A Meta-analysis of Within-Household Respondent Selection Methods

Ting Yan[1]
[1]NORC at the University of Chicago, 55 E. Monroe Street, Chicago, IL 60615

Abstract
Random selection of a respondent within a sampled household is essential for maintaining the probability nature of the resulted sample and for making reference from the sample to the general population. The latest review on within-household respondent selection methods shows that available methods differ on the ability to maintain a truly probability sample and the extent of perceived intrusiveness and burden to respondents. However, the review fails to produce effect sizes quantifying the impact of different within-household selection methods. This paper filled the gap by conducting a meta-analysis with the goal to generate quantitative effect sizes indicating the size of the impact of different selection methods. This research advances the survey literature on within-household respondent selection methods and will be of practical significance to survey organizations, who can make use of the quantitative effect sizes to guide their decisions on which within-household selection method to use in their surveys.

Key Words: within-household respondent selection methods, meta-analysis, demographic representativeness, cooperation rates

1. Introduction

Surveys of general population usually start with a probability sample of households identified by telephone numbers or mailing addresses. However, a probability sample of households doesn't automatically translate into a probability sample of people with characteristics of interest. Random selection of a respondent within a sampled household is essential for maintaining the probability nature of the resulted sample and for making reference from the sample to the general population. A good within-household respondent selection method should be able to randomly select a respondent within the household without appearing intrusive or burdensome to potential respondents. The latest review (Gaziano, 2005), the most comprehensive review of all existing within-household selection methods, listed at least 15 selection methods that have been used to generate a sample of persons from a probability sample of households. Gaziano grouped the 15 selection methods into three categories based on the probability nature of the resulted sample of persons – probability methods, quasi-probability methods, and non-probability methods (Gaziano, 2005).

Probability methods require a listing of all people living in the sampled household in order to compute the probability of selection for each individual. The advantages of probability methods lie in their ability to produce consistent and unbiased estimates and the property of measurability essential for calculating sampling variance. However, the biggest drawback of probability methods is the perceived intrusiveness and burden associated with the household listing. It is commonly believed that the listing of all household members add to the length of the interview and increase the likelihood of encountering household refusals due to the sensitiveness of the listing questions. Kish method (Kish, 1949), Age-order or Age-only method (Denk and Hall 2000; Forsman, 1993), and full enumeration methods (Denk et al., 2000; Srinivasan et al., 1996) fall into this category.[1]

Quasi-probability methods bypass household listing in order to reduce the perceived intrusiveness and sensitivity associated with household listing and to decrease the administration time. However, the bypass of household listing increase cooperation and reduces cost at a sacrifice of randomness. Typical examples of quasi-probability selection methods are birthday methods, including next birthday method (Salmon & Nichols, 1983) and last birthday or most recent birthday method (Salmon & Nichols, 1983). According to Gaziano (2005), birthday methods allow all household members to have an equal chance of selection under the assumption that births are random. However, as Gaziano (2005) and others acknowledged, births are not truly random and tend to heap in certain months. Last birthday method is believed to be the most widely used and is most prevalent in comparative studies

Non-probability methods trade randomness for increased cooperation and reduced cost. They were created to streamline selection process and only tried to approximate population age and gender distributions (Gaziano, 2005). Troldahl and Carter method (Troldahl & Carter, 1964) is a typical example of non-probability methods, so are various modifications or variants on Troldahl and Carter method.[2]

Gaziano (2005) reviewed 16 comparative studies where two or more within-household selection methods were compared in experiments. Different methods were compared in terms of their impact on demographic representativeness, cooperation and refusal rates, and costs. The review focused on 5 dimensions for comparison: Kish vs. Last Birthday, Last Birthday vs. Youngest Male and Oldest Female, Next Birthday vs. any other method, Troldahl and Carter method and its variations, and no selection vs. any selection method.

Gaziano's review is comprehensive and extensive. However, a major drawback is that the review fails to produce effect sizes quantifying the impact of different within-household selection methods. As concluded by Gaziano, "little systematic, accessible evidence exists to guide choice of respondent selection method" (2005, 124). This paper filled the gap by conducting a meta-analysis of within-household respondent selection methods with the goal to generate quantitative effect sizes indicating the size of the impact of different selection methods.

2. Meta-analysis of Within-Household Selection Methods

2.1 Selection of Studies

I searched for empirical reports of studies comparing two or more within-household selection methods, focusing on studies using random assignment of the survey

[1] Please refer to Gaziano (2005) for detailed description of each method.

[2] Other variations or modifications to the Trodhal and Carter method include Paisley and Parker Standard Modification (1965), Bryant's Correction for Too Many Females (1975), Groves and Kahn's Modification (1979), Czaja-Blair-Sebestik modification (1982), Hagan-Collier Alternative (1983), Youngest Male/Oldest Female (YMOF). Again, please refer to Gaziano (205) for detailed description of each method.

respondents to each within-household selection method being compared. I searched various databases available (e.g., JSTOR, Ebsco, LexisNexis, PubMed) and supplemented this with online search engines (e.g., Google Scholars), using as key words within-household, respondent, selection, method, survey. I also searched *the Proceedings of the Survey Research Methods Section of the American Statistical Association*. These proceedings publish papers presented at the two major conferences of survey methodologists (the Joint Statistical Meeting and the annual conferences of the American Association for Public Opinion Research) where survey methods studies are often presented.[3] In order to uncover unpublished studies or new studies in press, I posted emails to relevant communities (e.g., AAPORNET) looking for studies evaluating or comparing within-household section methods. I also contacted Ms Gaziano, who is kind enough to hand over hard-copy papers or presentation slides.

I included studies in the meta-analysis if they randomly assigned survey respondents to one of the within-household respondent selection methods and if they provided quantitative numbers. However, I dropped a few studies that didn't include enough information to estimate effect sizes; for instance, some reported means but not standard deviations. I also dropped some studies that provided statistics not appropriate for meta-analysis. A total of 23 studies met my inclusion criteria. Most of the papers were conference papers presented at the two major conferences of survey research. Two were published in the *Proceedings of the Survey Research Methods Section of the American Statistical Association*. I used the Duval and Tweedies's (2000) trim-and-fill procedure to test for the omission of unpublished studies. In none of the meta-analyses reported here did I detect publication bias.

2.2 Analytic Procedures

As in Gaziano (2005), the meta-analysis was conducted to assess the following comparisons:

- Kish method vs last birthday method
- Kish method vs. any other selection methods
- Last birthday vs. Youngest Male Oldest Female
- Birthday methods (including both Last birthday and Next Birthday) vs. any other selection methods
- Probability methods vs. any other selection methods
- No selection method vs. any selection method

Since most studies provided the proportion of respondents by sex in the resulted sample, I compared proportions of females and males in the resulted sample by different selection methods as a key measure of demographic representativeness.

Studies report different forms of response rates, refusal rates, and completion rates. I adopted a simple completion rate used in Gaziano (2005) in order to compare across studies. This simple completion rate is simply the number of completed interviews over the sum of completes and refusals. This is roughly equivalent to COOP3 in AAPOR standards (AAPOR, 2008). Thus, the second comparison measure is the proportion of completes and refusals in the resulted sample by different selection methods as a measure of response rates and cooperation rates.

---

[3] As a matter of fact, one session was devoted to within-household selection methods at the 1993 AAPOR conference.

For the meta-analysis, effect sizes could be the raw proportion differences (e.g., the difference in the proportion of females or completed interviews resulted from different selection methods). However, raw proportions are in effect unstandardized means and can inflate the apparent heterogeneity across studies (see the discussion in Lipsey and Wilson, 2001). As a result, odds ratios (OR) were calculated as follows:

$$OR_{sex} = \frac{p_{f,method1} * (1 - p_{f,method2})}{p_{f,method2} * (1 - p_{f,method1})},$$

$$OR_{COOP} = \frac{p_{C,method1} * (1 - p_{C,method2})}{p_{C,method2} * (1 - p_{C,method1})}$$

Odds ratios compare the odds of obtaining females (or completed interviews) under different selection methods. I then took the log of each odds ratio and used the log odds ratio as the effect size to be summarized in the meta-analysis.

I carried out each of the meta-analyses reported in this paper using the Comprehensive Meta-Analysis (CMA) package (Borenstein, Hedges, Higgins, & Rothstein, 2005). That program uses the simple average of the effect sizes from a single study (study *i*) to estimate the study-level effect size (ES_i below); it then calculates the estimated overall mean effect size ($\overline{ES}$) as a weighted average of the study-level effect sizes:

$$\overline{ES} = \frac{\sum_i^k ES_i}{\sum_i^k w_i}$$

The weight w_i is the inverse of the variance V_i for study *i* if a fixed effects model is assumed, which considers only sampling error associated with the sampling of individuals within study *i*. However, the fixed effects model assumption may not be correct in all circumstances. For instance, it is possible that the variability across effect sizes is greater than expected from sampling error alone. In this case, a random effects model should be used instead, which takes into account the excess variability stemmed from random differences between the studies that can not be modeled. The random effects model adds a between-study variance (the random effects variance component) to the adjusted variance:

$$V_i^* = V_i + \frac{Q-k-1}{\sum w_i - (\sum w_i^2 / \sum w_i)} \text{ where } Q = \sum w_i * (ES_i - \overline{ES})$$

and w_i is the inverse of V_i^* for study *i*.

The standard error of $\overline{ES}$ is computed as:

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum_i^k w_i}}.$$

2.3 Results

Study level effect sizes, the overall mean effect sizes under both the fixed effects and random effects models are shown in a series of tables below. These tables also present the homogeneity test results for each comparison. A statistically significant Q at *p*=.05 level suggest that the null hypothesis of homogeneity is rejected and random effects model is more appropriate than the fixed effects model. In that case, mean effect sizes under the random effects model are discussed. Otherwise, mean effect sizes under the fixed effects model are discussed.

2.3.1 Comparing Kish method to the Last Birthday method

Table 1 displays the studies, study effect sizes, and the overall mean effect sizes under both the fixed effects and random effects model. The overall mean effect size is .89 under both models. It is significantly different from 1 (*p*=.014). It suggests that the odds for getting females with the Kish method are statistically smaller than the odds of getting females with the Last Birthday method. In other words, the Kish method effectively reduces the percent of females in the resulted sample and has a better male representation. This finding confirms the advantage of a true probability selection method since it reduces the chance of women self-selecting themselves into the sample as they are more likely to be at home and answer the phone than the male counterparts.

Table 1. Comparison on Proportion of Females, Kish vs. Last Birthday Method

Study	Effect Size (Odds Ratio)	p-value
Kennedy (1993)	0.98	
Denk et al. (1996)	0.99	
Oldendick et al. (1998)-Experiment 1	0.87	
Oldendick et al. (1998)-Experiment 2	0.94	
Oldendick et al. (1998)-Experiment 3	0.89	
Denk et al. (2000)-Experiment 1	0.72	
Denk et al. (2000)-Experiment 1	0.77	
Fixed Effects Model	0.89	0.01
Random Effects Model	0.89	0.01
Q-value	5.02	0.53

In terms of the odds of obtaining completed interview (versus refusals), the mean overall effect size is .78 under the random effects model and is statistically significantly from 1 as seen in Table 2. The significant overall mean effect size suggests that the odds of obtaining completed interviews are smaller for Kish method than for the Last Birthday method. In other words, as survey researcher worried, Kish method has a significantly higher chance of eliciting refusals from the respondents than the less intrusive Last birthday method. It seems that the household listing process in the Kish method does appear more intrusive to respondents than the last birthday method.

Table 2. Comparison on Proportion of Completes, Kish vs. Last Birthday Method

Study	Effect Size (Odds Ratio)	p-value
Denk et al. (1996)	1.00	
Binson et al. (2000)	0.88	
Oldendick et al. (1988) -Experiment 1	0.95	
Oldendick et al. (1988) -Experiment 2	1.16	
Oldendick et al. (1988) -Experiment 3	0.37	

O'Rouke et al. (1983)	0.67	
Tarnai et al. (1987)	0.70	
Fixed Effects Model	0.78	0.00
Random Effects Model	0.78	0.01
Q-value	14.30	0.03

2.3.2 Comparing Kish Method to all other selection methods

Tables 3 and 4 display the study level effect sizes and the overall mean effect sizes for the comparison on cooperation rates and demographic representativeness. The overall mean effect size for the ratio of the odds of obtaining females in the resulted sample is .92 under the fixed effects model and .91 for the random effects model. Both are statistically significant from 1. It is clear that Kish method significantly reduces the odds of recruiting female respondents compared to any other selection method.

Table 3. Comparison on Proportion of Females, Kish vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
	Odds ratio	
Carr et al. (1993)	0.89	
Kennedy (1993) -Experiment 1	0.93	
Kennedy (1993) -Experiment 2	0.89	
Denk et al. (1996)	0.81	
Oldendick et al. (1998)-Experiment 1	0.87	
Oldendick et al. (1998)-Experiment 2	0.94	
Oldendick et al. (1998)-Experiment 3	0.89	
Czaja et al. (1982)	1.19	
Denk et al. (2000)-Experiment 1	0.82	
Denk et al. (2000)-Experiment 1	0.79	
Fixed Effects Model	0.91	0.01
Random Effects Model	0.91	0.01
Q-value	11.02	0.27

As in the previous comparison to the Last Birthday method, Kish method again has a significantly smaller odds ratio than any other selection method with regard to producing completes (the overall mean effect size is .80 under the fixed effects model and .79 for the random effects model. In other words, Kish method yields a higher refusal rate (and a lower completion rate) than any other selection method.

Table 4. Comparison on Proportion of Completed Interviews, Kish vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Denk et al. (1996)	1.00	
Binson et al. (2000)	0.84	
Oldendick et al. (1988)-Experiment 1	0.95	
Oldendick et al. (1988)-Experiment 1	1.16	
Oldendick et al. (1988)-Experiment 1	0.37	
Czaja et al. (1982)	0.89	

O'Rouke et al. (1983)	0.67	
Tarnai et al. (1987)	0.70	
Fixed Effects Model	0.80	0.00
Random Effects Model	0.79	0.00
Q-value	15.10	0.03

2.3.3 Comparing the Last Birthday Method to the Youngest Male Oldest Female Method (YMOF)

Four studies compared the two methods and presented information on gender distribution and on refusal and completed interviews. Table 5 shows that the overall mean effect size for the female proportions in the resulted sample 1.3 under both models, which means that the odds of obtaining female respondents are much higher under the Last Birthday method than under the YMOF method. It seems that asking for youngest males first and then oldest females ensured a better representation of male respondents in the sample than the Last Birthday method which doesn't have any control on gender.

Table 5. Comparison on Proportion of Females, Last Birthday Method vs. Youngest Male/Oldest Female Method

Study	Effect Size (Odds Ratio)	p-value
Keeter et al. (1997-98)	1.26	
Srinivsan et al. (1996)	1.38	
Hill et al. (1999)	1.46	
Keeter et al. (1997)	1.16	
Fixed Effects Model	1.33	0.00
Random Effects Model	1.31	0.00
Q-value	5.55	0.14

As shown in Table 6, the overall mean effect size for the odds ratio of obtaining completed interviews is .92 under both models. However, it is not statistically significantly different from 1 (p=.08). Thus, the Last Birthday method doesn't seem to have distinctive advantages over YMOF method in eliciting completed interviews.

Table 6. Comparison on Proportion of Completed Interviews, Last Birthday Method vs. Youngest Male/Oldest Female Method

Study	Effect Size (Odds Ratio)	p-value
Hill et al. (1999)	0.92	
Keeter et al. (1997)-Experiment 1	0.93	
Keeter et al. (1997)-Experiment 2	0.97	
Keeter et al. (1997)-Experiment 3	0.85	
Fixed Effects Model	0.92	0.08
Random Effects Model	0.92	0.08
Q-value	0.48	0.92

2.3.4 Comparing Birthday methods to all other selection methods

Altogether 13 studies report 16 experiments comparing birthday methods (including both the Last Birthday method and the Next Birthday Method) to other selection methods. Table 7 displays the study level effect sizes and the overall mean effect size. The overall mean effect size under the random effects model is 1.1 for the proportion of females and is statistically significant at p=.05 level. It suggests that birthday methods do have a higher chance of selecting female respondents than any other selection method.

Table 7. Comparison on Proportion of Females, Birthday Methods vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Forsman (1993)	1.17	
Carr et al. (1993)	1.12	
Praire Research Associates (2006)	1.11	
Kennedy (1993)	0.91	
Keeter et al. (1997-98)	1.26	
Battaglia et al. (2008)	1.00	
Denk et al. (1996)	0.84	
Beebe et al. (2007)	0.98	
Oldendick et al. (1988)-Experiment 1	1.15	
Oldendick et al. (1988)-Experiment 2	1.06	
Oldendick et al. (1988)-Experiment 3	1.12	
Goyder et al. (2001)	0.84	
Denk et al. ((2000)-Experiment 1	1.31	
Denk et al. ((2000)-Experiment 1	1.08	
Srinivsan et al. (1996)	1.28	
Hill et al. (1999)	1.46	
Keeter et al. (1997)	1.16	
Fixed Effects Model	1.15	0.00
Random Effects Model	1.12	0.01
Q-value	41.1	0.01

Information on completes and refusals is recorded in 13 experiments reported by 10 studies. As shown in Table 8, the overall mean effect size under the random model is 1.1 and it is not statistically significant at p=.05 level. The birthday methods are not advantageous to other selection methods in retaining completes.

Table 8. Comparison on Proportion of Completes Birthday Methods vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Praire Research Associates (2006)	0.83	
Denk et al. (1996)	1.00	
Beebe et al. (2007)	1.72	
Binson et al. (2000)	1.19	
Oldendick et al. (1988)-Experiment 1	1.05	
Oldendick et al. (1988)-Experiment 2	0.86	
Oldendick et al. (1988)-Experiment 3	2.68	
O'Rouke et al. (1983)	1.49	
Hill et al. (1999)	0.92	
Tarnai et al. (1987)	1.43	
Keeter et al. (1997)-Experiment 1	0.93	

Keeter et al. (1997)-Experiment 2	0.97	
Keeter et al. (1997)-Experiment 3	0.85	
Fixed Effects Model	1.10	0.00
Random Effects Model	1.13	0.14
Q-value	66	0.01

2.3.5 Comparing Probability Methods to all other selection methods

Eight studies report 11 experiments that compare a true probability selection method to any other selection method. The true probability method (e.g., Kish method) is believed to keep the randomness feature at the price of higher refusals. The overall mean effect size is 0.92 under both models. It is statistically significant at *p*=.05 level, suggesting that the true probability methods are more effective in reducing female representation in the resulted sample than any other methods. In other words, the random selection feature of the true probability methods ensures better sex representativeness.

Table 9. Comparison on Proportion of Females, Probability Methods vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Forsman (1993)	0.95	
Carr et al. (1993)	0.89	
Kennedy (1993) -Experiment 1	0.93	
Kennedy (1993) -Experiment 2	0.89	
Denk et al. (1996)	0.81	
Oldendick et al. (1988)-Experiment 1	0.87	
Oldendick et al. (1988)-Experiment 2	0.94	
Oldendick et al. (1988)-Experiment 3	0.89	
Czaja et al. (1982)	1.19	
Denk et al. (2000)-Experiment 1	0.77	
Denk et al. (2000)-Experiment 1	0.93	
Srinivsan et al. (1996)	1.00	
Fixed Effects Model	0.92	0.01
Random Effects Model	0.92	0.01
Q-value	12.26	0.34

Fewer studies report complete data on completes and refusals (see Table 10). The overall mean effect size under the random effects model is .79, suggesting that true probability methods suffer in terms of cooperation rates and tend to trade higher refusal rates for better representativeness than other selection methods.

Table 10. Comparison on Proportion of Completes, Probability Methods vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Denk et al. (1996)	1.00	
Binson et al. (2000)	0.84	
Oldendick et al. (1988)-Experiment 1	0.95	
Oldendick et al. (1988)-Experiment 2	1.16	

Oldendick et al. (1988)-Experiment 3	0.37	
Czaja et al. (1982)	0.89	
O'Rouke et al. (1983)	0.67	
Tarnai et al. (1987)	0.70	
Fixed Effects Model	0.80	0.00
Random Effects Model	0.79	0.00
Q-value	15.10	0.03

2.3.6 Comparing no selection method to any selection method

Six experiments compared the situation when no within-household selection was done to when any selection method was used. All experiments reported the sex distribution, but no complete data on completed interviews and refusals. Thus, this comparison is only done on the ratio of the odds of obtaining female respondents when no within-household selection was done versus when any selection method was used. As shown in Table 11, the overall mean effect size under the random effects model is 1.30. This finding confirms the concerns of survey researchers that women would be overrepresented if no within-household selection was done. Any selection method produces better representativeness than no selection.

Table 11. Comparison on Proportion of Females, No Selection vs. Any Other Method

Study	Effect Size (Odds Ratio)	p-value
Praire Research Associates (2006)	0.90	
Kennedy (1993) -Experiment 1	1.52	
Kennedy (1993) -Experiment 2	1.44	
Battaglia et al. (2008)	0.99	
Denk et al. (1996)	1.87	
Goyder et al. (2001)	1.19	
Fixed Effects Model	1.38	0.00
Random Effects Model	1.30	0.01
Q-value	15.41	0.01

3. Discussion and Conclusions

This paper reports a meta-analysis on the within-household selection methods. Unlike previous reviews, which tend to be narrative in nature, this paper took advantage of the analytical power of meta-analysis to summarize findings across different empirical studies and to produce quantitative effect sizes that can be used in future guidance. Gaziano's review concluded that "little systematic, accessible evidence exists to guide choice of respondent selection method" (2005, p124). However, the meta-analytic results seem to be clear cut as to which methods tend to over identify female respondents and which method has a lower odds of obtaining completes.

Several conclusions can be drawn through the meta-analysis:

1. Kish method reduces the chance to identify women than any other method (including the Last Birthday method). However, inconsistent with Gaziano's finding that "Kish also may be less intrusive than previously thought" (2005,

p150), the meta-analytic results make it clear that Kish has a lower odds of obtaining completed interviews than other methods, which suggest that Kish is as intrusive as previously thought.

2. Last birthday method has advantages over Kish in cooperation rates; but it doesn't have significant advantages over other methods (including the Youngest Male Oldest Female method and other methods) in producing completed interviews.
3. Birthday methods in general tend to over identify female respondents but do not seem to have obvious advantage in cooperation and completed interviews.
4. True probability methods reduce the odds to identify female respondents due to the random mechanism they employ but at a price of higher refusals.
5. Any selection method is better than no selection in reducing the odds of identifying female respondents.

This paper reports a partial meta-analysis on the within-household selection methods. The two key dependent variables in this paper are the proportion of females as a measure of demographic representativeness and the proportion of completed interviews as a measure of cooperativeness. Other important demographic variables such as race and age are not considered here. The next step is to conduct similar meta-analysis on those variables. The main limitation of this paper is that cost is not examined. This is because researchers measure cost in different metrics and report incomplete information. I urge future researchers to include more data in their paper and to report complete information whenever possible (e.g, standard deviations whenever a mean is reported). I also urge researchers to conduct and report experiments to thoroughly examine within-household selection methods.

Acknowledgements

I'd like to thank Dr. Roger Tourangeau for inspiring me and challenging me for the meta-analysis, Dr. Kirk Wolter and the Center for Survey Research Excellence of NORC for the financial support, and Ms. Gaziano for hand carrying several hard-copy conference papers to Chicago.

References

Battaglia, Michael, Michael W. Link, Martin F. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "An Evaluation of Respondent Selection Methods for Household Mail Surveys." *Public Opinion Quarterly* 72: 459-469

Beebe, Timothy J., Michael E. Davern, Donna D. McAlpine, and Jeanette K. Ziegenfuss. 2007. "Comparison of Two Within-Household Selection Methods in a Telephone Survey of Substance Abuse and Dependence." *Annals of epidemiology* 17:458-463.

Binson. Diane, Jesse A. Canchola, and Joseph A. Catania. 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.

Carr. Kathleen, and Joan Hertvik. 1993. "Within-Household Selection: Is Anybody Listening?" In *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 1119-1123.

Czaja. Ronald. Johnny Blair, and Jutta P. Sebestik. 1982. “Respondent Selection in a Telephone Survey: A Comparison of Three Techniques.” *Journal of Marketing Research* 19:381-85.

Denk. Charles E., Thomas M. Guterbock, and Daniel B. Gold. 1996. “Modelling Selection of Respondents Within Household in Telephone Surveys.” Paper presented at the annual meeting of the American Association for Public Opinion Research, Sail Lake City, UT.

Denk, Charles E., and John W. Hall, 2000. “Respondent Selection in RDD Surveys: A Randomized Trial of Selection Performance." Paper presented at the annual meeting of the American Association for Public Opinion Research, Poriland, OR.

Forsman, Gosta. 1993. “Sampling Individuals Within Households in Telephone Surveys.” In *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 1113-1118.

Gaziano, Cecile. 2005. “Comparative Analysis of Within-Household Respondent Selection Techniques.” *Public Opinion Quarterly* 69:124-157.

Goyder, John, Maryanna Basic, and M. E. Thompson. 2001. “Within Household Sample Selection: A Reassessment.” Paper presented at the annual meeting of the American Association for Public Opinion Research, Montreal, Canada. (Revised post-meeting version.)

Hagan, Dan E., and Charlotte Meier Collier. 1983. “Must Respondent Selection Procedures for Telephone Surveys Be Invasive?" *PublicOpinion Quarterly* 47:547-56.

Hill, Craig A., Karen Donelan, and Martin R. Frankel. 1999. “Within-Household Respondent Selection in an RDD Telephone Survey: A Comparison of Two Methods.” Paper presented at the annual meeting of the American Association for Public Opinion Research. St. Petersburg, FL.

Keeter, Scott. and Kevin Fisher. 1997. “A Comparison of ‘Last Birthday’ and ‘Youngest Male/Oldest Female’ Respondent Selection Procedures.” Paper presented at the annual meeting of the American Association for Public Opinion Research. Norfolk. VA.

____.1997-98. “Comparison of Two Respondent Selection Methods.” *National Network of State Polls (NNSP) Newsletter* 31 (Winter): 1-3. Available online at http://survey.rgs.uky.edu/nnsp/newsletters/nnsp31.pdf (accessed September 2009).

Kennedy, John M. 1993. “A Comparison of Telephone Survey Respondent Selection Procedures.” Paper presented al the annual meeting of the American Association for Public Opinion Research, St. Charles, IL. Available online at http://www.indiana.edu/~csr/aapor93.html (accessed September 2009).

Longstreth, Molly, and Todd Shields. 2005. “A Comparison of Within-Household Random Selection Procedures: The Rizzo, Brick & Park (RPB) and Last Birthday Methods.” Paper presented at the American Association for Public Opinion Research, Miami, FL.

Oldendick, Robert W., George F, Bishop, Susan B. Sorenson, and Alfred J. Tuchfarbor. 1988. “A Comparison of the Kish and Last Birthday Methods of Respondent Selection in Telephone Surveys.” *Journal of Official Statistics* 4:307-18.

O'Rourke, Diane, and Johnny Blair. 1983. “Improving Random Respondent Selection in Telephone Surveys,” *Journal of Marketing Research* 20:428-32

Prairie Research Associates, Inc. 2001. “Comparing Two Respondent Selection Methods in Telephone Surveys.” Manitoba, Canada. Available online at http://www.pra.ca/resources/birthday_e.pdf (accessed September 2009).

Sabin, Martha C., and Susan H. Godley. 1987. "Mental Health Citizen Surveys: A Comparison of Two Within Household Telephone Sampling Techniques." *Evaluation and Program Planning* 10: 137-141.

Salmon. Charles T., and John Spicer Nichols, 1983. "The Next-Birthday Method of Respondent Selection." *Public Opinion Quarterly* 47:270-76.

Srinivasan, Rajesh, Elaine H. Christiansen, and Robert Tortora. 1996. "Comparative Study of Three Respondent Selection Techniques for Telephone Survey Research." Paper presented at the annual meeting of the American Association for Public Opinion Research. Sail Lake City, UT.

Tarnai, John. Eugene A, Rosa, and Lesli Peterson Scott. 1987. "An Empirical Comparison of the Kish and the Most Recent Birthday Method for Selecting a Random Household Respondent in Telephone Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Hershey, PA.

Zukin, Cliff, Bob Carter, and Mark Schulman. 1987. "Trade-offs in Respondent Selection Methods: Theoretical Purity versus Real World Concerns." Paper presented at the annual meeting of the American Association for Public Opinion Research, Hershey, PA.

Appendix: Studies included in meta-analysis

	Study	Methods Compared
1	Battaglia, Link, Frankel, Osborn, and Mokdad (2008)	Next Birthday, no selection
2	Beebe, Davern, McAlpine, and Ziegenfuss (2007)	Rizzio-Brick-Park, Next birthday
3	Binson, Canchola, and Catania (2000)	Kish, Last birthday, Next birthday
4	Carr and Hertvik (1993)	Birthday, Kish (simulated)
5	Czaja, Blair, and Sebestik (1982)	Kish, T-C-B/men, T-C-B/woman
6	Denk and Hall (2000)	Last birthday, Kish (age-gender), age-order, full enumeration
7	Denk, Guterbock, and Gold (1996)	Kish (age-gender), last birthday, no selection
8	Forsman (1993)	Age-order, T-C/Woman, Birthday
9	Goyder, Basic, and Thompson (2001)	Next birthday, no selection
10	Hagen and Collier (1983)	T-C-B/men, H-C
11	Hill, Donelan, and Frankel (1999)	Last birthday, YMOF
12	Keeter and Fisher (1997)	Last birthday, YMOF
13	Keeter and Fisher (1997-1998)	Last Birthday, YMOF
14	Kennedy (1993)	Kish, T-C-B/men, H-C, Last Birthday, no selection
15	Longstreth M, Shields T. (2005)	Last Birthday, Rizzio-Brick-Park
16	Oldendick, Bishop, Sorenson, and Tuchfarbr (1998)	Kish, Last birthday
17	O'Rouke and Blair (1983)	Kish, Last birthday
18	Praire Research Associates (2001)	Next Birthday, no selection
19	Sabin and Godley (1987)	T-C, no selection
20	Salmon and Nicholis (1983)	Next Birthday, no selection, T-C, alternative male/female
21	Srinivsan, Christiansen, and Tortora (1996)	Last birthday, full enumeration, YMOF
22	Tarnai, Rosa, and Scott (1987)	Last birthday, YMOF
23	Zukin, Carter, and Schuman (1987)	Last birthday, no selection