

ESTIMATING RESPONSE BIAS IN AN ADULT EDUCATION SURVEY

J. Michael Brick and David Morganstein, Westat, Inc.

David Morganstein, Westat, Inc., 1650 Research Boulevard, Rockville, Maryland 20850

KEY WORDS: Reinterview, telephone survey, measurement errors

1. Introduction

Estimates from surveys are subject to both variable and systematic nonsampling errors. Variable nonsampling errors, or response variance, are those that might vary across repeated surveys administered to the same sample, assuming that the conditions of the interview could be controlled so that the surveys were independent. For example, the same respondent might report annual income differently when asked in repetitions of the same survey because the method used by the respondent to estimate income might vary (records might be used, recall might be used, or the value might be estimated using different schemes). These circumstances would lead to variable errors for estimates of income.

Systematic nonsampling errors, on the other hand, are those that have a particular direction. For example, if respondents tend to omit certain types of income, say interest income from savings, then the estimated income would be expected to be lower than the true income. In repetitions of the same survey, the estimated income would always be less than the true income. These types of systematic errors are called response bias. Survey estimates can be subject to both response variance and response bias.

Measuring response bias is typically very difficult. This study examines an intensive reinterview as a particular approach to estimating response bias. Other approaches for measuring response bias, the reasons for using an intensive reinterview, and the goals of the study are presented in the next section, after describing the source of the data. Section 3 outlines the methods used to collect the intensive reinterview data. Section 4 gives the estimates of the response bias and possible explanations of the findings. The last section summarizes the highlights of the study and the applicability of the method to other surveys. A more complete analysis of this study is given in Brick *et al.* (1996a).

2. Study Design

The source of the data for this analysis is a special methodological study undertaken as part of the 1995 National Household Education Survey (NHES:95). The NHES is an ongoing data collection

system of the National Center for Education Statistics (NCES) conducted by Westat, Inc. designed to address a wide range of education-related issues. It is a telephone survey of the noninstitutionalized civilian population of the US that has been conducted in 1991, 1993, 1995 and 1996.

In the NHES, households are selected for the survey using random digit dialing (RDD) methods and data are collected using computer-assisted telephone interviewing (CATI) procedures. Approximately 60,000 households are screened for each administration. The NHES survey for a given year typically consists of a Screener, which collects household composition and demographic data, and extended interviews on two substantive components addressing education-related topics. This study is based on the Adult Education (AE) component of the NHES:95. It was designed to estimate the percentage of adults participating in adult education activities and the characteristics of both participants and nonparticipants.

As noted above, it is often difficult to measure response bias. A frequently used method of doing this is to compare the results of the survey against answers from a more definitive source, such as an administrative record file. However, record checks have their own limitations, *e.g.*, record checks can only be used if records exist on the survey topic and those records can be accessed. Brick *et al.* (1994) found that, even for the well-defined topic of teacher certification, records were not complete and accurate and could not be matched to the survey respondents without error.

Another way of measuring response bias is through the use of reinterviews. Reinterviews are ordinarily undertaken to measure response variance rather than response bias. However, sometimes a process called reconciliation is used in reinterviewing to measure bias. If the original and reinterview responses are different, then the respondent is asked to reconcile the differences and the resulting response is called the reconciled response. The reconciliation is often conducted by a supervisor rather than a regular interviewer, assuming this will make the reconciled response less subject to error. Under these assumptions, the difference between the original and reconciled response has been used to estimate response bias (Forsman and Schreiner 1991).

Reconciliation has been used in earlier NHES studies to estimate response bias (Brick *et al.* 1996b) and a reinterview was also conducted for the NHES:95

AE interviews (Brick *et al.* 1996c). However, there is little evidence that reinterviews, even the reconciled responses, actually measure response bias. As a result of this, the NHES:95 reinterview study was designed to estimate the response variance.

It should not be too surprising that this approach does not provide reliable estimates of response bias. The methods used in the reinterviews, such as selecting interviewers from the original interviewer pool, asking the questions in much the same way as asked in the original interview, not informing the interviewer or the respondent of the answers from the original interview, and waiting at least 14 days between interviews so that the respondent will not remember the details of the original interview are all designed to support the measurement of response variance rather than response bias.

The intensive reinterview was designed to be an alternative method of estimating response bias that did not suffer from some of the shortcomings of the record check or regular reinterviews. The intensive reinterview method was pioneered by Belson (1986) who focused on difficult or sensitive topics primarily in opinion and marketing research.

The intensive reinterview differed from the regular reinterview in a number of ways. The interviewers were not selected from the regular pool of telephone interviewers, but were persons with previous experience in interviewing using less structured methods. The interviewers were trained to use a protocol and to conduct the reinterviews in a conversational mode, using probes and other devices to trigger recall and comprehension. The reinterview was focused on a few topics and ample time was allowed for discussing these few points. The respondents were encouraged to voice their opinions and understanding of the topics. Furthermore, attempts were made to engage the respondents in the interview by explicitly asking for their advice on ways to improve the interview. The hope was that these methods would lead to more complete and accurate reporting in the intensive reinterview.

Although there were four major research objectives of the study, only two of them are discussed in this paper. The first goal of the study was to examine the potential bias in the estimates of the percentage of adults who participated in adult education activities. The bias could be due to either underreporting participation or overreporting activities that took place outside of the time frame of the survey (i.e., the past 12 months prior to the original interview). Respondents might underreport participation either because they might not recall a qualifying activity during the 12 months before they

were interviewed or because they might not comprehend the range of activities that were included as adult education. These types of underreporting would lead to downward bias in participation rates. However, an upward bias could occur if respondents "telescoped" some activities. Telescoping is reporting activities that took place outside of the time frame of the survey as having taken place within that time frame. As described below, underreporting was expected to be minimal in all types of adult education (the six types of activities were: ESL, adult basic education/GED preparation classes, credential programs, apprenticeships, work-related courses, and personal development courses), except work-related and personal development courses. As a result, the intensive reinterview focused on these two types of participation in order to assess the bias in the overall participation rate.

The second goal was to obtain more accurate estimates of participation in work-related and personal development courses, separately. As a result of the differences between the estimates of participation from earlier AE surveys and cognitive laboratory work, it was suspected that work-related courses and personal development courses were susceptible to underreporting. One of the major concerns for reporting these types of courses is that respondents might not comprehend the full range of activities that are included as work-related and personal development courses. These types of comprehension problems could combine with recall problems and result in underreporting of work-related and personal development courses.

3. Intensive Reinterview Method

In an attempt to more closely determine the respondent's actual status or opinions, the intensive interview was more of a directed conversation between the respondent and the interviewer rather than a formally scripted interview. Respondents were reminded of their answers in the original survey and asked if the answers were still true for them. They were asked to recall other details related to their responses. Interviewers were fully knowledgeable about the original answers given by the respondent. Tactics similar to those used in cognitive laboratory work, such as asking open-ended questions and using probes to encourage the respondent to elaborate on his or her answer, were used. The goal was to obtain more detailed and accurate information by understanding the respondent's perspective and the reasons for his or her answers.

The intensive reinterview was a new undertaking and presented several challenges. For example, the interviewers who conducted the NHES interviews were thoroughly trained to read the questions verbatim and

to avoid affective behavior that might influence the respondent. Adopting the conversational and unstructured interviewing method called for in the intensive reinterview required major changes in their behavior. The interviewers were also called upon to implement some methods used in cognitive research, but they were not previously trained in these methods. The respondents also faced a challenge because the intensive reinterview differed significantly from the type of interview they had already done. They were called upon to give reasons for their responses and provide details rather than choose among response alternatives.

To address these challenges, a protocol and data collection methods were developed especially for this study. The full details of the protocol development, the methods used to select and train the interviewers, the sampling of respondents to the original interview, and data collection methods are provided by Brick *et al.* (1996a). These issues are very important but space limitations prevent giving but a few important features of the final sample in this paper.

Although the goal of the study was to develop estimates of bias, only a very limited sample size could be fielded. Because of the small sample sizes, it was decided that the typical design-based estimates gathered from the original interview would be subject to very large sampling errors and relationships would be obscured by these sampling errors. Thus, the results from this relatively small sample were analyzed assuming the observations were from independent, identically distributed random variables and sampling weights were not used. The sample was randomly selected from both participants and nonparticipants who completed the AE extended interview. In order for a case to be eligible for the study, certain conditions had to be met. For example, the original interview had to be conducted in English.

A sample of 230 adults was selected to meet specific targets by participation status and educational level. Of the 206 sampled adults who completed the intensive reinterviews (90 percent), 115 were nonparticipants in the original interview and 91 were participants.

4. Findings

4.1 Bias in Participation Rate Estimates

The first goal of the study was to estimate the response bias associated with estimates of the rate of participation in AE from the NHES:95. Adults were classified as participants in AE if they had participated in one or more of six different types of adult education activities during the past 12 months. Based on the responses in the NHES:95, 40 percent of all adults had

participated in one or more of these activities in the last 12 months (Kim *et al.* 1995).

As noted earlier, responding to the items about participation in work-related and personal development courses was identified as being problematic during the design phase of the NHES:95. After the survey was completed, the results from the reinterview confirmed that these two types of participation were much more likely to be reported inconsistently than any of the other types (Brick *et al.* 1996c). These results support the decision to restrict this study to an in-depth examination of reporting work-related and personal development courses.

Adults were classified as nonparticipants in the original interview if they said they had not taken any courses in the last 12 months. Of the 115 nonparticipants who responded to the intensive reinterview, 41 percent indicated that they had taken one or more work-related or personal development courses (in the intensive reinterview, respondents were not asked about other types of courses). Since none of the participants sampled for the intensive reinterview denied having taken courses, the response bias in the overall participation rate is one-directional and substantial.

Assuming the responding nonparticipants in the Bias Study are a simple random sample of all adults classified as nonparticipants in the NHES:95 (the analysis is thus unweighted), the bias in the NHES:95 estimate is 24 percent. The bias is estimated by multiplying the percent of all adults who were nonparticipants as reported in the NHES:95 by the percent of the nonparticipants who reported participating in the intensive reinterview. In general, the estimated bias is

$$\hat{b}(p_0) = p_0 \hat{y}^p - (100 - p_0) \hat{y}^{np} \quad (1)$$

where p_0 is the estimate of the percentage of adults classified as participants in the initial interview, \hat{y}^p is the estimate of the proportion of participants in the initial interview who reported not participating in the intensive reinterview, and \hat{y}^{np} is the estimate of the proportion of nonparticipants in the initial interview who reported participating in the intensive reinterview.

In this case, the last term of the estimated bias is zero ($\hat{y}^p = 0$), because no initial participants said they had not taken any courses during the intensive reinterview. If the bias in the estimated percentage of adults who participated in AE is 24 percent, then the bias-corrected estimate is that 64 percent of adults participated in AE in 1995. This is substantially larger than the 40 percent reported in the NHES:95. Both the bias and the percentage participating from the NHES:95 are

subject to sampling error and because of the sample size the sampling error of the bias is very large relative to that for the estimate from the NHES:95. Taking advantage of the fact that $\hat{y}^P=0$, the estimated bias can be written as

$$\hat{b}(p_0) = (p_0 - 100)\hat{y}^{np}.$$

Thus, the estimated variance of the bias is a product of random variables and the approximate variance for a product of independent random variables is (Hansen *et al.* 1953)

$$Var(\hat{b}) = (\hat{y}^{np})^2 Var(p_0) + (p_0 - 100)^2 Var(\hat{y}^{np}). \quad (2)$$

Substituting the estimated values into (2) and taking the square root, the standard error of the estimated bias of 24 percent is 2.7 percent. Thus, a 95 percent confidence interval for the estimated bias is from 19 percent to 29 percent, and for the percent of adults participating, the confidence interval is from 59 to 69 percent.

Considering the nonparticipating respondents to the intensive reinterview a simple random sample of all nonparticipants in the original survey is a key assumption in estimating the response bias. Usually, the sampling procedures would ensure that this assumption holds, but the sampling methods described earlier were primarily concerned with making sure the sample sizes for specific groups were large enough to provide some nonparticipants with various characteristics. In addition, the small sample size does not allow for broad generalizations. Thus, these estimates are exploratory and should not be used to make bias corrections to the NHES:95 estimates.

To evaluate the reasonability of this assumption, the characteristics of the responding nonparticipants from the intensive reinterview were compared to the characteristics of all nonparticipants from the NHES:95.

While the age and sex distributions are similar, the educational attainment distributions are different, with a much larger percentage of intensive reinterview nonparticipants having less than a high school education. This is a consequence of the sampling methods used for the intensive reinterview. This difference highlights the fact that the point estimates and confidence intervals from the study are subject to specification errors that cannot be measured. Despite this shortcoming, the findings clearly show that a relatively large fraction of the adults classified as nonparticipants in the original survey did identify AE activities in the intensive reinterview.

In addition to nonparticipants, participants in the original survey who were sampled were asked if there were any courses they had not reported in the initial

interview by using the same types of probes described above. All of the work-related and personal development courses participants reported in the NHES:95 were verified as being within the eligible 12 month time period. In addition, about one-third of the sampled participants identified additional courses that were not reported in the original interview.

The reporting of additional work-related and personal development courses by adults classified as participants in the original survey is a further indication that the respondents may have had a more restrictive understanding of the scope of activities than was intended. Drawing on the work of Schwarz (1995), one interpretation of this finding is that respondents might have reacted to the context of the original interview in determining what was an eligible activity. The NHES:95 interview began by asking about more formal types of participation and some respondents may have created a response paradigm before the questions about the less formal activities were asked. In the intensive reinterview the context was different because the only types of courses discussed were work-related and personal development courses.

4.2 Bias in Work-Related and Personal Development Participation Estimates

Overall, about half the adults who named additional courses reported work-related courses and half reported personal development courses. Participants were more likely to add personal development courses and the nonparticipants were more likely to add work-related courses, but these differences are not statistically significant.

Based on the NHES:95 responses, 21 percent of adults were estimated to have participated in work-related courses during the previous 12 months and 20 percent were estimated to have participated in personal development courses (Kim *et al.* 1995). The extent of the bias in these estimates can be estimated using equation (1). The bias for the work-related participation rate is

$$\hat{b}(p_{0,wr}) = (1 - p_{0,wr})\hat{y}_{wr}^{np} \quad (3)$$

where $p_{0,wr}$ is the estimate of the percentage of adults classified as work-related participants in the initial interview, and \hat{y}_{wr}^{np} is the estimate of the proportion of adults who did not report participating in work-related activities in the initial interview but reported participating in the intensive reinterview. Because we are now dealing with participation in a particular type of adult education, the value of \hat{y}_{wr}^{np} has two components: those classified as nonparticipants who reported taking

work-related courses in the intensive reinterview, and participants in the initial survey who reported taking work-related courses for the first time in the intensive reinterview. In the intensive reinterview, 23 percent of the nonparticipants reported taking work-related courses and 8 percent of the participants reported taking work-related courses for the first time.

Substituting the values into (3), the estimated bias for the percent of adults participating in work-related courses is 16 percent. The standard error can be computed using (2), where \hat{y}_{wr}^{np} is treated as a sum of the two components described above. Using this approach, the standard error of the estimated bias is 3 percent and the 95 percent confidence interval for the estimated bias is from 10 to 22 percent.

The same calculations can be performed for personal development courses to compute the estimated bias and its standard error. The estimated biases are summarized in Table 1 below.

These estimates show that the underreporting bias is approximately the same for both work-related and personal development courses. While these estimates are subject to the same caveats as the overall estimates of participation rates, they also have an interesting implication because of the difference in the wording of the questions about the two types of participation. The introduction to the question about work-related courses does not include specific examples, but does mention courses taken at work, taken somewhere else but related to work or career, and courses taken to obtain a license or certificate related to work or career. On the other hand, the introduction to the question about personal development specifically mentions courses including "arts and crafts, sports or recreation, first aid or childbirth, Bible study, or any other course."

One way of interpreting the equal biases for the two types of participation is that adding examples does not improve the quality of reporting in this situation. This interpretation is consistent with the hypothesis that respondents develop a response paradigm in the original interview that includes only more formal courses. If this is true, then the addition of the specific examples may not cause respondents to change their paradigm and a different approach might be needed to address the exclusion of less formal courses in reporting.

5. Summary

The estimated bias in the overall participation rate of adults was 24 percent and the bias-corrected estimate is that 64 percent of adults participated in AE in 1995. This is substantially larger than the 40

percent reported in the NHES:95. The underreporting of participation for work-related and personal development courses was also substantial and of about the same magnitude for each of these types of participation.

A reason for the underreporting interpretation may be related to how respondents react to the context of the interview. Some respondents may have created a response paradigm that restricted their answers to more formal courses before the questions about the work-related and personal development courses were asked. Despite the fact that more examples were used for the personal development courses than for the work-related course question, the estimated biases were approximately the same for the two types of participation. This suggests that simply adding examples to the wordings of the questions may not improve the quality of reporting and that other approaches to the underreporting problem may be needed.

If the adults have developed a response paradigm that focuses on formal types of participation (*i.e.*, traditional schooling or formal programs), then a relatively drastic intervention may be needed to modify this behavior. For example, a modification in which the respondents are asked to actively cooperate in changing the focus, for example by giving examples of less formal courses, might be more effective.

The intensive reinterview methodology appears to have good potential as a method for detecting biases, especially if more traditional methods like record check studies are not feasible. The alternative approach of using reconciled reinterviews, on the other hand, has not proven to be successful for estimating bias. However, from an operational perspective, it is important to understand that this method is more costly than a regular reinterview. As a result, this method should be used primarily when there is an indication of reporting errors and the estimates subject to the biases are important to the survey objectives.

References

- Belson, W.A. (1986). *Validity in survey research*. Cambridge, England: University Press.
- Brick, J.M., Cahalan, M., Gray, L., Severynse, J., and Stowe, P. (1994). "A study of selected sampling errors in the 1991 Survey of Recent College Graduates," Technical Report. U.S. Department of Education, Office of Educational Research and Improvement, NCES 95-640.
- Brick, J.M., Kim, K., Nolin, M.J., and Collins, M. (1996a). "Estimation of response bias in the NHES:95 Adult Education Survey." Working Paper ,96-13. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

- Brick, J.M., Rizzo, L., and Wernimont, J. (1996b). "The 1993 National Household Education Survey: Reinterview results for the School Readiness and School Safety and Discipline components." Technical report. U.S. Department of Education, Office of Educational Research and Improvement, NCES. (in progress).
- Brick, J.M., Wernimont, J., and Montes, M. (1996c). "The 1995 National Household Education Survey: Reinterview results for the Adult Education component." Working Paper, 96-14. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Forsman, G., and Schreiner, I. (1991). "The design and analysis of reinterview: An overview." In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (eds.), *Measurement error in surveys*, 279-302. New York: John Wiley & Sons.
- Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). *Survey Methods and Theory, Vol. I: Methods and Applications*. New York: John Wiley and Sons.
- Kim, K., Collins, M., Stowe, P., and Chandler, K. (1995). "Forty percent of adults participate in adult education activities: 1994-95." Statistics in Brief. U.S. Department of Education, Office of Educational Research and Improvement, NCES 95-823.
- Schwarz, N. (1995) "What respondents learn from questionnaires: the survey interview and the logic of conversation." *International Statistical Review*, 62, 2, 153-177.

Table 1. Estimates of bias in overall, work-related, and personal development participation rates.

Type of participation	NHES:95 estimate	Estimated bias	Sampling error	Bias-corrected estimate
Overall	40%	24%	2.7%	64%
Work-related	21%	16%	3.0%	37%
Personal development	20%	14%	3.0%	34%

Source: U.S. Department of Education, National Center for Education Statistics, Bias Study of the National Household Education Survey, 1995.