

2015

Using a Calendar and Explanatory Instructions to Aid Within-household Selection in Mail Surveys

Mathew Stange

Mathematica Policy Research, Ann Arbor, MI, mstange@mathematica-mpr.com

Jolene D. Smyth

University of Nebraska-Lincoln, jsmyth2@unl.edu

Kristen Olson

University of Nebraska-Lincoln, kolson5@unl.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/sociologyfacpub>

 Part of the [Community-Based Research Commons](#), [Family, Life Course, and Society Commons](#), [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#), and the [Social Psychology and Interaction Commons](#)

Stange, Mathew; Smyth, Jolene D.; and Olson, Kristen, "Using a Calendar and Explanatory Instructions to Aid Within-household Selection in Mail Surveys" (2015). *Sociology Department, Faculty Publications*. 278.

<http://digitalcommons.unl.edu/sociologyfacpub/278>

This Article is brought to you for free and open access by the Sociology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Sociology Department, Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Field Methods* 28:1 (2015), pp. 64–78.

doi: 10.1177/1525822X15604825

Copyright © 2015 Mathew Stange, Jolene D. Smyth, and Kristen Olson. Published by SAGE Publications. Used by permission.

An earlier version of this article was presented at the American Association for Public Opinion Research (AAPOR) annual meeting in Boston, MA, May 17, 2013.



Using a Calendar and Explanatory Instructions to Aid Within-household Selection in Mail Surveys

Mathew Stange,¹ Jolene D. Smyth,² and Kristen Olson²

1 Mathematica Policy Research, Ann Arbor, MI, USA

2 Department of Sociology, University of Nebraska–Lincoln, Lincoln, NE, USA

Corresponding author — Mathew Stange, Mathematica Policy Research,
220 East Huron Street, Suite 300, Ann Arbor, MI 48103, USA;
email mstange@mathematica-mpr.com

Abstract

Although researchers can easily select probability samples of addresses using the U.S. Postal Service’s Delivery Sequence File, randomly selecting respondents within households for surveys remains challenging. Researchers often place within-household selection instructions, such as the next or last birthday methods, in survey cover letters to select respondents. Studies show that up to 30% of selections are inaccurate and samples tend not to represent key demographics. This article tests two design elements—a calendar and explanatory wording of selection instructions in cover letters—to aid and motivate households to carry out selection procedures accurately. We empirically examine these elements in two mail surveys of Nebraskans—the 2012 Nebraska Annual Social Indicators Survey and 2013 Nebraska Trees and Forest Survey.

We find that neither the calendar nor the explanatory wording adversely affected response rates but that neither improved the representativeness of the completed samples and that the calendar actually reduced selection accuracy.

Introduction

Probability samples underlie scientific survey research (Baker et al. 2013). Although households can be selected with known probabilities, obtaining a probability sample of adults within households in mail surveys remains challenging. Researchers typically use a probability or quasi-probability technique, such as asking the household member with the next birthday to complete the survey (Gaziano 2005; Yan 2009). Several of these techniques—next birthday, last birthday, and oldest/youngest adult—yield similarly unrepresentative respondent pools (e.g., underrepresenting younger, non-Whites, and less-educated individuals; see Battaglia et al. 2008; Olson et al. 2014). Additionally, up to 30% of respondents in mail surveys may be incorrectly selected (Battaglia et al. 2008; Olson and Smyth 2014; Olson et al. 2014; Schnell et al. 2007). This article reports the results of two experiments in general population surveys of Nebraska residents that test the effectiveness of methods intended to help household members correctly implement the next birthday selection procedure—a calendar and explanatory instructions in the survey cover letter.

Households require several steps to select a respondent. A household member first must open the mail and then read and understand the selection procedure in the cover letter. Next, this household informant must be motivated to implement the selection procedure and be able to follow the instructions. They then must persuade the selected individual to complete the survey. Within-household selection may fail at any stage, leading to inaccurate selections and unrepresentative samples. With no interviewer present, it is difficult to know why the process fails.

Households may incorrectly select respondents in the birthday methods because they are unable to determine which household member has the next birthday. Thus, our first experiment tests the effectiveness of providing a calendar on the cover letter to aid the household informant in keeping track of household members' birthdays (calendar experiment). We hypothesized that the cover letter with a calendar would yield a more representative completed sample and more households making accurate selections than the cover letter without a calendar.

Alternatively, households may incorrectly select respondents in the birthday methods because they may not understand what a representative sample is, why it is important, and/or how birthdays relate to obtaining a representative sample and thus be unmotivated to follow the instructions. In fact, Battaglia et al. (2008) found that 13% of those making inaccurate next birthday selections reported that the decision for who would answer for their household was random. These households did not understand random in the same way researchers do and they did not understand that the birthday procedure is intended to produce a (quasi)randomly selected respondent. In our second experiment, we compare a standard version of the within-household selection instruction to a version that explains the rationale and procedures behind the instruction in detailed lay terminology (cover letter experiment). We hypothesized that the explanatory version would yield a more representative completed sample. At the same time, however, we worried that the additional length of the letter would make the survey seem difficult, reducing response rates.

For both experiments, we examine the response rates and sample composition across the experimental treatments and compare the sample pools to 2011 American Community Survey (ACS) Nebraska benchmark data. In addition, one of the surveys included a household roster that we use to assess whether the correct household member was selected. We also examine whether the treatments or the correct respondent influenced substantive reports and item nonresponse rates.

Data and Methods

Our calendar experiment was embedded in the 2012 Nebraska Annual Social Indicators Survey (NASIS), and our cover letter wording experiment was embedded in the 2013 Nebraska Trees and Forest Survey (Forest Survey). Both used the next birthday within-household selection method.

Calendar Experiment

NASIS is a 16-page, 105-question, omnibus mail survey of 3,600 Nebraskans whose addresses were selected from the U.S. Postal Service Delivery Sequence File (DSF). The survey covered a range of topics such as roads, trees, water, social issues, household finances, and demographics. It included an invitation, reminder postcard, and two replacement mailings sent between June 14 and August 6, 2012, with 959 respondents (American Association for Public

Opinion Research Response Rate 1 [AAPOR RR1] = 26.6%). Each sampled address of NASIS was randomly assigned to one of two treatments (Figure 1). The cover letter of the “no calendar” treatment ($n = 1,800$) had the instructions:

Please have the **adult age 19 or older in your household who will have the next birthday that will take place after July 1st, 2012**, complete the questionnaire and return it in the enclosed envelope. Hearing from the person with the next birthday is very important because it ensures that we get responses from all different types of Nebraskans—men and women, the young and old, those who typically read the mail and those who do not.

The “calendar” treatment ($n = 1,800$) was otherwise identical but included a sentence stating, “We have printed the calendar at the right in case it helps you identify the right person in your household.” The calendar began on July 1, 2012.

Cover Letter Experiment

The Forest Survey was a four-page questionnaire asking questions about Nebraskans’ use of trees and forests. Three mailings (survey packet, reminder postcard, and replacement packet) were sent to 3,000 addresses sampled from the DSF between May 17 and June 17, 2013, with 908 respondents (AAPOR RR1 = 30.4%).

Each sampled address in the Forest Survey was randomly assigned to one of two cover letters. The standard treatment ($n = 1,500$) included the wording:

To make sure that our results accurately reflect the opinions of all Nebraskans, we ask that the enclosed survey be completed by *the adult (age 19 or older) in your household who will be the next to celebrate a birthday*.

The “explanatory” treatment ($n = 1,500$) included the wording:

Some people like filling out surveys and others do not, but hearing from only certain types of people can lower the quality of our results. To make sure that our results accurately reflect the opinions of all Nebraskans, we need to randomly pick someone within your household to answer the survey. Because the timing of birthdays is pretty random, we can use them to determine who should answer. Please take a moment to think about the birthdays of all the adults (age 19 or older) in your home. Who will be the next to celebrate a birthday? We ask that the enclosed survey be completed by the adult (age 19 or older) in your household who will be the next to celebrate a birthday. To ensure the quality of our results, it is very important that this is the person to complete the survey.

One limitation with this cover letter experiment is that the extra explanatory sentences both made the letter longer and added information. However, the Flesch-Kincaid Reading Grade Level decreased from 18.3 (above college) to 9.9 (between 9th and 10th grade), meaning it was easier to read (Kincaid et al. 1975). Thus, while we hypothesized that the explanatory treatment as a whole would yield a more representative sample because it explained the importance of the selection procedure, the increased length could discourage some recipients from reading the instructions.

Analysis Plan

We first examine whether including the calendar in NASIS and the explanatory language in the Forest Survey affects response rates. We then assess whether the treatments improve within-household selection by examining sample composition across key demographics: sex, race, ethnicity, age, education level, household income, and presence of children. We also compare the completed samples to 2011 ACS benchmark data.

The household roster in NASIS allows us to assess accuracy (i.e., determine if the respondent was the household member with the next birthday after July 1, 2012). An accurate selection occurs when person 1 on the roster (the respondent) has the first birthdate (day and month) after July 1 among all adult household members. We examine accuracy rates overall and by demographic characteristics.

We also examine whether the calendar and cover letter wording treatments affect item nonresponse rates or reports to substantive items. Further, in NASIS, we examine whether estimates differ between the correctly and incorrectly selected respondents.

Results

Response Rates

Response rates did not differ across the calendar (NASIS—AAPOR RR1 = 26.6% vs. 26.7%, $\chi^2(1) = .01, p = .92$) or cover letter treatments (Forest Survey—AAPOR RR1 = 30.1% vs. 30.5%, $\chi^2(1) = .06, p = .81$). In logistic regression models using information from the NASIS and Forest Survey sample frames, we found no suppression effect for the treatments and no interaction effect between information on length of residence, home ownership, children in the household, and Hispanic individuals living in the

household and the treatment groups. Having children in the household and Hispanic identity significantly predicted response in both surveys, and estimated length of residence also predicted response in NASIS ($p < .05$; analyses available on request).

Sample Composition

Table 1 displays the unweighted final sample composition for each survey overall and by treatment and the 2011 ACS benchmark values. For both experiments, the control and treatment conditions produced similar sets of respondents ($p > .10$ for all comparisons).

Almost all demographic characteristics were statistically different from the ACS benchmarks for both experiments, suggesting differential nonresponse for these subgroups overall. This differential nonresponse did not differ across the experimental treatments. All treatments overrepresented Whites, non-Hispanics, individuals 50 and older, those with a bachelor's or graduate degree, and those with income levels US\$50,000 and over ($p < .05$). For example, 88.2% of Nebraskans identify their race as white, according to the ACS, but 93.5% of the calendar sample and 91.3% of the no calendar sample identified as white. Results were similar in the Forest Survey, where 94.2% of the standard instructions sample and 93.2% of the explanatory instructions sample identified as white ($p < .05$). We suspect that the underrepresentation of Hispanics also resulted from only fielding the surveys in English.

All treatments were representative of Nebraskans with some college relative to the ACS benchmark. The no calendar treatment and both cover letter treatments produced samples that were equally representative of males and females, while the calendar treatment significantly underrepresented males and overrepresented females ($p < .05$). We also examined the sample composition of the standard and explanatory cover letter treatments at each step of the mailing process (results not shown). Additional mailings did not change sample composition.

Across the seven demographic characteristics we examined for the calendar experiment, we found that the no calendar treatment had an average absolute difference from the ACS of 8.23 percentage points, whereas the calendar treatment had an average absolute difference of 9.27 percentage points. The no calendar produced a sample that was less different from the ACS than the calendar treatment. The average absolute difference of the cover

Table 1. Sample Composition on Select Demographics by Treatment, NASIS, and Forest Survey (Unweighted Percentages).^a

	NASIS			Forest Survey			ACS estimates	
	Total sample	Calendar	No calendar	χ^2/T value (p value)	Total sample	Standard wording		Explanatory wording
Sex								
Male	45.82*	44.42*	47.22	.7382 (.390)	46.1*	46.0	46.3	.0087 (.926)
Female	54.18*	55.58*	52.78		53.9*	54.0	53.7	
Race								
White	92.79*	94.17*	91.42*	2.6287 (.105)	93.7*	94.2*	93.2*	.3443 (.557)
Nonwhite	7.21*	5.83*	8.58*		6.3*	5.8*	6.8*	
Ethnicity								
Hispanic	1.62*	1.30*	1.94*	.5903 (.442)	1.8*	1.6*	2.0*	.2732 (.601)
Not Hispanic	98.38*	98.70*	98.06*		98.2*	98.4*	98.0*	
Age								
Mean	57.45	57.53	57.38	.1385 (.890)	57.0	56.5	57.6	-.9725 (.3311)
19-34	11.36*	11.21*	11.52*	1.8794 (.598)	13.4*	14.9*	12.0*	2.2875 (.515)
35-49	18.18*	18.75*	17.61*		15.6*	16.4*	15.1*	
50-64	36.06*	34.05*	38.04*		33.6*	32.8*	34.4*	
65+	34.42*	35.99*	32.83*		37.2*	35.9*	38.6*	
Education								
≤ High school	22.92*	23.03*	22.81*	.5356 (.765)	24.0*	22.4*	25.6*	1.2726 (.529)
Some college	33.55	32.46	34.65		34.7	35.1	34.3	
BA+	43.53*	44.52*	42.54*		41.3*	42.5*	40.1*	

(continued)

Table 1. (continued)

	NASIS			Forest Survey					
	Total sample	Calendar	No calendar	χ^2/T value (p value)	Total sample	Standard wording	Explanatory wording	χ^2/T value (p value)	ACS estimates
Family income									
Under US\$50,000	43.11*	42.55*	43.65*	.1697 (.919)	43.2*	43.6*	42.8*	2.9802 (.225)	51.3
US\$50,000–99,999	37.62*	38.30*	36.95		37.8*	35.4	40.3*		33.0
Over US\$100,000	19.28*	19.15	19.40		19.0*	21.0*	16.9		15.7
Have children									
Yes	18.14*	17.36*	18.92*	.3903 (.532)	–	–	–	–	32.1
No	81.86*	82.64*	81.08*		–	–	–		67.9

a. The χ^2 p value indicates whether the distributions are independent of the experimental treatment. The asterisks (*) denote a significant difference from the American Community Survey (ACS) benchmark ($p < .05$) based on a t test. The results for both surveys were consistent between the unweighted and weighted data (household size) with the exception of the 35–49 age group not differing from the ACS benchmarks for both Nebraska Annual Social Indicators Survey (NASIS) treatments among the weighted data.

letter treatments in the Forest Survey from the six ACS benchmarks was almost identical (7.94 percentage points—standard wording; 7.96 percentage points—explanatory wording).

Selection Accuracy

Among all households, 69.51% of respondents were selected accurately, with the no significant difference in accuracy between the no calendar (72.42%) and calendar (66.67%) treatments ($\chi^2 = 3.29, p = .07$). We examine accuracy for households with two or more adults because the default accurate selections of one-adult households inflate accuracy rates. Overall, 58.04% of NASIS respondents from households with at least two adults were accurately selected. In the calendar treatment, 53.16% of respondents were accurately selected, compared to 62.91% in the no calendar treatment ($\chi^2 = 5.89; p < .05$). Thus, counter to our hypothesis, the calendar resulted in less accurate selections. Both accuracy rates are similar to previous research (Olson et al. 2014).

We also examined accuracy rates across treatments for key demographics. As seen in Table 2, there are few statistically significant differences in selection accuracy across the experimental treatments for most of the demographic characteristics, but the pattern of results is almost uniform in the direction of the no calendar treatment yielding higher accuracy rates than the calendar treatment.

Substantive Results and Item Nonresponse

For both NASIS and the Forest Survey, we examined whether the experimental treatments influenced substantive reports and item nonresponse rates. We examined attitudes about roads, natural resources, interpersonal trust, and political and social issues in NASIS, and the questions about trees and forests in the Forest Survey. Substantive estimates and item nonresponse rates did not differ across the calendar treatments in the NASIS survey, nor between the standard and explanatory cover letters in the Forest Survey (results available on request).

We also compared substantive estimates on the same questions from the entire completed sample to estimates from only those who were correctly selected in NASIS and found few differences ($p < .05$). For example, 92.0% of the correctly selected respondents reported being satisfied with snow and ice removal by the Nebraska Department of Roads compared to 91.9% of

Table 2. Accuracy Rates by Cover Letter Treatment across Demographics, Two Plus Adult Households, and Nebraska Annual Social Indicators Survey.

	Calendar	No calendar	Calendar versus no calendar <i>T</i> value
Sex			
Male	56.9	62.5	-1.10
Female	51.8	62.8	-1.98*
Race			
White	52.8	62.4	-2.30*
Nonwhite	69.2	66.7	0.55
Age			
19-34	40.5	51.3	-0.93
35-49	60.0	61.7	-0.19
50-64	50.0	63.9	-2.11*
65+	59.0	66.7	-1.05
Education			
≤High school	63.4	66.1	-0.31
Some college	47.3	57.6	-1.44
BA+	53.4	64.9	-1.89
Family income			
Under US\$50,000	49.4	62.2	-1.71
US\$50,000-99,999	54.9	64.3	-1.50
Over US\$100,000	44.6	55.4	-1.23
Have children			
Yes	50.0	60.0	-1.02
No	53.8	63.5	-2.19*

The asterisks (*) denote accuracy rate between the calendar and no calendar treatments is statistically different ($p < .05$) within each demographic category.

all respondents, and 84.1% of correctly selected respondents and 83% of all respondents support wind energy in Nebraska. Item nonresponse rates also did not differ. Thus, in NASIS, a high inaccurate within-household selection rate had little impact on substantive estimates among the items examined here. We recognize, however, that these analyses are limited because the accurately selected respondents are not the ideal sample that would result had the entire sample been accurately selected according to the next birthday selection technique. Unfortunately, we lack true values or a rich sample frame that would provide the necessary data for a better evaluation of the effect of inaccurate within household selections on substantive estimates and item nonresponse rates.

Discussion

Neither the calendar nor the explanatory language improved the representativity of the samples, and the calendar actually lowered the selection accuracy rate. Neither treatment affected response rates nor made the samples more unrepresentative. Overall, the sample compositions in these studies are similar to previous research (e.g., Battaglia et al. 2008; Hicks and Cantor 2012; Olson et al. 2014) and the findings for address-based samples in general (e.g., Link et al. 2008). In particular, all of the treatments yielded samples that underrepresented non-Whites, younger individuals, those with lower education, and those earning less than US\$50,000. Thus, weighting is required to make the sample reflect the target population on these characteristics. Our finding of accuracy rates (NASIS) is also consistent with previous research (Olson et al. 2014).

Why did the experimental treatments fail to improve within-household selection? Our experimental treatments were developed under the assumption that respondents were unable to follow the instructions and/or did not understand the rationale behind our need for a representative sample. Battaglia et al. (2008) found that few respondents expressed difficulty following within-household selection instructions. We wondered whether this finding was driven by the social undesirability of admitting such difficulties. Our findings suggest that our treatments were not strong enough (i.e., not helpful, educational, or motivational enough), the increased cover letter length of the explanatory treatment counteracted the effect of presenting the rationale for selection, or that, as Battaglia et al. (2008) suggest, the within-household selection process breaks down for other reasons.

One reason for inaccurate within-household selections is that the selected household member does not want to participate (Battaglia et al. 2008). In this study, neither survey had an engaging topic, and the length of NASIS (105 items) may deter participation. Although it is an unlikely factor here, asking sensitive questions may also be a reason that correct respondents choose not to participate. As Battaglia et al. (2008) suggested, adding incentives may have helped increase the accuracy rate by increasing the selected household member's motivation to participate. With rising nonresponse generally (Massey and Tourangeau 2013; Rookey et al. 2012), the selected household member's reluctance to respond is likely part of the larger nonresponse challenge facing surveys.

Future research should examine alternative within-household selection instructions and other survey design features that aim to motivate adherence to within-household selection procedures and to garner participation from those groups usually underrepresented in mail surveys. Additionally, the selection procedures in NASIS and the Forest Survey used slightly different wording. This research would benefit from cognitive interviews focusing on how sampled persons understand and complete different within-household selection instructions and where the selection process breaks down (Willis 2005).

As previous research has shown that the framing of survey requests can influence response rates (Fazekas et al. 2013; Redline et al. 2004; Tourangeau and Ye 2009), cover letters with different appeals for participation may also influence within-household selection accuracy. Future research should also examine how survey topic, sensitivity, length, and incentives influence selection accuracy and the makeup of completed samples in mail surveys.

Our findings suggest that other aspects of survey design should be investigated to improve selection quality. The calendar aid on the cover letter did not increase the accuracy of selecting respondents, and wording cover letters in lay terminology to communicate the rationale and procedures behind within-household selection instructions was of little utility in terms of aiding within-household selection. This research suggests that practitioners can keep their cover letters short and uncluttered from visual elements, such as calendars, and maintain the same respondent pool.

Acknowledgments — We thank the Bureau of Sociological Research at the University of Nebraska- Lincoln and the Nebraska Forest Service for funding the data collection. Mathew Stange contributed to this research while a graduate student in the Survey Research and Methodology Program at the University of Nebraska-Lincoln. The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article and received no financial support for its research, authorship, or publication.

References

- Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. A. Gile, and R. Tourangeau. 2013. Report of the AAPOR task force on nonprobability sampling. <http://aapor.org/AM/Template.cfm?Section=Reports1&Template=/CM/ContentDisplay.cfm&ContentID=5963> (accessed March 9, 2013).
- Battaglia, M. P., M. W. Link, M. R. Frankel, L. Osborn, and A. H. Mokdad. 2008. An evaluation of respondent selection methods for household mail surveys. *Public Opinion Quarterly* 72:459–69.
- Fazekas, Z., M. T. Wall, and A. Krouwel. 2013. Is it what you say, or how you say it? An experimental analysis of the effects of invitation wording for online panel surveys. *International Journal of Public Opinion Research* 26:235–44.
- Gaziano, C. 2005. Comparative analysis of within-household respondent selection techniques. *Public Opinion Quarterly* 69:124–57.
- Hicks, W., and D. Cantor. 2012. Evaluating methods to select a respondent for a general population mail survey. Paper presented at the annual meeting of the American Association for Public Opinion Research, Orlando, FL, May 18.
- Kincaid, J. P., R. P. Fishburne, R. L. Rogers, and B. S. Chissom. 1975. Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel (Branch Report 8-75). Millington, TN: Chief of Naval Training.
- Link, M. W., M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad. 2008. A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys. *Public Opinion Quarterly* 72:6–27.
- Massey, D. S., and R. Tourangeau. 2013. New challenges to social measurement. *The ANNALS of the American Academy of Political and Social Science* 645: 6–22.
- Olson, K., and J. D. Smyth. 2014. Accuracy of within-household selection in web and mail surveys of the general population. *Field Methods* 26:56–69.
- Olson, K., M. Stange, and J. D. Smyth. 2014. Assessing within-household selection methods in household mail surveys. *Public Opinion Quarterly* 78:656–78.
- Redline, C., J. Oliver, and R. Fecso. 2004. The effect of cover letter appeals and visual design on response rates in a government mail survey. Paper presented at the annual meeting of the American Association for Public Opinion Research, Phoenix, AZ, May 11.
- Rookey, B. D., L. Le, M. Littlejohn, and D. A. Dillman. 2012. Understanding the resilience of mail-back survey methods: An analysis of 20 years of change in response rates to national park surveys. *Social Science Research* 41:1404–14.

- Schnell, R., S. Ziniel, and E. Coutts. 2007. Inaccuracy of birthday respondent selection methods in mail and telephone surveys. Presentation at the conference of the European Survey Research Association, Prague, Czech Republic, October 7.
- Tourangeau, R., and C. Ye. 2009. The framing of the survey request and panel attrition. *Public Opinion Quarterly* 73:338–48.
- Willis, G. B. 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage.
- Yan, T. 2009. A meta-analysis of within household respondent selection methods. *Proceedings of the Survey Research Methods Section of the American Statistical Association* 14:6134–47.