Title: Factors Associated with Participation in the Collection of Saliva Samples by Mail in a Survey of Older Adults

Authors: Jennifer Dykema[1]*, Kerryann DiLoreto[1], Kenneth D. Croes[1], Jeremy Beach[2], Dana Garbarski[1,3,4]

[1]University of Wisconsin Survey Center, University of Wisconsin-Madison
[2]USDA National Agricultural Statistics Service
[3]Department of Sociology, University of Wisconsin-Madison
[4]Department of Population Health Sciences, University of Wisconsin-Madison

***** Corresponding author:
Dana Garbarski
Department of Sociology
University of Wisconsin-Madison
8128 Sewell Social Science Building
1180 Observatory Drive
Madison, WI 53706
Phone: 608-262-1913
E-mail: dgarbars@ssc.wisc.edu

ABSTRACT

Biological measures are increasingly collected along with survey responses in household surveys. While the promise of collecting these measures is rich, there are many barriers to gathering data accurately and reliably, which may lead to a winnowing of the sample to a more select group. Research is needed to determine what factors are associated with respondents' willingness to provide biological measures in a survey and what barriers impede successful collection.

We examine factors associated with compliance to a request to provide a salivary DNA sample as part of a separate data collection effort for the highly successful Wisconsin Longitudinal Study. Over 50% of eligible respondents returned a self-administered saliva sample through the mail. Using survey responses from prior waves and paradata from the survey data collection process, we examine factors associated with returning the saliva sample including respondents' socio-demographic characteristics, cognitive ability, health, religiosity, and prior participation in the survey.

BACKGROUND

Increasingly biological measures are collected along with survey responses in household surveys that measure health. However, while the promise of collecting these measures in surveys is rich, there are many barriers to gathering representative and high-quality data including: nonresponse bias that may result if different groups, such as those in poorer health, participate at lower levels than those in better health; ensuring respondents follow directions and understand the nature of the task; technological hurdles; and increased costs (Beebe 2007; Weinstein, Vaupel, and Wachter 2008; Weinstein and Willis 2000). Research is needed to identify and isolate variables that are associated with successful participation from those that reduce participation in order to develop methods to increase participation, and to understand the potential for nonresponse bias.

One critical problem in indentifying variables that are correlated with successful participation in survey-based requests is that information on nonparticipants is typically unknown (Groves and Couper 1998). Longitudinal survey designs, however, offer the unique ability to use data from previous waves of data collection to predict nonresponse to later requests for participation (Lepkowski and Couper 2002), and that is the strategy we employ here. Using survey data from the highly successful Wisconsin Longitudinal Study (WLS) (Sewell, Hauser, Springer, and Hauser 2003), for which data were first collected from a random sample of all graduating high school seniors in Wisconsin in 1957, we examine variables that are associated with a request of sample members to provide a saliva sample for the purposes of extracting DNA in 2007, approximately 50 years later.

To isolate respondent and survey characteristics with the potential to predict when respondents might be more likely to provide a biological sample like saliva using mail as the mode of administration, we reviewed the literature on patterns of participation in household surveys that use probability sampling (Groves and Couper 1998), including the few (published) studies in which biological measures have been collected along with survey responses in a household survey (e.g., Jaszczak, Lundeen, and Smith 2009; Sakshaug, Couper, and Ofstedal 2010). We also examined the literature that explores factors associated with collecting biological measures in clinical settings using volunteer or quota samples (e.g., Henderson et al. 2008). Our review highlights the importance of variables that describe characteristics of the respondent, such as socio-demographic characteristics, cognitive ability, health status, and religiosity, and paradata, or survey process variables that describe characteristics of the respondent's prior and current participation in the survey.

METHODS

Data Collection

Data are provided by the Wisconsin Longitudinal Study (WLS), a longitudinal study of 10,317 randomly selected respondents who graduated from Wisconsin high schools in 1957. Since 1957, follow-up interviews have been conducted with graduates or their parents in 1964, 1975, 1992-1993 (phone with mail follow-up), and 2003-2005 (phone with mail follow-up). The survey covers a variety of topics with a focus on educational plans, occupational aspirations, social influences, and more recently, physical and mental health. A high survey response rate characterizes all waves of survey data collection (Hauser 2005).

In 2007, WLS sample members were selected to participate in the collection of salivary DNA, including graduates who were nonparticipants from earlier waves of data collection. Sample members were only excluded from this phase of data collection if they were included in the pilot study described above, deceased prior to or died during the field period, were not able to

be located during tracing, were unable to complete the task, or had requested no further contact prior to the saliva collection effort. Saliva collection kits were sent to all sample members who did not refuse during the advance calling and for whom we had an address. Kits consisted of a cover letter explaining that salivary DNA was being collected to facilitate research into genetically-based illnesses and disorders such as Alzheimer's Disease, cancer, and depression; two consent forms (one to be signed and returned, the other for the sample member's records); a sealable plastic bag with a patch of cloth to absorb potential sample leakage; an Oragene-brand saliva self-collection kit (type OG-250); a self-addressed, first-class-stamped padded return envelope; and a sheet that provided written and illustrated instructions on how to use the saliva self-collection kit. Saliva kits were returned by 4,356 of the 8,112 graduates included in the production phase of the salivary DNA collection effort for an overall participation rate of 53.7 percent. Nonparticipation primarily resulted from sample members actively refusing to participate during the advance calling or failing to return kits sent to them.

Our analytic sample of the effects of self-reported and administrative survey variables uses reports only from the graduates in the salivary DNA collection who also completed the 2003-2005 WLS telephone and mail surveys, and were among the 80 percent randomly selected to receive the module on religion in the telephone survey (N = 4,587). Saliva kits were returned by 3,053 these respondents for a participation rate of 66.6 percent. (Note that the participation rate for this subset of respondents was higher than for the full sample described above (66.6 percent vs. 53.7 percent). We selected this sample because it allowed us to examine a wide variety of measures, some of which were only available in the religion module or in the mail survey follow-up to the 2003-2005 telephone survey. We note, however, that while levels of some variables for which we have measures for all 8,000ish graduates are different from our analytic sample of 4,587, the associations reported here are stable. This study was approved by the Institutional Review Board at the University of Wisconsin-Madison.

PRELIMINARY RESULTS

Table 1 presents the results from logistic regression models in which the outcome is coded with a value of "1" if the sample member returned a saliva kit and "0" otherwise. Cell entries provide the odds ratios (Exp[B]) of returning a saliva kit on various combinations of the survey predictors calculated using STATA (Version 10). The first column in Table 1 presents the results from bivariate logistic regression models predicting the effect of each independent variable on the odds of returning a saliva kit. This allows us to examine the effect of each survey predictor on its own and before controlling for other indicators.

Reduced Model 1 shows the effects on saliva kit response for the variables measuring socio-demographic characteristics. All of these variables have large and statistically significant effects on participation. The odds of responding are 11% lower among women than among men (OR, 0.89; CI, 0.79-1.02) and 22% lower for graduates who live alone versus those who live with at least one other person (OR, 0.78; CI, 0.66-0.92). In contrast, having a college education increases the odds of participating by 47% (OR, 1.09; CI, 1.06-1.12). Furthermore, with each additional organization reported, the odds of participating increase by 9% (OR, 1.09; CI, 1.05-1.14).

In Reduced Model 2, we add the measure for cognitive ability – adolescent IQ score – to the equation, and consistent with its effect in the bivariate model, a one unit increase in IQ is associated with a 1% increase in the odds of returning a saliva kit (OR, 1.01; CI, 1.00-1.01). The addition of IQ to the model also attenuates the effect of education.

In the bivariate models, the indicators for health status are all significant and in the predicted direction of respondents in poorer health being less likely to provide a saliva sample. In Reduced Model 3, we examine the effects of the measures of health status while controlling for socio-demographic characteristics and cognitive ability. We find significant effects remain such that respondents who lack private health insurance have lower odds of participation as well as those with low health care utilization as indicated by not having been to the doctor in the year prior to the 2004 survey. Interestingly, the effects of our composite measures of physical health – self-reported health and the HUI – are accounted for by socio-demographic characteristics, cognitive ability, and the other measures of health status.

We examine indicators for religiosity in Reduced Model 4, finding that respondents with low and high religious attendance have lower odds of participating in comparison with respondents who report more moderate religious attendance even after controlling for other respondent characteristics. However, while respondents who "agree" or express neutrality with the statement that the "Bible is God's word" have lower odds of participating in the bivariate model, these effects are not significant when we control for other factors.

Reduced model 4 also controls for all respondent characteristics simultaneously. In this model, women are significantly less likely to participate than men, those who participate in social organizations and higher IQ scores are more likely to participate, respondents with low health care utilization in the previous year were less likely to participate, and respondents with low and high religious attendance were less likely to participate.

In both the bivariate and multivariate model in which we control for the survey process characteristics simultaneously, we find strong effects of the indicators for patterns of participation and resistance in the 2004 survey interview. For all variables examined the trend is that respondents who demonstrate reluctance to a prior request continue to be more likely to not participate in subsequent tasks. The final column in Table 1 presents a full model in which all variables are estimated simultaneously. Few respondent characteristics continue to be significantly associated with participation net of the effect of the survey process variables.

DISCUSSION

In contrast to findings about participation in household surveys for which some evidence indicates women are more likely to participate, we find female graduates are less likely to provide saliva samples through the mail. We are still investigating possible mechanisms, but expect that this could be due to women being less likely to want to provide DNA or women being less desirous or able to spit. Our findings for household composition mirror those of Sakshaug, Couper, and Ofstedal (2009) who reported that respondents in the HRS who lived with another eligible respondent were more likely to consent to provide biological measures. These findings are consistent with studies of response rates in household surveys more generally (see review in Couper and Groves 1996); respondents who live alone are hypothesized to be more socially isolated and consequently less likely to engage in a predominantly social exchange like participating in survey.

Respondents' cognitive ability, as indicated by their IQ scores assessed during high school, has important relationships to participation. Respondents with lower cognitive abilities are much less likely to participate. An avenue for future investigation would be to explore the relationship between more current measures of cognitive functioning and respondents' willingness to participate in the collection of biomarkers like saliva.

It is interesting to note that respondents with no physical or dental exam were less likely to participate in the saliva collection after controlling for other health and sociodemographic

characteristics, cognitive ability, and religiosity (Reduced model 4). Of major concern to researchers collecting biological measures is the potential for nonresponse bias based on the respondents' health. If respondents in poorer health are less likely to provide biomeasures than respondents in better health, results from studies that combine self-reports with biological data could be biased. However, it is unclear whether those with no exams are healthier than those with exams or whether health care utilization is a proxy for socioeconomic status. We plan to explore this effect further.

The prior survey experience variables we examine strongly indicate that respondents who express reluctance to participate in a prior wave of data collection – such as by being harder to attempt to interview or refusing at some point in the calling effort – participate in lower levels to a subsequent request to provide a biological sample. These findings are consistent with others. For example, Lepkowski and Couper (2002) developed a model of factors to use in predicting retention in the second waves of America's Changing Lives Survey and the National Election Surveys. They showed that interviewer assessments of respondents' survey experience at Wave 1 – such as comments that the respondent "enjoyed the interview," was "too busy," or demonstrated "reluctant behavior"— predicted cooperation at Wave 2. These findings complement those of Groves and Couper (1998) and Campanelli et al. (1997) who reported that negative comments by respondents during an initial interviewer contact in a cross-sectional study predicted subsequent refusal.

We are still at the beginning stages of understanding how respondents make decisions about participating in survey research in which biological measures are gathered and what factors influence when that decision-making process will lead to consent. Taking advantage of a longitudinal survey design in which a wealth of information exists on respondents, our research culls a limited number of respondent characteristics – socio-demographic characteristics, cognitive ability, health-status measures, religiosity, and prior participation – and examines their effect on a request to provide a salivary DNA sample. Ultimately, we believe a fuller understanding of how the characteristics of respondents affect their current and prior (if relevant) survey participation and how they evaluate the characteristics of the task and ultimately make a decision to participate is needed. This research will not only need to isolate which respondent characteristics are important but how they interact with task characteristics in order to leverage responses to optimize participation.

REFERENCES

Beebe, Timothy J. 2007. "Challenges to collecting survey-based biomarker and genetic data." *Survey Research* 38:1-4.

Campanelli, Pamela, Patrick Sturgis, and Susan Purdon. 1997. "Can you hear me knocking: An investigation into the impact of interviewers on survey response rates." The Survey Methods Centre at SCPR.

Couper, Mick P. and Robert M. Groves. 1996. "Household-level determinants of survey nonresponse." Pp. 63-79 in *Advances in Survey Research*, edited by M. Braverman and J. K. Slater. San Francisco: Jossey-Bass.

Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley and Sons.

Hauser, R. M. 2005. "Survey Response in the Long Run: The Wisconsin Longitudinal Study." *Field Methods* 17:3-29.

Henderson, Gail, Joanne Garrett, Jada Bussey-Jones, Mairead Eastin Moloney, Connie Blumenthal, and Giselle Corbie-Smith. 2008. "Great expectations: views of genetic research participants regarding current and future genetic studies." *Genetics in Medicine* 10:193-200.

Jaszczak, Angela, Katie Lundeen, and Stephen Smith. 2009. "Using Nonmedically Trained Interviewers to Collect Biomeasures in a National In-Home Survey." *Field Methods* 21:26-48.Lepkowski, James M. and Mick P. Couper. 2002. "Nonresponse in the Second Wave of Longitudinal Household Surveys." Pp. 259-272 in *Survey Nonresponse*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York: John Wiley & Sons.

Sakshaug, Joseph W., Mick P. Couper, and Mary Beth Ofstedal. 2010. "Characteristics of Physical Measurement Consent in a Population-Based Survey of Older Adults." *Medical Care* 48:64-71.

Sewell, William H., Robert M. Hauser, Kristen W. Springer, and Taissa S. Hauser. 2003. "As We Age: The Wisconsin Longitudinal Study, 1957-2001." in *Research in Social Mobility*, edited by K. Leicht. London: Elsevier.

Weinstein, Maxine, James Vaupel, and Kenneth Wachter. 2008. *Biosocial Surveys. Committee on Advances in Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys.* Washington, D.C.: National Academies Press.

Weinstein, Maxine and Robert J. Willis. 2000. "Stretching Social Surveys to Include Bioindicators: Possibilities for the Health and Retirement Study, Experience from the Taiwan Study of the Elderly." Pp. 250-275 in *Cell and Surveys: Should Biological Measures be Included in Social Science research?*, edited by C. E. Finch, J. W. Vaupel, and K. Kinsella. National Research Council Commission on Behavioral and Social Sciences and Education: Washington D.C.: National Academy Press.

Table 1. Bivariate, reduced, and full models for WLS survey variables as predictors of participation in the saliva collection among WLS graduates who participated in the 2003-2005 phone and mail follow-up surveys, including 80% random sample, including social participation as a demographic characteristic

			Respondent Characteristics											
	Bivariate Models		Reduced Model 1		Reduced Model 2		Reduced Model 3		Reduced Model 4		Survey Process Characteristics		Full Model	
Survey predictors	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Respondent Characteristics														
Socio-demographic characteristics														
Female [vs. male]	0.81**	0.72-0.92	0.89^{+}	0.79-1.02	0.88^{+}	0.77-1.00	0.82**	0.72-0.94	0.78**	0.68-0.90			0.77**	0.66-0.90
Educational attainment														
[High school graduate]	1.00		1.00		1.00		1.00		1.00				1.00	
Some college	1.06	0.88-1.28	0.97	0.80-1.18	0.93	0.76-1.13	0.90	0.73-1.10	0.91	0.74-1.12			0.84	0.67-1.06
College degree or more	1.61***	1.39-1.89	1.47***	1.25-1.73	1.32**	1.10-1.57	1.17^{+}	0.97-1.41	1.17	0.96-1.43			1.11	0.90-1.37
Lives alone [vs. with another]	0.71***	0.60-0.83	0.78**	0.66-0.92	0.78**	0.66-0.92	0.84^{+}	0.71-1.01	0.86	0.72-1.03			0.89	0.72-1.08
Social participation	1.13***	1.08-1.17	1.09***	1.05-1.14	1.09***	1.05-1.14	1.09***	1.04-1.13	1.06**	1.02-1.11			1.05^{+}	1.00-1.10
Cognitive ability														
Adolescent IQ	1.01***	1.01-1.02			1.01**	1.00-1.01	1.01**	1.00-1.01	1.01**	1.00-1.01			1.00	0.99-1.01
Health status														
Self-reported health														
[Excellent]	1.00						1.00		1.00				1.00	
Very good	0.85*	0.72-1.00					0.89	0.75-1.06	0.89	0.75-1.07			0.86	0.71-1.04
Good	0.72***	0.61-0.86					0.82^{+}	0.67-1.01	0.82^{+}	0.67-1.01			0.82^{+}	0.65-1.03
Fair	0.79^{+}	0.61-1.03					1.05	0.76-1.44	1.06	0.77-1.47			1.26	0.88-1.82
Poor	0.45**	0.29-0.71					0.62^{+}	0.36-1.05	0.61^{+}	0.36-1.06			0.72	0.39-1.33
HUI	1.61**	1.20-2.17					1.05	0.73-1.53	1.09	0.75-1.60			1.05	0.69-1.60
Chronic conditions														
[0 conditions]	1.00						1.00		1.00				1.00	
1 to 2 conditions	1.12	0.96-1.29					1.11	0.94-1.31	1.13	0.95-1.34			1.12	0.93-1.35
3 conditions or more	0.79*	0.65-0.97					0.83	0.65-1.06	0.84	0.65-1.07			0.82	0.62-1.08
No physical exam [vs. had one]	0.76***	0.66-0.88					0.77**	0.65-0.90	0.80**	0.68-0.94			0.72***	0.60-0.87
No dental exam [vs. had one]	0.70***	0.60-0.81					0.84^{+}	0.71-1.00	0.84*	0.71-1.00			0.95	0.78-1.15
Current smoker [vs. not]	0.86	0.71-1.04					0.96	0.78-1.17	0.96	0.78-1.18			1.02	0.81-1.28
No private insurance [vs. any]	0.73***	0.62-0.87					0.81*	0.67-0.98	0.83^{+}	0.68-1.01			0.90	0.73-1.12

Table 1 is continued on the following page.

Table 1. Continued

	Bivariate Models		Respondent Characteristics: Reduced Model 1		Reduced Model 2		Reduced Model 3		Reduced Model 4		Survey Process Characteristics		Full Model	
Survey predictors	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Religiosity														
Religious attendance														
Never	0.70***	0.58-0.84							0.72**	0.58-0.90			0.81+	0.63-1.03
Less than weekly	0.81**	0.70-0.94							0.81**	0.69-0.95			0.86	0.72-1.03
[About weekly]	1.00								1.00				1.00	
More than weekly	0.68**	0.54-0.85							0.69**	0.53-0.88			0.71*	0.54-0.93
"Bible is God's word"														
Strongly agree	0.67***	0.56-0.82							0.83+	0.66-1.04			0.82	0.64-1.05
Agree	0.85+	0.70-1.02							1.05	0.85-1.31			1.12	0.88-1.43
Neither agree nor disagree	0.76**	0.64-0.90							0.94	0.78-1.14			0.92	0.75-1.14
[Disagree/strongly disagree]	1.00								1.00				1.00	
Survey Process Characteristics														
Incomplete participation [vs. complete]	0.48***	0.42-0.55									0.65***	0.55-0.75	0.73***	0.61-0.87
Medicare questionnaire not returned [vs. returned]	0.12***	0.11-0.14									0.14***	0.12-0.17	0.14***	0.11-0.16
Call attempts in 2004 (number)	0.93***	0.92-0.94									0.97***	0.95-0.98	0.97***	0.95-0.98
Refused in a 2004 call [vs. not]	0.32***	0.25-0.40									0.49***	0.38-0.65	0.52***	0.38-0.71
Likelihood ratio chi-squared			71.44***		80.42***		116.06***		134.56***		856.10***		782.23***	
Log Likelihood			-2735.04		-2730.55		-2556.51		-2456.51		-2495.09		-2132.68	
Degrees of freedom			5		6		17		23		4		27	
N			4374		4374		4176		4059		4587		4059	

+ p < .10; * p < .05; ** p < .01; *** p < .001