* question wording or test design for *particular* ratings services. Dr. Lavrakas' proposals should be considered as suggestive within the larger context of each Service's methods, samples, and overall bias-reduction priorities.

For example, the MRC has not yet seen the "Pew method" described in the attachment used in any Accredited rating service, even in the limited way mentioned by Dr. Lavrakas. If a rating service chose to investigate the technique, that would need careful offline research prior to implementation, we believe.

However, this Point of View document is intended to make clear the importance we place on each relevant Service *taking action* on reducing this potential source of bias. The research literature is now clear that material respondent selection errors *can and often do* occur, and the MRC has an obligation to assure that ratings services are responding accordingly.

About Dr. Lavrakas

Paul J. Lavrakas, Ph.D. is a research psychologist and currently is serving as a methodological research consultant for several organizations. He served as Vice President and chief methodologist for Nielsen Media Research from 2000-2007. Previously, Dr. Lavrakas was a Professor of Journalism and Communication Studies at Northwestern University (1978-1996) and at Ohio State University (1996-2000). During his academic career he was the founding faculty director of the Northwestern University Survey Lab (1982-1996) and the OSU Center for Survey Research (1996-2000).

Among his publications, he has written a widely read book on telephone survey methodology and served as the lead editor for three books on election polling, the news media, and democracy, as well as co-authoring four editions of *The Voter's Guide to Election Polls*. He served as guest editor for a special issue of *Public Opinion Quarterly* on "Cell Phone Numbers and Telephone Surveys" published in December 2007, and also is the editor of the *Encyclopedia of Survey Research Methods* which Sage published in September, 2008. Dr. Lavrakas is a current elected member of the AAPOR Executive Council and was a co-recipient of the 2003 AAPOR Innovators Award for his work on the standardization of survey response rate calculations.

About the Media Rating Council (MRC)

The MRC is a non-profit industry association established in 1964 composed of leading television, radio, print and Internet companies, as well as advertisers, advertising agencies and trade associations whose goal is to ensure measurement services that are valid, reliable and effective. Measurement services desiring MRC Accreditation are required to disclose to their customers all methodological aspects of their service; comply with the MRC *Minimum Standards for Media Rating Research*; and submit to MRC-designed audits to authenticate and illuminate their procedures. In addition, the MRC membership actively pursues research issues they consider priorities in an effort to improve the quality of research in the marketplace. Currently approximately 50 research products are audited by the MRC.

Media Rating
Council, Inc.

Within-Household Respondent Selection:

How Best to Reduce Total Survey Error?

Prepared for:

Media Rating Council, Inc.

Prepared by:

Paul J. Lavrakas, Ph.D.

November 2008

Media Rating Council, Inc.

Within-Household Respondent Selection: How Best to Reduce Total Survey Error?

This report has been prepared for the Media Rating Council (MRC) to help them, and the audience measurement services they audit, make more informed decisions about how to best handle the thorny issue of within-household respondent selection in landline telephone and mail surveys. (Cell phone issues are also discussed briefly.) The report addresses the following major questions:

1. Why should survey organizations be concerned about how best to select a survey respondent within a household?

2. What is the nature of the errors that are associated with the well-known and often used within-household respondent selection methods?

3. Which within-household selection method is the best for a landline telephone survey in which telephone numbers are selected in the first stage of sampling and which method is best for a mail survey in which addresses are selected in the first stage of sampling?

4. What new research studies can be conducted by MRC audited research companies in order to further improve the within-household respondent selection procedures they use?

This report has been prepared by a consultant to the MRC, Paul J. Lavrakas, who has contributed to the survey industry's state of knowledge on within-household respondent selection through his publications and conference papers during the past two decades.

Why Representative Within-Household Respondent Selection is Important

Many surveys require the selection of one, and only one, "designated" respondent within a sampled household. This person is the one respondent in that household from whom data are to be gathered. Ideally, this person is selected randomly from among all eligible persons (as defined in the survey design) in the household (cf. Gaziano, 2005).

The proper selection of a designated respondent is important if survey estimates are meant to represent parameters at the person-level in the target population. If a systematic method is not used for selecting one respondent per household in such surveys, then the resulting sample will be comprised of the "most willing" and "most readily available" persons at the sampled households. This in turn will yield an unrepresentative final sample of persons that contains proportionally too many (a) females, (b) older adults, (c) those without fulltime employment, and (d) those who, in general, are more disposed towards cooperating with a survey request. It also may be unrepresentative in other ways that bias the final survey estimates because of the use of an unsystematic and/or nonrandom respondent selection method.

Furthermore, although weighting adjustments can be used to balance a sample along the lines of key demographic characteristics, there is no guarantee that these adjustments will correct for any existing within-unit coverage bias nor is there any certainty that population parameters (i.e., universe estimates [UEs]) will exist to properly weight against – e.g., many surveys will not be able to weight for employment status as no reliable UEs will be available for the geopolitical area being sampled, nor are there population parameters available for the host of psychographic variables that are related to survey cooperation.

Thus it is important that surveys select one respondent per sampling unit by deploying a systematic approach for this selection – one that does not allow the "most willing" or "most readily available" person to select themselves. It also is important that surveys with interviewer-administered data collection utilize a systematic approach that avoids interviewers merely choosing whomever within the household they prefer to interview.

However, just because a systematic within-household respondent selection technique is deployed does not mean it will work as it is intended. And, even if it does work as intended, it does not mean it will work well enough for the needs of the survey. If, for example, the respondent eligibility criterion for a survey is simply "18 years of age and older," then all the adult residents are eligible for selection as the designated respondent in their respective household. If there is only one eligible adult resident, then by definition s/he becomes the designated respondent. If there is more than one eligible adult who resides in the household, then a respondent selection method must be used to choose the one who will be treated as the designated respondent.

Ideally, this respondent selection method will yield a representative and unbiased within-household sample across all the households that are selected for surveying. But for this to happen, the method itself must

- *in theory* be capable of achieving a representative within-household sample, and
- *in practice* not lead to non-ignorable (biasing) *misselection* errors.

Yet, as discussed below, many of the within-household selection techniques as they are used by survey organizations do not achieve both of these goals. That is, they either are not devised to yield representative samples and/or they are not (or cannot be feasibly) implemented so as to avoid non-ignorable misselection errors.

Noncoverage versus Nonresponse vis-a-vis Within-Household Selection

The issue of within-household respondent selection is further complicated by the effect the respondent selection technique may have on survey cooperation. Unfortunately, the choice of a within-household respondent selection procedure often pits concerns about coverage and *possible coverage bias* against concerns about nonresponse and *possible nonresponse bias*.

The research literature is generally in agreement that as the rigor of a within-household respondent selection method increases, the chances of gaining a representative within-unit sample also increases, thereby improving coverage and possibly lowering the chance of noncoverage bias. But too often the trade-off is an increase in nonresponse and the possibility of increasing the chance of nonresponse bias. Furthermore, as the rigor of the respondent selection methods increases, too often there is an associated increase in the percentage of households in which the respondent is misselected and the "wrong" person ends up serving as the respondent who provides data.

Research on the Challenges of Representative Within-Household Respondent Selection

The most rigorous respondent selection approach that has been implemented by survey researchers was described by Kish (1949) more than half a century ago. The Kish method is thought to be the "gold standard" for within-unit respondent selection in interviewer-administered surveys. However, it is regarded as too complicated to deploy in self-administered surveys, such as mail surveys.[1]

The Kish method began being used in face-to-face surveys conducted in people's homes. The method requires an enumeration of all persons residing in a household who meet the survey's eligibility criteria (e.g., being 18 years of age or older). For households with only one eligible, that person becomes the designated respondent. For households with more than one eligible, the interviewer asks for the gender and age of each of them, one at a time. The interviewer then uses one of a series of selection grids pegged to the number of eligible residents in the household to choose one of the eligibles as the designated respondent. Kish acknowledged that his method was not perfect in achieving the selection of a random respondent in households of all sizes. But the Kish method breaks down slightly only among very large households (those with seven or more eligible residents), which constitute an extremely small proportion of all survey households. This potential for nonrepresentation is believed to be wholly ignorable in almost all surveys.[2]

As the public's willingness to participate in surveys began to decline and as telephone surveys became the most frequently used mode to gather survey data from the general public – with both trends occurring by the mid-1980s – researchers became concerned that the Kish method was contributing to survey nonresponse due to the invasive nature of the information that interviewers needed to gather in order to deploy the method properly, especially given that respondent selection needs to be completed as part of the initial contact an interviewer makes with a household. As such, less invasive within-household selection methods were devised and began to be

[1] The Kish selection technique requires the use of a number of grids to make the final respondent selection in a given household and it is impractical to send all these grids to a respondent or to try to explain to the respondent why such a complex approach is being used. In contrast, the complexity of the Kish method remains essentially invisible to respondents when the method is deployed by an interviewer.

[2] Recent research has suggested that the set of selection grids that Kish created nearly 60 years ago may no longer work as well as when they first were devised due to changing demographic composition of households in our modern era (cf. Nemeth, 2001). Thus, there may be need to modify these grids.

reported in the research literature (Bryant, 1975; Hagan and Collier, 1983; Salmon and Nichols, 1983).

Many researchers liked these less invasive methods because interviewers found it easier to gain cooperation from households with them than when deploying the more rigorous Kish method. These simpler methods were especially appealing to commercial survey organizations because they saved time and thus survey costs. However, none of these methods were as scientifically rigorous or as successful as the Kish method in selecting a randomly chosen, representative respondent from within the household, thereby adequately covering the within-household population of interest.

Among all of these alternative methods, the so-called "birthday" methods, in theory, are at least quasi-random in their selections. One version of the birthday method selects the person in the household (among all eligibles) who had the *last* (most recent) birthday, whereas the other version selects the eligible person with the *next* (forthcoming) birthday. Yet the birthday approach is not truly random as it skews respondent selection to those eligibles born in months closest to the day that respondent selection is made for a given household. Whether this nonrandom aspect of the birthday methods is linked to any bias has not been reported in the research literature, but it seems improbable that it would be.[3] Furthermore, as noted by Battaglia, Link, Frankel and Mokdad (2007), by randomly applying *both* the next and last birthday selection methods in the same survey, selection within two person households becomes truly random (in theory, if not in practice), and selection within larger-sized households moves closer to true randomness (again, in theory, if not in practice).

None of the other alternative selection methods are random, although they all utilize a systematic method to select one respondent so as to generate better within household representation along the lines of gender and age than occurs when no selection method is applied.

Concerns about the nonrandomness of alternative selection methods notwithstanding, as methodological studies on various respondent selection methods were conducted in the 1980s and 1990s, another and greater concern mounted. This was the finding that *misselection errors* in the application of these less invasive procedures may be leading to other within-unit coverage errors that may be offsetting any reduction in Total Survey Error that was being gained from the assumed reduction in nonresponse associated with use of the less invasive selection methods.

However, until the mid-1990s, no one had reported a rigorous study that was able to dimension the nature of the survey errors associated with within-unit selection, including the errors that might be associated with trade-offs between nonresponse vs. noncoverage. Two studies reported that with the last birthday method more than 20 percent of the respondents whose households were sampled in RDD landline surveys were misselected and that these misselections were not all random error (cf. Lavrakas, Bauman, and Merkle, 1994; Lavrakas, Harpuder, and Stasny, 2000). Lind, Link and Oldendick (2001) also found that the last birthday method misselected re-

[3] New analyses conducted in early 2008 specifically for this MRC report using the three RDD landline studies from Lavrakas et al. (2000) found no meaningful correlations between substantive data provided by respondents and the number of months their birthday was from the month the interview was conducted. These findings held across a diverse range of demographic and media-related variables.

spondents in 20 percent of households, whereas the next birthday method misselected in 27 percent of households. Grandjean, Leighty, Taylor and Xu (2004) reported results from a survey experiment that selected from adults and nonadults in households, and found that the birthday approach (compared to a true random selection ala Kish) significantly over-selected children, in particular female children. They speculated that these selection errors occurred because children's birthdays are "more memorable." Battaglia et al. (2007) reported that nearly 40 percent of the respondents in their mail surveys were misselected by the birthday method.

No other studies have been reported on the misselection rate or the nature of the misselection errors that occur with the other less invasive methods, but there is no reason to assume that these other methods do not also lead to selection errors. Battaglia et al. (2007) reported that many of the selection errors in their mail surveys occurred because the person who was chosen within a household refused to participate, whereas someone else in the household was willing to do so. This explanation is likely to hold also for selection errors that occur with the other less invasive methods.

Using data from three RDD landline surveys, Lavrakas et al. (2000) reported that selection errors in the last birthday method were not linked systematically to interviewer behavior. They also reported that these errors were more likely to occur in households with lower educational attainment and larger-sized households. Lind et al. (2000) also found that misselection with the last birthday method was correlated with lower educational attainment, whereas Goyder, Basic and Thompson (2004) found it correlated with being foreign born.

Most importantly, the Lavrakas et al. (2000) research found evidence for some biases linked to the selection errors associated with the birthday method. These biases were statistically significant and large enough to be non-ignorable and affected some political variables and some media-related variables, although there was not a consistent trend of significant variables across their three studies.

Two studies reported by Denk and his colleagues in the late 1990s made a rigorous effort to measure the trade-offs in Total Survey Error and in survey costs associated with different respondent selection methods (Denk, Guterbock, and Gold, 1996; Denk and Hall, 2000). These authors reported two major findings:

1. Respondent selection methods that use names or gender to identify the selected respondent *increase* the chance by upwards of one-third that the interview will be completed by someone other than the person within a household who is initially contacted. However, this improvement in coverage which resulted from "handoffs" to another household member led to a *decline* in cooperation.

2. Selection procedures which require handoffs do result in a *reduction* of the proportion of interviewed respondents who are female, but this improvement in sample characteristics mostly led to negligible effects on survey estimates, as well as *lowering* overall cooperation and thus *raising* costs.

Summary

The Kish method is accepted as the "gold standard" for within-household selection for interviewer-administered surveys. It is not deemed practical to deploy in mail surveys because of its operational complexity and often is considered too costly for use in telephone surveys. When used by skilled interviewers it yields random and representative within-household coverage of the target population. However, there are legitimate concerns that it raises nonresponse due to its invasive nature, and thereby raises survey costs, but it is not known if it raises nonresponse bias.

Among the less invasive selection methods that have been devised, implemented, and tested over the past three decades, the birthday methods have the greatest *potential* for selecting a within-household sample that is as near as possible to the randomness of the Kish method. These methods generally are favored by interviewers as they are far less invasive than the Kish method. They generally are favored by researchers as they are perceived as being less likely to cause nonresponse and they are less costly to utilize.

However, the birthday methods have been found to lead to a non-ignorable amount of selection errors in which the "wrong" person becomes the designated respondent for her/his household. These selection errors are not all random and lead to an overrepresentation of females and those most willing to participate in surveys. More importantly, these selection errors are disproportionately associated with lower educated, foreign-born, and large-sized households. There also is some evidence these errors contribute bias associated with some political and media-related variables.

Respondent Selection and Cell Phone Surveying

The issue of respondent selection in telephone surveys reaching cell phone numbers is not addressed in this report, as the state of knowledge on this topic is not yet advanced enough for there to be an accepted need to use such a selection method when reaching someone on a cell phone or an acceptable method for doing so (see Brick, Edwards, and Lee [2007]; and Lavrakas, Shuttles, Steeh, and Fienberg [2008]). Most telephone researchers in the U.S. appear to assume that the cell phone is a personal device and that within-unit respondent selection is not needed when sampling someone via a cell phone number.

At present, there appears to be no valid empirical information regarding the proportion of personal cell phones in the U.S. that are shared by more than one person to receive incoming calls, such as a survey request. However, Brick et al. (2007) reported that 11% of cell-phone only households with two or more adults shared a cell phone number in their California health pilot study. These authors noted that their "study was not definitive about the ability to interview a different sampled adult when the cell phone is shared" (p. 810). Thus, until more empirical studies have been conducted, the issue of whether a within-unit respondent selection procedure should be used when sampling cell phone number remains uncertain. Furthermore, if it is determined in the future that a selection procedure is required in cell phone surveys, it is not certain whether the same selection procedures used to sample a respondent in a landline survey will be appropriate for cell phone surveys.

A Promising New Approach to Within-Household Respondent Selection

Subsequent research reported from Westat (Rizzo, Brick, and Park, 2004), reinforced the findings of Denk et al. (1996; 2000), Lavrakas et al. (2000), and others, that the popular birthday methods too often lead to the misselection of respondents which result in "too many" women ending up as the respondent for their households.

More importantly, Rizzo et al. (2004) reported on a new method for respondent selection that has a solid scientific foundation and a great deal of commonsensical appeal. The new method essentially treats households as one of three types, based on the number of eligibles in the household. The method begins by determining how many eligible persons reside in a household. For example, for landline telephone surveys in which the only eligibility criterion is being 18 years of age or older – which is the case for the majority of general population surveys – in order to start the within-unit selection process, all that needs to be asked of the adult household member who first is contacted, is a question along the following lines:

Including yourself, how many people aged 18 or older currently live in this household?

Households with One Eligible. If there is only one eligible person then the interview commences with the person already being spoken to.

Households with Two Eligibles. If there are two eligible persons, then a random process is used to alert the interviewer to either proceed to interview the person already being spoken to or to ask for the other eligible person (which for most survey designs would occur half of the time when there are two eligibles).

Households with Three or More Eligibles. If three or more eligible people reside in the household, then another selection process is used to choose only one of them to interview. Several processes can be used for this selection depending on the rigor with which the survey organization wants to have the selection made.

In the U.S. less than one in six households has three of more adults residing in them. As such, this within-unit procedure is noninvasive for the vast majority (> 80%) of U.S. households. To date, the Westat method or variations of it appear to offer the greatest benefits between balancing concerns about within-unit coverage against concerns for nonresponse.

The Westat method also can be implemented in ways that will substantially reduce misselection in interviewer-administered surveys.

Recommended Methods to Use for Respondent Selection

Considering the state of existing knowledge in this topic area, four overarching conclusions can be drawn:

1. A systematic within-household respondent selection method should be used in surveys that first sample households and then pick one eligible person within the household to serve as a designated respondent. Otherwise the resulting sample at the person-level will be comprised of the "most available" and "most willing" persons. This will lead to samples that over-represent females, the elderly, those without fulltime employment, and those generally most disposed to cooperate with surveys. It will under-represent the hardest to reach respondents (e.g., young adult males) and those generally least disposed to cooperate with surveys. Because of these coverage errors survey estimates often will be biased.

2. There is no compelling evidence to suggest that the Kish method needs to be used in order to avoid nonignorable coverage error.

3. Unless a survey organization has reliable empirical evidence to the contrary specific to a given survey they conduct, deploying a birthday method to *all* households is not advisable. This follows from the compelling evidence in the literature that the birthday methods lead to a misselection of the proper respondent in a sizable minority of households with two or more eligibles and that these errors are not random and thus very likely introduce bias into survey estimates.

4. All factors considered, some variant of the procedure described by Rizzo et al. (2004) appears to be the "best" all around respondent selection method to deploy in interviewer-administered landline telephone surveys.

5. More experimental research is needed to determine what method is best to use for mail surveys. (See the later section for suggestions.)

The following sections describe a variation of the Rizzo et al. (2004) within-unit respondent selection method for RDD landline telephone surveys. An approach for selection within mail surveys is also described, but this approach requires further testing before it can be used with full confidence. Of note, the variations described in the following sections "blend" in some procedures that are have been used successfully by the Pew Research Center for more than a decade (cf. Keeter and Fisher, 1997; Dimock and Craighill, 2005).

For a Landline Telephone Survey

For a landline telephone survey of households, including all RDD surveys, in which one and only one eligible person within the household is selected to be interviewed, a variation of the Westat method, as shown in Table 1, is recommended. This assumes that a CATI system is used for interviewing and that the CATI software will control the required randomizations. The example in Table 1 also assumes that the only eligibility criterion is that the selected respondent be 18 years of age of older. If there are other eligibility criteria, then the first question in the sequence would need to be modified in order to determine the total number of eligibles in the household.

The method shown in Table 1 (the variation with a randomized next/last birthday selection used in households with three or more eligibles) was used by SRBI/ABT for a national RDD survey of

1,002 adults conducted for the Associated Press and AOL in late March and early April, 2008. It also was used by Eastern/GfK for eight state pre-election surveys for the Associated conducted in October, 2008, with 6,347 adults.[4] Using this modified version of the Westat method led to an unweighted mix of 46% male respondents and 54% female respondents in the March/April survey and an unweighted mix of 43% male and 57% female (using a short field period in which it is known to be harder to contact males) in the October survey. This compares to male-female gender ratios of 40:60 (or even more extreme) that result with many RDD surveys that utilize merely one of the birthday methods for respondent selection in all households. (Of note, no increase in nonresponse was observed using this selection method compared to what the AP typically experiences using just the birthday method.)

Table 1

Recommended Within-Household Respondent Selection Script for a Landline Telephone Survey of Households

(Assumes Eligibility is Only that Someone is 18 years of Age or Older)

1. Interviewer needs to be speaking with an adult member of the household.

2. The adult member is asked:

Including yourself, how many people aged 18 or older currently live in this household?

If 1, then interviewer proceeds to interview the adult being spoken to.

If 2, then CATI randomly selects either the adult being spoken to for the interviewer to proceed to interview or alerts the interviewer to ask for the other adult to interview.

If > 2, then CATI proceeds to #3 below.

3. FOR A RANDOM SELECTION METHOD:

Please tell me the name or initials of each of these persons.

After that information is recorded, CATI randomly selects one of these adults and the interviewer either (a) proceeds to interview the adult being spoken to if s/he is the one selected or (b) asks for the other adult who has been selected to interview.

FOR A QUASI-RANDOM METHOD:

Randomly assign these households to either the last or the next birthday method to select one of the adults.

FOR A SYSTEMATIC, BUT NONRANDOM METHOD:

[4] Lavrakas was a consultant on these studies and devised the modified Westat selection method that was deployed.

> *Among these persons, may I please speak with the youngest male adult currently at home?*
>
> If no adult male is at home, the interview asks (with the phrasing "youngest" and "oldest" being randomly assigned:
>
> *Then, may I please speak with the (oldest)(youngest) female adult currently at home?*

The nonrandom method in Table 1 that is used to select one respondent in households with more than two eligibles is adapted from the method used by the Pew Research Center (cf. Keeter and Fisher, 1997; Dimock and Craighill, 2005). In this method, there is a randomization applied that varies whether it is the "youngest female" or "oldest female" that is selected if the youngest male is not at home at the time of interviewer contact. Pew reports that this method does a very good job of selecting a sample whose age and gender characteristics closely match census UEs for the target population. The advantages of the Pew method is that it identifies a specific person by gender and in the case of selecting a female it does so with some regard to age.

All things considered, use of the modified-Westat method with the systematic nonrandom selection for households with three or more eligibles, as shown in Table 1, is appealing for many reasons. First, in one- and two-person households – which constitute more than 80 percent of U.S. households – it achieves the same end as Kish, but by asking only one relatively noninvasive question. In households with three or more eligibles, it actually does better than Kish when using the randomized selection procedure because its effectiveness is not altered in households with seven or more eligibles, and it accomplishes this by asking only a second question, thereby avoiding the time consuming invasiveness of Kish. The randomized selection procedure in table 1 for large households does not yield a within-household sample stratified by gender, as Kish does, but that potential benefit should be negligible in most surveys.

Second, from a nonresponse standpoint, the modified-Westat method should lead to fewer refusals than with Kish in most surveys.

Third, unlike the birthday selection methods, there will be no chance of misselection in households with two or more eligibles, as the person being spoken to is not actively determining whether the correct person is being chosen. The misunderstandings that occur with the birthday methods also are avoided with the randomized approach (when deployed in households with three or more eligibles) since there are no instructions that must be followed and no choices are being made by the person being spoken to.

Fourth, the potential basing effects of the misselections associated with the birthday methods are avoided.

Fifth, the resulting sample may have a better gender and age balance than the birthday methods and certainly will be better that the other non-invasive selection methods that have been deployed as alternatives to Kish for the past three decades.

For a Mail Survey

It is much more of a challenge to accurately implement a within-unit respondent selection technique in a self-administered mail survey than in surveys that are interviewer administered. Furthermore, there has been very little reported (in publications or conference papers) about within-household respondent selection in mail surveys. As Battaglia et al. (2008) observe, "…there is virtually no literature on random respondent selection for address-based sampling other than an occasional mention of using a birthday method or an assertion that the Kish method does not work well" (p. 460).

No matter what researchers do, the rate at which the "wrong" person in the household – due to both misselection and to someone else completing the questionnaire because the selected person is not willing or available to do so – will be greater in a mail survey than in a telephone or an in-person survey. For example, Battaglia et al. reported that when using a next-birthday selection method, 36% of households in a large mail survey had the questionnaire completed by someone who was not the selected respondent, and that these mistakes occurred in 60% of households with three or more adults. These researchers also determined that few of these errors were due to mistakes in identifying the correct person, but rather were because the correct person did not want to particulate and someone else was willing to do so instead.

This notwithstanding, there are levels of rigor that can be used to reduce the proportion of times the "wrong" person completes the questionnaire and improve the sample characteristics of the final sample in a mail survey compared to what happens currently when mail surveys try to use a birthday selection method for all households.

Two different approaches are considered here, both of which require a randomization of the selection instructions sent households. Since neither of these approaches has been formally tested, it is recommended that a survey organization use an experiment, such as that described in the following section of this report as part of their initial deployment of either mail method.

1. Battaglia, Link, Frankel, and Mokdad (2007) investigated within-household selection for a mail survey. A conclusion these researchers reached is that if the birthday method is used, then it should be a split-half *mix* of "birthday" methods – one in which a random half of the sample gets a cover letter explaining that the questionnaire should be completed by the adult in the household who will have the *next* birthday and the other random half of the sample would receive a cover letter explaining that the questionnaire should be completed by the adult in the household who had the *last* birthday. This mix will not eliminate all selection errors that occur, but will help make the selection more representative of the target population. This mix of next/last birthday methods will improve the distribution (balance) of the months in which the designated respondents are born throughout the entire year, rather than having the skewed distribution of months of birth that results when only one of the birthday methods is deployed.

   Furthermore, under no circumstances should the word "celebrate" be used in selection instructions, because some people do not celebrate their birthdays in the sense that

they do not have a party or other festivities to recognize their birthday. Instead, wording to the following effect should be used:

> *This questionnaire should be completed by the adult (18 years of age or older) who lives in your household whose birthday* ***(is occurring today or occurred most recently prior to today)(will be the first to occur after today).***

Finally, these respondent selection instructions should be printed on the first page of the questionnaire and *the questionnaire should start by asking the person filling out the questionnaire to answer affirmatively* that s/he is in fact the correct person that was selected within the household. This additional precaution will not eliminate all instances of misselection but certainly should reduce them.

2. An alternative approach for mail surveys would be to adapt the method used by Pew to systematically select a within-household sample that is more in line demographically with the gender and age UEs of the target population. To this end, a random half of the households in the sample would be sent the following instructions in a cover letter:

> *This questionnaire should be completed by the* ***youngest male adult*** *(18 years of age or older) who lives in your household. If there are no male adults living in your household, the questionnaire should be completed by the* ***oldest female adult*** *living in your household.*

The other random half of the sample would be sent the following alternative instructions:

> *This questionnaire should be completed by the* ***youngest male adult*** *(18 years of age or older) who lives in your household. If there are no male adults living in your household, the questionnaire should be completed by the* ***youngest female adult*** *living in your household.*

These respondent selection instructions also should be printed on the first page of the questionnaire and the questionnaire should start by asking the person filling out the questionnaire to *answer affirmatively* that s/he is in fact the correct person that was selected within the household. This additional precaution will not eliminate all instances of misselection but certainly should reduce them.

New Research to Conduct to Improve Respondent Selection

This section identifies two new research studies which MRC audited companies that use a within-household respond selection method should consider conducting in order to improve the empirical knowledge available to them to decide how best (from a Total Survey Error standpoint) to select a respondent within a household.

New Landline Telephone Survey Research Study

The purpose of this study is to provide a rigorous dimensioning of the effect of within-unit respondent selection on nonresponse error vs. coverage error in a landline telephone survey, something that never has been done adequately.

The decision about which selection approach is best to use is best done on a survey-by-survey basis, as there are idiosyncrasies across surveys that could lead to different findings and conclusions that are unique to the specific survey of interest. Given that the surveys that are audited by the MRC are conducted over and over by their companies, studying this issue within the context of a specific survey would be highly beneficial to the company and its clients.

This study design would randomly assign sampled telephone numbers to one of five conditions, as follows:

> **Condition 1:** This would be the control condition and would utilize the company's current within-home respondent selection scheme (e.g., the birthday method).
>
> **Condition 2:** This would utilize a modified version of the Westat approach with the random procedure for households with > 2 eligible members, modeled after the one described earlier in this report.
>
> **Condition 3:** This would utilize a modified version of the Westat approach with the quasi-random birthday procedure for households with > 2 eligible members, modeled after the one described earlier in this report, with half of the households with three or more eligibles being assigned to the last birthday condition and the other half being assigned to the next birthday condition.
>
> **Condition 4:** This would utilize a modified version of the Westat approach with the Pew nonrandom procedure for households with > 2 eligible members, modeled after the one described earlier in this report, with the "youngest woman" and "oldest woman" being randomly alternated.
>
> **Condition 5:** This would utilize the Pew procedure as Pew deploys it, *not* as it is described earlier in this report. In this condition, for all households, the interviewer does *not* ask how many eligibles there are in the household, but instead merely asks to speak with the youngest male adult currently at home. If there is no male adult resident at home at the time of the call then the interviewer asks to speak with either the youngest or oldest female adult at home. A randomization procedure is used to decide whether the interviewer asks for the "youngest' or "oldest" for a given household.

To investigate the effects on coverage of these five selection methods, comparisons of the demographic characteristics of the selected respondents, in particular for gender, age, and education, will be made for the subsamples that result against what the person-level UEs are for the characteristics. The "best" method will be the one that most closely matches the UEs along with having the most heterogeneous sample characteristics.

To investigate the effects on nonresponse of these five methods, various response rates will be computed for each subsample and compared with those achieved in the other samples. The "best" method will be the one that achieves the highest response rate with the fewest break-off/refusals and requires the least number of average callbacks per completed case.

The possibility of coverage error and/or nonresponse error can be investigated by observing differences in key survey estimates associated with each selection method. This is not as rigorous an analysis as those above, but it will need to suffice as gathering more rigorous data for studying these errors would be prohibitively expensive.

Of great importance, in order to maintain the integrity of this experimental design, *interviewers and telephone numbers must be randomly assigned to one, and only one, of these five conditions*. The need to randomly assign interviewers follows from the fact that if this is not done, the experiment will be confounded. The confound will stem from the uncontrollable and differential reactions that individual interviewers would have to the various selection methods they were using, if they were assigned to multiple selection methods. These differential reactions would make it impossible to sort out whether it was the effects of the selection procedures that led to any observed differences in the survey measures or whether it was merely that interviewers performed differentially under the various selection methods – for example, because of idiosyncratic preferences they had for one selection procedure over the other. Ideally, interviewers should be kept blind to the fact that the experiment is being conducted and that there are other selection methods being used beyond the one they are using.

New Mail Survey Research Study

Similar to the study described above, the purpose of this study is to provide a rigorous dimensioning of the effect of within-unit respondent selection on nonresponse error vs. coverage error in a mail survey, something that never has been done adequately.

There would be two experimental factors that would be manipulated and controlled by the researchers, crossed in all possible ways, thus yielding a full factorial design. One factor would be the selection method. The other factor would be how the method is deployed within the mailing. This design would allow for the testing of two main effects and one interaction effect for each dependent variable.

This study design would randomly assign sampled addresses to one of six conditions, as follows:

> **Condition 1:** This would be the control condition and would utilize the company's current within-home respondent selection scheme (e.g., the birthday method) along with mentioning it in the cover letter and at the beginning of the questionnaire.
>
> **Condition 2:** This would be a variation of the control condition in which the company's current selection procedure is used, but in addition a question will be added to the very beginning of the questionnaire asking that the respondent actively affirm that s/he is in fact the correct adult in the household.

Condition 3: This would utilize the mixed "birthday" selection approach suggested by Battaglia et al. (2007) and described earlier in this report. The selection instructions would appear in the cover letter and at the beginning of the questionnaire.

Condition 4: This would be similar to Condition 3, but in addition, a question will be added to the very beginning of the questionnaire asking that the respondent actively affirm that s/he is in fact the correct adult in the household.

Condition 5: This would utilize the Pew selection approach reported by Dimock and Craighill (2005) and described earlier in this report. The selection instructions would appear in the cover letter and at the beginning of the questionnaire.

Condition 6: This would be similar to Condition 5, but in addition, a question will be added to the very beginning of the questionnaire asking that the respondent actively affirm that s/he is in fact the "youngest male adult" or either the "youngest female adult" or "oldest female adult," depending on which of those two latter conditions the address is randomly assigned.

To investigate the effects on coverage of these two factors, comparisons of the demographic characteristics of the selected respondents, in particular for gender, age, and education, will be made for the six subsamples that result against what the person-level UEs are for the characteristics. The "best" methods (recall that two main effects are being tested in this study) will be the ones that most closely match the UEs along with having the most heterogeneous sample characteristics.

To investigate the effects on nonresponse of these two factors, various response rates will be computed for each subsample and compared with those achieved in the other subsamples. The "best" methods will be the ones that achieve the highest response rate and require the least number of follow-up mailings to previously nonresponding households, if such mailings are deployed in the company's survey design.

The possibility of coverage error and/or nonresponse error can be investigated by observing differences in key survey estimates associated with each factor. This is not as rigorous an analysis as those above, but it will need to suffice as gathering more rigorous data for studying these errors would be prohibitively expensive.

References

Battaglia, M. P., Link. M. W., Frankel, M. R. and Mokdad, A. H. (2007). An Evaluation of Respondent Selection Methods for Household Mail Surveys. *2006 Proceedings of the American Statistical Association Section on Survey Research Methods*, pp. 2727-2731; Alexandria VA: American Statistical Association.

Battaglia, M. P., Link. M. W., Frankel, M. R., Osborn, L., and Mokdad, A. H. (2008). An Evaluation of Respondent Selection Methods for Household Mail Surveys. *Public Opinion Quarterly*, 72(3), 459-469.

Brick, J. M., Edwards, W. S. and Lee, S. (2007). Sampling Telephone Numbers and Adults, Interview Length, and Weighting in the California Health Interview Survey Cell Phone Pilot Study. *Public Opinion Quarterly*, 71(5), 793-813.

Bryant, B. E. (1975). Respondent Selection in a Time of Changing Household Composition. *Journal of Marketing Research*, 12, 129-135.

Denk, C. E., Guterbock, T. M. and Gold, D. B. (1996). Modeling Selection of Respondents Within Households in Telephone Surveys. Paper presented at the 51st annual conference of the *American Association for Public Opinion Research*; Salt Lake City, UT.

Denk, C. E., and Hall, J. W. (2000). Respondent Selection in RDD Surveys: A Randomized Trail of Selection Performance. Paper presented at the 55th annual conference of the *American Association for Public Opinion Research*; Portland, OR.

Dimock, M. and Craighill, P. (2005). Unpublished memo on "2005 Household Randomization Experiment." Pew Research Center, Washington DC; October 18.

Hagen, D. E. and Collier, C. M. (1983). Must Respondent Selection Procedures for Telephone Surveys Be Invasive? *Public Opinion Quarterly*, 47, 547-556.

Gaziano, C. (2005). Comparative Analysis of Within-Household Respondent Selection Techniques. *Public Opinion Quarterly*, 69, 124-157.

Goyder, J., Basic, M. and Thompson, M. (2001). Within Household Sample Selection: A Reassessment. Paper presented at the 56th annual conference of the *American Association for Public Opinion Research*; Montreal.

Grandjean, B. D., Leighty, M. G., Taylor, P. A. and Xu, Y. (2004). Is Target Selection by Last Birthday "Random Enough"? A Split-Ballot Test. *ASA/AAPOR 2004 Conferences Proceedings*, 4789-4793.

Keeter, S and Fisher, K. (1997). A Comparison of "Last Birthday" and "Youngest Male/Oldest Female" Respondent Selection Procedures. Paper presented at the 52nd annual conference of the *American Association for Public Opinion Research*; Norfolk, VA.

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

Lavrakas, P. J., Bauman, S. L. and Merkle, D. M. (1994). The last birthday selection method and within-unit coverage problems. *American Statistical Association 1993 Conference Proceedings,* 1107-1112.

Lavrakas, P. J., Harpuder, B., and Stasny, E. A. (2000). The Last-Birthday Selection Methods and Within-Unit Coverage Error. Paper presented at the 55th annual conference of the *American Association for Public Opinion Research*; Portland, OR.

Lavrakas, P. J., Steeh, C., Blumberg, S., Boyle, J., Brick, J. M., Callegaro, M. et al. (2008). *Guidelines and Considerations for Survey Researchers When Planning and Conducting RDD and Other Telephone Surveys in the U.S. With Respondents Reached via Cell Phone Number.* Lenexa, KS: American Association for Public Opinion Research. http://www.aapor.org/uploads/Final_AAPOR_Cell_Phone_TF_report_041208.pdf

Lind, K. Link, M. and Oldendick, R. (2000). A Comparison of the Accuracy of the Last Birthday versus the Next Birthday Methods for Random Selection of Household Respondents. *American Statistical Association 2000 Conference Proceedings,* 887-889.

Nemeth, R. (2001). "Respondent Selection Within the Households: A Modification of the Kish Grid." Paper presented at 6th *Austrian, Hungarian, Italian, and Slovenian Meeting of Young Statisticians*; Ossiach, Carinthia, Austria.

Rizzo, L. J., Brick, J. M. & Park, I. (2004). A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys. *Public Opinion Quarterly*, 68, 267-274.

Salmon, C. T. and Nichols, J. S. (1983). The Next-Birthday Method of Respondent Selection. *Public Opinion Quarterly*, 47, 270-276.