ARTICLES

# Do In-person Interviews Reduce Bias in a Mixed-Mode Survey of Persons with Disabilities?

Eric Grau[1]

[1] Mathematica Policy Research, Princeton, NJ 08540

## Survey Practice

## Background

For people with certain disabilities, completing an interview by telephone or mail may be difficult or impossible, making in-person interviews necessary. However, in-person interviews are generally more expensive than telephone or mail and may be cost prohibitive for large samples. The National Beneficiary Survey (NBS), sponsored by the Social Security Administration (SSA), is a multiwave survey of persons who receive Supplemental Security Income (SSI) or Social Security Disability Income (SSDI). The survey was conducted four times between 2004 and 2010 by Mathematica Policy Research using computer-assisted telephone interviewing (CATI) with computer-assisted personal interviewing (CAPI) follow-up for telephone nonrespondents.[1] We achieved significant cost savings by limiting field operations to only those beneficiaries who could not be reached or interviewed by telephone. The mixed mode design also furnished higher response rates than would be achievable with a telephone-only design.

The purpose of this analysis was to assess whether survey estimates would be affected if CAPI was scaled back or eliminated in the NBS. Specifically, we compared respondent attributes and responses for three groups of overlapping respondents:

1. **CATI-only (no CAPI) respondents.** Beneficiaries who completed an interview by telephone

2. **Two-month CAPI respondents.** All CATI-only respondents plus beneficiaries who completed a CAPI interview within two months after the start of field operations[2]

---

[1] A new iteration of the NBS is currently being conducted, where the sample for the first round will be selected at the end of 2014. This study uses data from the prior NBS.

[2] If field operations were deemed necessary for a particular sample member, the case was assigned to the field at some point after the start of data collection. We recorded the responses within 60 days of the assignment to the field for each case. The date of the assignment to the field could occur at any time after the first three months of data collection.

3. **Full-field respondents.** Beneficiaries who completed an interview by the end of the full-field period (all CATI and CAPI respondents)

We looked for differences in attributes between these groups that were larger than the two standard deviations.

## Methodology

The NBS as conducted in 2004, 2005, 2006, and 2010 was a 45-minute survey that gathered information on the health, insurance, employment, income, and demographic characteristics of SSI and SSDI beneficiaries. Interviews were attempted first by telephone, followed by face-to-face interviews with people for whom no telephone interview could be completed. The survey instrument was identical in both modes. (See Thornton et al. 2004 for details on the NBS.)

The NBS was composed of two samples: a nationally representative sample of SSDI and SSI beneficiaries (the beneficiary sample) and a sample of participants in SSA's Ticket to Work program (the participant sample). This study was based on data collected from the beneficiary sample during the third wave of the NBS (conducted in 2006).[3] At round 3, interviews were conducted over an eight-month data collection period, starting in February and ending in September.

CATI occurred throughout the entire data collection period. CAPI did not begin until May, three months into data collection, and lasted for five months. The sample included 2508 completed interviews with a weighted response rate of 81%.[4] This sample represents more than 10 million SSI and SSDI beneficiaries.

For this analysis, we used the 2006 survey data from the representative beneficiary sample to represent the full-field respondents. For the remaining two sets of respondents, we reassigned sample members as respondents or nonrespondents using details about the date and outcome of every locating and interview attempt. For the CATI-only condition, we excluded all cases that were completed by CAPI. For the two-month CAPI condition, we excluded all cases that were completed by CAPI more than two months after the initial field assignment.

---

3  Data from the fourth round were not yet available when this research was conducted.

4  This response rate is the weighted count of sample members for whom a completed interview was obtained or who were determined to be ineligible divided by the weighted sample count of all sample members (# of completed interviews + # partially completed + # of ineligibles/# of cases in the sample). It can be determined by taking the product of the weighted location rate and the weighted cooperation rate, also known as the weighted completion rate among located sample members. This response rate is equivalent to the American Association of Public Opinion Research (AAPOR) standard response rate calculation: RRAAPOR=# of completed interviews/# of cases in the sample – estimated # of ineligible cases. By including the ineligible cases in the numerator and denominator, we avoid using this estimation stage and the response rate computation is more clearly explicated.

**Table 1**  Prevalence of disability type in frame compared with sampled respondents, by level of CAPI effort.

| Respondent's disability | Frame (%) | Full-field | | | 2-Month CAPI | | | CATI-Only | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Est. (%) | S.E. (%) | Deviation from frame* | Est. (%) | S.E. (%) | Deviation from frame* | Est. (%) | S.E. (%) | Deviation from frame* |
| Deaf | 0.87 | 0.90 | 0.27 | 0.1 s.e.s | 0.98 | 0.31 | 0.4 s.e.s | 0.85 | 0.35 | –0.1 s.e.s |
| Blind | 2.39 | 2.80 | 0.41 | 1.0 s.e.s | 2.91 | 0.44 | 1.3 s.e.s | 2.87 | 0.50 | 1.2 s.e.s |
| Psychiatric disability | 30.4 | 28.9 | 1.28 | –1.2 s.e.s | 28.9 | 1.32 | –1.2 s.e.s | 29.3 | 1.67 | –0.9 s.e.s |
| Intellectual disability | 13.5 | 13.9 | 0.91 | 0.4 s.e.s | 14.0 | 0.95 | 0.6 s.e.s | 13.6 | 0.93 | 0.1 s.e.s |
| Other physical disability | 52.8 | 53.5 | 1.44 | 0.5 s.e.s | 53.2 | 1.46 | 0.2 s.e.s | 53.4 | 1.66 | 0.4 s.e.s |

*All deviations from the frame are expressed in terms of the number of full-field standard errors.

The weights for the original survey data (full-field condition) were adjusted for nonresponse to reduce the potential for nonresponse bias and poststratified to frame counts for selected variables. For this analysis, we also created weights for the CATI-only and two-month-CAPI conditions, adjusted for nonresponse, and poststratified to frame counts.

To assess potential bias associated with scaling back or eliminating CAPI efforts, we compared attributes of respondents for these three groups. We also compared selected survey estimates between the full-field effort sample and the two CAPI-truncated groups of respondents. In another study of NBS data, Sloan, Wright, and Barrett (2011) identified a variety of variables that could potentially give different estimates between CAPI and CATI,[5] which we used for the comparisons in this study.

## Results

### *Comparison with Frame*

Comparisons with the frame are limited to a five-level variable identifying the beneficiary's disability type.[6] When comparing weighted proportions of beneficiaries with sample members using the original weighted totals, the proportion with each disability differed by increasing amounts as field efforts decreased. However, as Table 1 indicates, the nonresponse adjustments to the weights appeared to successfully adjust the samples to the frame.

---

5  These authors identified variables for which item nonresponse, providing socially desirable responses, and acquiescence (a tendency to agree) differed between CAPI and CATI.

6  Weights were poststratified to age and gender, so comparisons with the frame for these variables would not be informative. Information about the sample members' race in the sample frame was too messy to be useful.

**Table 2**  Comparisons of estimates for demographic variables between CATI and CAPI, by level of CAPI effort.

| | Full-field | | 2-Month CAPI* | | | CATI-Only* | | |
|---|---|---|---|---|---|---|---|---|
| | Est. (%) | S.E. (%) | Est. (%) | S.E. (%) | Deviation from full-field est. | Est. (%) | S.E. (%) | Deviation from full-field est. |
| Father's Education: High School Graduate | 34.0 | 1.74 | 33.7 | 1.71 | −0.20 s.e.s | 34.2 | 2.05 | 0.12 s.e.s |
| Father's Education: 4-Year Degree or Higher | 10.8 | 1.22 | 10.8 | 1.25 | 0.02 s.e.s | 12.9 | 1.67 | 1.76 s.e.s |
| Education Level: High School Graduate | 36.1 | 1.40 | 35.5 | 1.45 | −0.37 s.e.s | 36.4 | 1.79 | 0.28 s.e.s |
| Education Level: 4-Year Degree or Higher | 7.0 | 0.74 | 7.2 | 0.78 | 0.19 s.e.s | 7.9 | 0.93 | 1.17 s.e.s |
| General Health: Very Good | 6.8 | 0.77 | 7.0 | 0.78 | 0.24 s.e.s | 7.3 | 0.87 | 0.64 s.e.s |
| Health Insurance: Medicare | 68.0 | 1.26 | 67.9 | 1.27 | −0.14 s.e.s | 67.8 | 1.52 | 0.20 s.e.s |
| Health Insurance: Private Health Insurance | 20.7 | 1.36 | 20.7 | 1.37 | −0.04 s.e.s | 23.7 | 1.67 | 2.20 s.e.s |

*For all deviations from the full-field estimate, the estimated standard error is the full-field standard error.

## Assessment of Survey Estimates

When comparing survey estimates, the demographic variables we examined included race, father's and mother's education, education level of beneficiary, general health of beneficiary, beneficiary's type of health insurance, and household income. A comparison of most of these estimates across the full-field, two-month CAPI, and CATI-only estimates is given in Table 2. Because many of these variables have multiple levels, we present the level with the maximum deviation from the full-field estimate for the two-month CAPI and the CATI-only estimates. For the sake of comparability, the same standard error is used for all estimates (the full-field standard error).

To preserve space, we limit attention to those variables with the highest deviations from the original estimates in terms of the number of full-field standard errors. As is apparent from this table, the estimates derived after two months of CAPI effort do not differ substantially from estimates obtained under full-field conditions.

The CATI-only estimates differ more, but in most cases, the deviation is fewer than two full-field standard errors. The exception is private health insurance, in which the CATI-only estimate is 2.20 standard errors higher than the original estimate. For median household income, which is not presented here, the CATI-only estimate is 1.41 five-month-field standard errors higher than the full-field estimate. These deviations might reflect a true difference in CAPI vs. CATI respondent characteristics.

We also examined variables related to work goals:

**Table 3**   Comparisons of estimates for variables indicating knowledge of TTW programs, by level of CAPI effort.

| Survey variable | Full-field | | 2-Month CAPI | | | CATI-Only | | |
|---|---|---|---|---|---|---|---|---|
| | Est. (%) | S.E. (%) | Est. (%) | S.E. (%) | Deviation from full-field est.* | Est. (%) | S.E. (%) | Deviation from full-field est.* |
| Heard of impairment-related work expenses exclusion | 9.8 | 0.99 | 10.1 | 1.02 | 0.30 s.e.s | 11.3 | 1.14 | 1.60 s.e.s |
| Heard of expedited reinstatement | 14.9 | 1.10 | 15.1 | 1.18 | 0.17 s.e.s | 17.2 | 1.38 | 2.08 s.e.s |
| Heard of work incentive and planning assistance programs | 11.0 | 0.92 | 11.3 | 0.95 | 0.37 s.e.s | 13.8 | 1.19 | 3.04 s.e.s |
| Heard of TTW | 25.5 | 1.24 | 25.8 | 1.27 | 0.28 s.e.s | 27.0 | 1.47 | 1.20 s.e.s |

*Deviations from the full-field estimates are presented in terms of the number of standard errors of the given estimate, assuming the full-field estimate is constant.

- Goals included moving up[7]
- See yourself working for pay next year[8]
- See yourself working for pay in the next five years[8:1]

The estimate for the first variable was defined by the proportion with an affirmative response; the other two were on a Likert scale. As with the demographic variables, the estimates obtained in the two-month CAPI condition did not differ appreciably from those obtained under the full-field condition. The CATI-only and full-field estimates varied more, but no differences exceeded one standard error.

Finally, we looked at estimates for variables indicating knowledge about the TTW program (see Table 3). Again, the two-month CAPI estimates do not vary much from those achieved from the full-field period. However, "yes" responses were more likely for all of the variables examined in the CATI-only condition compared with full-field condition. This indicates either that (1) we are excluding a part of the population that is less informed about the programs if no CAPI efforts are made or (2) a bias due to acquiescence (a tendency to agree) that is found in CATI is at least somewhat alleviated by including field efforts and, by excluding fielded cases, this bias is more serious. These results agree with those from Sloan, Wright, and Barrett (2011), who found that "yes" responses were more likely with CATI respondents when compared with CAPI respondents.

---

7  The survey question reads, "Do your personal goals include moving up in a job or learning new job skills?"

8  The second and third bullet points are from the same survey question. The survey questions reads, "Please tell me how much you agree with the following statements. Would you say you strongly agree, agree, disagree, or strongly disagree?: You see yourself continuing to work for pay in the next year/next five years."

## Conclusion

Comparisons of sampled respondents at various levels of CAPI efforts revealed that the nonresponse adjustments to the weights effectively accounted for any under- or overrepresentation overrepresentation of various disability groups. Additionally, although the effects were not large, several differences between CATI-only survey estimates and the estimates that were obtained with some CAPI follow-up. These differences were largest when comparing estimates obtained from CATI-only to those obtained under the full-field period, suggesting that even a truncated CAPI follow-up could be effective in reducing potential bias in survey estimates.

A limitation of this study is that it is not possible to determine the source of the differences in estimates obtained under the three conditions. These differences might reflect the experiences of the beneficiaries participating under varying degrees of locating and interviewing effort. However, it is also possible that the differences we observed reflect differences in the quality of the data obtained (Holbrooke, Green, and Krosnick 2003; Krosnick et al. 2002; Voogt, Saris, and Saris 2005).

Some researchers have found increased measurement error with a mixed-mode approach when compared with a single-mode approach. This could be due to a mode switch bias (Sakshaug, Yan, and Tourangeau 2011) or poorer quality data from difficult-to-contact or reluctant respondents (Fricker and Tourangeau 2010). Our findings, combined with other studies, suggest that bias could be exacerbated if CAPI follow-up efforts are eliminated. On the other hand, relying too much on field efforts to increase response rates could increase measurement error. Future research should focus on identifying the ideal length of such follow-ups.

## Funding & Acknowledgements

# REFERENCES

Fricker, S., and R. Tourangeau. 2010. "Examining the Relationship between Nonresponse Propensity and Data Quality in Two National Household Surveys." *Public Opinion Quarterly* 74 (5): 934–55.

Holbrooke, A.L., M.C. Green, and J.A. Krosnick. 2003. "Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Bias." *Public Opinion Quarterly* 67 (1): 79–125.

Krosnick, J.A., A.L. Holbrooke, M.K. Barent, R.T. Carson, W.M. Hanemann, R.J. Kopp, R.C. Mitchell, et al. 2002. "The Impact of 'No Opinion' Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66 (3): 371–403.

Sakshaug, J.W., T. Yan, and R. Tourangeau. 2011. "Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items." *Public Opinion Quarterly* 74 (5): 907–33.

Sloan, M., D. Wright, and K. Barrett. 2011. "Data Comparability in a Mixed Mode Telephone and Face to Face Interviewer of Persons with Disabilities." Washington, DC: Mathematica Policy Research.

Thornton, C., G. Livermore, D. Stapleton, J. Kregel, T. Silva, B. O'Day, T. Fraker, W. Revell, H. Schroeder, and M. Edwards. 2004. "Evaluation of the Ticket to Work Program Initial Evaluation Report. Document No. PP06-122." Washington, DC: Mathematica Policy Research.

Voogt, R., J. Saris, and W.E. Saris. 2005. "Mixed Mode Designs: Finding the Balance between Nonresponse Bias and Mode Effects." *Journal of Official Statistics* 21: 367–87.