REINTERVIEW PROGRAM FOR THE 1991 NATIONAL HOUSEHOLD EDUCATION SURVEY

J. Michael Brick, Westat, Inc. and Jerry West, National Center for Education Statistics
J. Michael Brick, 1650 Research Blvd., Rockville, MD 20850

Key Words: reinterview, telephone survey, measurement error

1. Introduction

In the spring of 1991, the first, full-scale National Household Education Survey (NHES:91) was conducted for the National Center for Education Statistics (NCES) by Westat, Inc. The NHES:91 was a national, random digit dial (RDD) telephone survey of about 60,000 households designed to estimate characteristics of the educational experiences of young children and adults. The survey was conducted using computer-assisted telephone interviewing (CATI).

A reinterview program was included in the NHES:91 in order to examine the impact of measurement errors on estimates of the characteristics of children's early educational experience. A sample of parents who completed the original telephone interview concerning their 3- to 8-year-old child was recontacted and asked to respond to a subset of the questions asked in the original interview. Responses to the original interview and the reinterview are the source of the statistics on measurement errors presented in this paper.

The primary objectives of the NHES:91 reinterview program were to identify survey items that were not reliable, to quantify the magnitude of the response variance for groups of items, and to provide feedback for improving the design of future NHES surveys. Since the interviewing was a closely monitored CATI survey conducted from centralized telephone centers, there was no need to use the reinterviews to prevent the falsification of interviews.

The reinterview program had a goal of completing 500 reinterviews of the nearly 14,000 interviews of parents of 3- to 8-year-olds. Only a subset of the full set of items included in the original interview was included in the reinterview to reduce the burden on the respondents and to control the cost of the reinterview. The items selected for the reinterview were ones that were important substantively and were not highly dependent on the circumstances surrounding the time of the interview.

Sometimes, respondents give answers during reinterviews that differ from the original interview responses. These differences, or discrepancies, could arise as a result of several different causes, and not all discrepancies are errors. In the NHES:91 reinterview program, the interviewers attempted to categorize the discrepancies into four categories:

- Circumstances related to the child changed between the time of the first and the second interview; both answers, although different, may be correct,
- The original response was recorded (interviewer error) or reported (respondent error) incorrectly,
- The reinterview response was recorded or reported incorrectly,
- Both the original and reinterview responses were recorded or reported incorrectly.

Because the reinterview was also computer-assisted, the responses to the original interview and reinterview were automatically compared and displayed for the interviewer at the end of the reinterview, not after each item was asked. If the reinterview response was incorrect, the reconciled value was entered by the interviewer at this time. This paper compares the responses to the original interview and the reconciled reinterview, discusses the reliability of the respondent's answers, and discusses the reasons for errors.

2. Design of the NHES:91 and Reinterview Program

The NHES:91 was a RDD telephone survey conducted with persons in a sample of telephone households in the 50 States and the District of Columbia between February and April of 1991. A reinterview program was included for the Early Childhood Education (ECE) component of the survey which interviewed the parents of children from 3 to 8 years old.

The survey covered the noninstitutional civilian population of 3- to 8-year-olds in the United States. Since only persons in telephone households were surveyed, the estimates were adjusted so that the totals were consistent with the total number of persons in all households. Household screening interviews were completed with 60,314 households, including 13,257 households with at least one 3- to 8-year-old in the household. A total of 13,892 ECE interviews were completed for the survey. The completion rate for the ECE interview, or the percent of interviews conducted, was 94 percent. The overall response rate for the ECE interview, the

product of the screening response rate and the ECE completion rate was 76 percent. Further details on the sample design and results of the ECE component of NHES:91 are given in Brick et al. (1991).

A random sample of completed ECE interviews was selected for reinterview. Not all ECE interviews were eligible for reinterview. The case was eligible if it met all of the following conditions: 1) the original interview was completed at least 6 weeks after the start of data collection, 2) the case was not included in a special longitudinal sample selected for other purposes; 3) no more than one case was sampled for reinterview per household; and 4) all other extended interviews scheduled for the household were complete.

A sample of 604 cases was selected for reinterviews, and 534 of these were completed, for a response rate of 88 percent. About half of the nonresponse was due to persons who refused to participate in the reinterview.

The reinterview was originally designed to be conducted 14 days after the completion of the original ECE interview. However, toward the end of the data collection period, the threshold was reduced in an attempt to complete additional reinterviews. Table 1 below shows the number of days between the original interview and the reinterview.

The reinterview was conducted using the same CATI system used in the original interview. Identical items were read to the parent/guardian who completed the original interview. After all of the items for the reinterview were asked, a reconciliation of the original and reinterview responses was done automatically by the computer. Up until the end of the interview and the appearance of the reconciliation screens, the interviewer was unaware of the responses given by the respondent to the original interview.

As mentioned in the introduction, discrepancies in responses were grouped into four categories. A total of 1,618 discrepancies occurred during the 534 reinterviews, or about 3 per interview. The number of items varied significantly from interview to interview due to skip patterns. The reasons for the discrepancies, as reported by the respondent, were distributed across the four categories as shown in Table 2.

Table 1. Number of days between completion of original interview and reinterview

Number of Days	Frequency[1]	Percent
less than 10	21	4%
10 to 13	36	7
14	126	24
15 to 20	261	49
21 to 27	77	15
28 to 41	9	2
Total	530	100

Note, that for the 207 discrepancies where the child's situation changed, the reinterview answer and the original answer were not the same, but this was not an indication of an error. However, since differences between the reconciled reinterview responses and the original interview responses were used to indicate an error in the analysis that follows, these cases somewhat inflate the estimates of the measurement errors for the NHES:91. The data could be re-analyzed without counting these as errors, but our preliminary analyses of these data indicate that the differences are minor in nearly all cases.

One of the interesting methodological features of the NHES:91 reinterview was the fact that the results of the original interview were unknown to the interviewers until the completion of the reinterview. If we assume that the interviewers conducting the reinterviews were of equal quality to the original interviewers (a reasonable assumption since the interviewers worked both

Table 2. Number of discrepancies between original and reinterview responses, by reason

Reason for discrepancies	Number Reported	Percent
Child's Situation Changed	207	13%
Original Interview Answer was Incorrect	1034	64
Reinterview Answer was Incorrect	320	20
Both Interview Answers were Incorrect	41	3
Didn't Know How to Explain Discrepancy	4	<1
Some Other Explanation for Change	12	1
	1,618	100

surveys) and that the chance of making an error was equal in both the original and reinterview, we would expect the percent of errors made in the reinterview to approximate the percent of errors made in the original interview. This is clearly not the case; in the reconciliation process, about 3 times as many errors were associated with the original interview as with the reinterview.

The finding of excess errors in the original interviews is a typical result for reinterviews. It has led many designers of reinterview programs to designate a large part of the reinterview sample to be conducted without reconciliation, at least partially due to the assumption that the interviewers might either perform differently or use the original values to skew the results to improve their (reinterview) performance. These results from a situation in which the interviewer does not have any opportunity to glance at the original responses suggest that the role of reconciliation in reinterviews may need to be reconsidered.

It is possible that just knowing that a reconciliation process will follow makes interviewers more careful and less prone to error. However, the alternative hypothesis that the respondent is the source of this inequality in the assignment of the errors is at least as feasible. In other words, respondents may wish to be internally consistent with their latest responses, making it more comfortable to report that the original interview is in error. If this hypothesis is correct, there are important implications for the design and analysis of reinterview data.

3. Methods Used for the Analysis of the Reinterview

The statistics computed to examine various aspects of reporting in the original ECE survey and its reinterview are the set of statistics developed for assessing response reliability based upon reinterview data. The statistics include the gross difference rate (G), the net difference rate (E), and the index of inconsistency (I).

The gross difference rate measures the proportion of cases that had different responses in the two administrations of the interview. The net difference rate measures the bias after the offsetting misclassifications have been taken into account. The index of inconsistency is a less familiar statistic. In some circumstances, the index can be used to measure the proportion of the total variability that arises due to random response error. Descriptions of these statistics and their interpretation are given by Biemer, et al. (1991).

These statistics are computed based on the number of sample cases reported as having the characteristic in the original survey and in the reinterview. No weights are used in the analysis.

4. Findings

The sample size, the gross difference rate, the net difference rate and the index of inconsistency for the items collected in the reinterview are shown in Table 3. The sample size varies from item to item because of skip patterns in the interviews. The table presents the items that are common to both the preprimary (children not yet in first grade) and the primary (children in first grade or beyond) interviews, followed by items found only in the preprimary interviews, and finally those only in the primary interview.

Overall Assessment

Before going into the details of the statistics presented, some comments on the overall nature of the response variability are in order. The net difference rate is probably the most direct measure of bias of the estimates among the three reinterview statistics presented. For over 80 percent of the items given in Table 3, the net difference rate is less than 5 percent. Only 4 items had net differences rates greater than 10 percent, and these four items were restricted to subgroups of children with small sample sizes (between 30 and 60 cases).

The gross difference rate, which includes the non-offsetting errors, follows much the same pattern. About three-fourths of the items have gross difference rates that are less than 10 percent. Several items have gross difference rates in excess of 15 percent and many of these were for items for subgroups of the population.

The index of inconsistency is not as easily generalized, since the size of this statistic is related to the size of the estimate (the denominator of the index is a function of the percent of persons with the characteristic). For items which are present in between 20 and 80 percent of all persons, the following general rule used by the Census Bureau is reasonable: an item with an index of inconsistency less than 20 has a low level of response variance; an item with an index between 20 and 45 has a moderate response variance, and; an item with an index over 45 is considered highly inconsistent. Using these guidelines, 54 percent of the items included had low response variability, 33 percent had moderate response variability, and 13 percent had high response variability.

Items with Large Measurement Errors

The gross difference rate, the net difference rate, and the index of inconsistency are very often related to each other. An item which has a high estimate for one of the statistics is usually found to have at least one of the other two statistics which is larger than average. This finding helps

in accomplishing the goal of identifying particular survey items that are not very reliable. Some of the items which exhibit relatively high measurement errors are discussed below.

Of all the items asked for both preprimary and primary school children, only two could be considered to have large measurement errors. The item about how many hours the child spent watching television has relatively large index of inconsistency and difference rates. This may be due to several factors, including the general ambiguity of the item, the crude measurement scale (whole hours) relative to the internal variability in the item, and differing circumstances (32 percent of the differences for this item were attributed to the situation changing).

The other item in this series which is worth noting is the one about how often the child is read to. This item has a large gross difference rate, but a moderate index of inconsistency and net difference rate. About 27 percent of the difference noted between the original and reinterview were attributed to changes in the child's situation. This item had specific pre-coded response options which the respondent was asked to use in their response. Nearly all the differences reported involved a difference of plus or minus one value of the scale.

In the preprimary series of questions, the two items that ask whether the daycare center or the nursery school/prekindergarten is a Head Start program have large measurement errors. While these items were only asked for 31 (for the daycare centers) and 52 (for the nursery school/ prekindergartens) children, all three of the statistics used indicate that the questions have response problems. The cause of the response problems for these items may be the parent's lack of knowledge about what constitutes a Head Start program. The child's situation changing is not a contributor to the response problems for these items.

Another related item that had large measurement errors was the one that asked parents to classify the program as a nursery school, prekindergarten, or Head Start program. The classification of these preschool programs is not simple and the measurement errors reveal that parents may not be able to do this very well.

The only other item in the preprimary series that showed very large measurement errors was the one that asked how often the parent talked with the daycare center provider. The same item for children who attended a nursery school/prekindergarten had a large gross difference rate, but small net difference rate. The daycare center question was only asked for 33 children. One of the problems respondents might face with this item is defining what constitutes talking to the provider. Some parents might include conversations with the provider when picking the child up at the end of the day while some might restrict it to more formal discussions.

Among the primary school children items, no items were observed to have a large gross difference rate, net difference rate, and index of inconsistency. Despite this, three items are worth noting mainly because they have a large index of inconsistency. One is the age when the child started kindergarten, which has a large gross difference rate and index of inconsistency. This item asked parents to give the month and year when the child started kindergarten. Most parents probably did not have this date memorized and thus were required to mentally construct the answer. This construction could have contributed to much of the problem.

The other two items that had large indexes of inconsistency were the one that asked how often the parent talked to the child about school and the one that asked if any of the child's previous daycare programs had an educational program. While the results raise some questions about the reliability of these items, the relatively low gross and net difference rates do not indicate that substantial problems are present.

Items Requiring Recall

About 10 items in the primary school interview and a few items in the preprimary school interview asked the parent to recall past activities of the child, such as whether the child ever attended a daycare center. The items concerning retention in kindergarten and primary school are discussed in a later section.

Except for question about an educational program in the daycare center which was discussed earlier, the statistics for the recall items are very similar. The gross difference rates run from 5.2 to 9.9 percent, the net difference rates range from -3.3 to 4.5 percent, and the indices of inconsistency range from 15.8 to 31.3. These relatively low measurement error statistics indicate that the recall items worked well. The items were well-defined for the parents and they typically repeated the same response in the reinterview as given in the original interview. This finding suggests that limited recall of well-defined and salient activities of children for future administrations of NHES are reasonable.

Enrollment and Retention Items

About 9 items were asked about children's current or past enrollment or retention in kindergarten or first grade and above. The items for preschool arrangements and the item which

asked when the child started kindergarten, which were already mentioned, are excluded from this discussion.

For virtually all of the enrollment and retention items, the three statistics used to approximate measurement errors are very small. In general, parents responded consistently to these items over both administrations of the interview. The statistics suggest that the items related to enrollment and retention are very reliable.

The initial item which asks if the child is attending or enrolled in school has larger measurement errors than any of the other items of this type. Even for this item, the gross difference rate is only 4.5 percent, the net difference rate is 0.8 percent, and the index of inconsistency is 12.7. This item has the same wording as used in the Current Population Survey.

Response problems for this item, which is asked for all children regardless of their age, appear to be associated with almost entirely preprimary school age children. In particular, children who are in nursery school or prekindergarten programs may be sometimes classified as enrolled while at other times as not enrolled. Sixteen of the 17 response errors were found in the 197 preprimary interviews. In the NHES:91, this was not a problem since other questions were used to direct the flow of the interviews and classify the child. However, these results do indicate that the item may have high response errors when used for young children.

5. Summary

The reinterview program for the Early Childhood Education component in the NHES:91 was designed to help identify specific items in the interviews that were not reliable, to quantify the response variance for groups of items, and to provide feedback for future administrations of the interviews. The reinterview program accomplished all three of these objectives.

The results of the reinterview are encouraging. Most of the items included in the reinterview had small to moderate measurement errors. For the specific items with high measurement errors potential problems associated with most of these items included vague or ambiguous classifications, and parents' lack of knowledge about the information requested by the item.

One of the weaknesses of the reinterview program was its limited scope. Only slightly over 500 reinterviews were conducted and this limits the ability to look more closely at the distribution of errors by characteristics of the respondents. For example, the type of analysis done by O'Muircheartaigh (1986) on the correlates of response errors can not be measured with a sample of this size. In future administrations of the NHES, reinterviews will still be conducted using the same basic methods, but the size of the program may be increased if these types of analyses are viewed as important.

References

Biemer, P. and Stokes, L., (1991) "Approaches to the Modeling of Measurement Errors," Chapter 24, *Measurement Errors in Surveys*, Ed. by Biemer, et al., John Wiley and Sons.

Brick, Collins, Celebuski, Nolin, Squadere, Ha, Wernimont, West, Chandler, Hauskens, and Owings, (1991) "NHES:91 Preprimary and Primary School Data Files User's Manual," U.S. Department of Education, National Center for Education Statistics.

O'Muircheartaigh, C.A., (1986) "Correlates of Response Inconsistency in the Current Population Survey," Proceedings of the Second Annual Research Conference, U.S. Bureau of the Census.

[1] The number of days was missing for four of the cases and these are not included in the table.

Table 3. Gross difference rate, net difference rate, and index of inconsistency for items included in the ECE reinterview.

	Sample Size	Estimate* (Percent)	Gross Difference Rate	Net Difference Rate	Index of Inconsistency
Attending or enrolled in school?	534	78	4.5	0.8	12.7
Grade or year attending?	392	–	1.0	-0.5	1.3
Age as of January 1, 1991?	534	–	0.2	0.2	0.2
Missing grade? (derived)	534	–	5.6	0.0	15.4
How often read to child?	532	–	23.3	3.4	34.1
How many children's books?	533	–	5.1	-0.6	27.9
Family get daily newspaper?	534	54	6.2	0.2	12.5
Hours per day watching TV?	528	–	32.8	-5.5	41.7
Receive care from relative?	311	16	5.1	0.0	17.9
Relative care in own home?	44	32	4.5	4.5	8.3
Ever receive care from relative?	248	16	8.5	3.6	26.7
Receiving care from non-relative	311	14	5.8	2.6	22.2
Non-relative care in own home?	31	19	3.2	-3.2	10.0
Ever receive care from non-relative?	255	25	9.8	2.0	24.4
Going to daycare center?	311	18	6.8	-1.0	27.3
Number of daycare centers	35	–	2.9	2.9	-
Daycare center have instruction?	32	89	6.3	-6.3	18.3
Daycare center head start?	31	21	12.9	-12.9	107.0
How often talk w/care provider?	33	–	18.2	18.2	28.0
Ever gone to daycare center?	250	23	5.2	2.8	14.6
Going to Nursery/prek/head start?	311	27	8.7	3.5	21.6
Number of nursery/prek/head start?	77	–	11.7	-9.1	60.9
Nursery school/prekindergarten/head start?	60	–	23.3	16.7	40.5
Nursery/prek/head start have education?	60	84	1.7	1.7	-
Head start program?	52	15	13.5	-13.5	67.7
How often talk w/teacher?	60	–	18.4	1.7	28.7
Ever gone to nursery/prek/head start?	209	31	7.2	1.4	15.5
One or two kindergartens?	435	99	0.0	0.0	-
Full- or part-day kindergarten?	99	42	0.0	0.0	0.0
After kindergarten care program?	99	8	5.1	5.1	25.6
1st or 2nd year of kindergarten?	99	96	2.0	-2.0	34.4
Enroll kindergarten when old enough?	99	94	9.9	0.0	0.0
Grade attending last year?	221	–	0.0	0.0	0.0
Ever attend kindergarten?	221	98	0.5	0.5	-
Age when started kindergarten?	209	–	27.4	0.6	56.2
Attend kindergarten 1 or 2 years?	220	95	0.0	0.0	0.0
Enrolled kindergarten when old enough?	220	95	0.5	-0.5	5.8
Changed schools since started 1st?	219	34	1.8	0.9	3.8
Repeated grade since starting 1st?	222	6	0.9	0.9	0.8
Attend public or private school?	222	89	0.0	0.0	0.0
How often talk w/child ab. school?	222	–	2.7	-2.7	60.2
Teacher contacted about behavior?	222	23	10.2	-1.3	26.1
Teacher contacted about schoolwork?	221	25	6.3	0.9	16.9
Number of days has homework?	221	–	10.4	5.0	15.2
Ever receive care from relative?	222	29	9.0	3.6	20.9
Ever receive care from non-relative?	223	26	9.9	4.5	22.3
Ever go to daycare center?	223	36	8.1	-2.7	16.7
Ever go to nursery/prek/head start?	221	55	7.2	-0.9	14.7
Any daycare have educational?	152	91	5.3	4.0	52.8
Any daycare head start?	150	29	6.0	-3.3	19.2

* If the estimate is not a dichotomy, then a – is entered for the estimated percent.