

Data Comparability in a Mixed Mode Telephone and Face to Face Survey of Persons with Disabilities

Matt Sloan, Debra Wright, Kirsten Barrett
Mathematica Policy Research, Inc.

Abstract

The purpose of this analysis is to compare data collected via telephone and face-to-face interviewing on a national survey of disability beneficiaries. Krosnick (2002) and others have suggested that telephone interviewing may pose more of a cognitive burden on respondents than face-to-face interviewing and that this effect may be more pronounced for individuals with reduced cognitive abilities. Since the sample for this survey is comprised entirely of individuals with physical and mental impairments, we were interested in whether survey mode would have an impact on data quality indicators such as acquiescence, social desirability, item non-response, and non-differentiation in response.

To determine whether there were differences in data quality between telephone and face-to-face modes, chi-square analyses were performed to determine whether there was an association between mode and response for select items. Since participants were not randomly assigned to mode, we controlled for the non-experimental nature of the study by comparing subsets of respondents matched by demographic characteristics on several indicators of data quality.

The results indicate that for the key items tested, there was more evidence of acquiescence, more social desirability in response, and more item non-response in CATI than CAPI. We conclude that mode of data collection may have effected data quality in this survey. However, we found that this effect was greatest for items that were vague or cognitively demanding. Factual questions about behavior appear to show the minimal, if any, mode effects.

Keywords: Mode effects, Data quality, Telephone interviewing, Face-to-face interviewing

1. Introduction

Multi-mode designs are increasingly employed to boost response rates while controlling costs. While computer assisted telephone interviewing (CATI) remains a cost-effective mode of data collection, response rates have dropped in recent years. On the

other hand, face-to-face or computer assisted personal interviewing (CAPI) generally yields higher response rates, but is often prohibitively expensive. Combining these (and other) modes of data collection can result in a design which takes advantage of the benefits of both, while minimizing the downsides often associated with each alone. In the case of the National Beneficiary Survey (NBS), a survey of disability beneficiaries sponsored by the Social Security Administration (SSA), telephone interviewing was combined with face-to-face interviewing to control costs while providing adequate coverage of the population, maintaining high response rates, and enhancing accessibility to the interview.

While mixed mode methodologies may reduce non-response bias, measurement error can be introduced if the mode of data collection has an independent effect on the interview process and the data collected (Voogt & Saris, 2005). The purpose of this analysis is to compare the data collected in telephone and face-to-face interviews on a variety of measures of data quality. Since the NBS sample is comprised entirely of individuals with physical and mental impairments, we were particularly interested in whether the effect of mode on data quality would be evident in this survey.

2. Background

While multi-mode surveys offer several benefits such as providing multiple channels for reaching sample persons and allowing sample persons to respond in the mode most convenient for them, there is a risk that the data collection methods will not produce equivalent results (Lyberg & Kasprzyk, 1991). In the case of a dual mode telephone and face-to-face survey such mode effects could be the result of inherent differences in the attributes of these two modes. While telephone and face-to-face interviews are similar in many ways since both involve an interviewer, they are different in the channels of communication available to the interviewer and respondent (de Leeuw, 2005). This, combined with norms of conversation via telephone compared to face-to-face, can result in differences in ability to build rapport and engage respondents, respondent's willingness to reveal information, the pace of the interaction, and the cognitive complexity of the task experienced by the respondent.

Because face-to-face interviewers are physically present, there is more opportunity to develop rapport and maintain the respondent's interest and motivation. The face-to-face format allows interviewers to read and use body language to more easily pick up on respondent confusion or frustration. Telephone interviewers on the other hand, must rely on pauses and tone of voice which may make it more difficult to notice inadequate responses (de Leeuw & van der Zouwen, 1988).

The pace of a telephone versus a face-to-face interview may also vary. To keep up with the pace of conversation and avoid lapses in conversation, telephone respondents may not easily be able to spend as much time as they would like on any given item. In a face-to-face setting, interviewers may find it easier to match pace and communication style by using body language cues. This feeling of time pressure may interfere with cognitive processing and increase reliance on strategies which simplify the task such as acquiescence; defined as the tendency to agree with an item, regardless of its content (Krosnick, 1991).

There is evidence that there are differences in data quality between modes. For example, the quantity of information provided in response to an open-ended question is often less in telephone interviews compared to face-to-face interviews. Telephone respondents have also been found to display more acquiescence, to choose more extreme categories, refuse more items, and display more evidence of recency effects than face-to-face respondents (Jordan, Marcus, & Reeder, 1980; Locander & Burton, 1976). This suggests that face-to-face interviewing may have an advantage when it comes to collecting complex information on topics that are not immediately in the forefront of the respondent's consciousness. Face-to-face interviewers may be better equipped to provide clarification or to probe responses. Additionally, the increased ability of face-to-face interviewers to follow respondent cues and modify the pace of the interview as necessary, may also give respondents more time to process and respond to complex questions. Krosnick (2002) suggests that the higher cognitive demand placed on respondents, the more likely they are to take shortcuts to simplify the task, or to engage in "satisficing". Additionally, respondents with limited cognitive abilities and low motivation may be most likely to exhibit "strong satisficing" behavior and make efforts to give a seemingly reasonable answer while putting in minimal effort.

Despite these findings, few differences in actual estimates of behavior have been found for non-

threatening questions. The impact of mode of administration on response is greatest when the topic of the questions are sensitive. Although comparisons of estimates obtained in face-to-face and telephone modes are not entirely consistent, results of previous studies generally support the notion that respondents to face-to-face surveys over-report socially desirable behavior. Sykes and Collin's (1988) review of four comparative studies reveals that more socially desirable answers were consistently given in face-to-face than telephone. However, Holbrook, et al. (2003) report that telephone respondents were more likely to present themselves in socially desirable ways than face-to-face respondents. The direction of this mode effect may depend on the specific needs of the population being studied. If social distance is more important in encouraging valid response, telephone interviewing may lead to more accurate reporting since interviewer presence can decrease the level of anonymity and privacy provided. However, if establishing legitimacy and rapport is more important, face-to-face interviewing, which allows interviewers to increase the respondents' belief in the confidentiality of the data and the legitimacy of the survey institution, may make respondents more inclined to reveal sensitive information.

2.1 Research Question

The purpose of this investigation was to compare data collected in the telephone and face-to-face modes to determine if there were differences in data quality by mode for selected items. Because the sample for the NBS includes people with mental as well as physical impairments, we tested the hypothesis that data collected face-to-face would be of higher quality than that collected via telephone for particular items. This hypothesis suggests that telephone interviewing poses a greater cognitive burden and leads to more satisficing behavior than face-to-face interviewing for this population.

We addressed the research questions by examining differences in the quality of data collected in telephone and face-to-face mode on the NBS as evidenced by variation in item non-response, socially desirable responses, non-differentiation, and acquiescence for select items in the survey.

2.2 Description of Survey

The National Beneficiary Survey (NBS), conducted by Mathematica Policy Research, Inc. and sponsored by the Social Security Administration (SSA), is a nationally representative survey of 18 to 64 year old

SSA disability beneficiaries. The 45-minute, dual-mode (CATI/CAPI) survey gathered information on health, insurance, employment, income, and demographic characteristics. Interviews were attempted first by telephone. Face-to-face interviews were then conducted with people who could not be located, requested or required an in-person interview, were evasive to telephone attempts, or who refused to participate by phone. There is both a cross-sectional component and a longitudinal component in which a cohort of beneficiaries are followed for several rounds.

In 2004, the first round of the survey was fielded with a sample of 10,530 SSA beneficiaries. If the beneficiary was incapable of responding for him or herself, a proxy interview with a knowledgeable informant was attempted. A total of 7,603 cases were completed for an overall weighted response rate of 77.6%. Of those, 6,302 were completed by CATI and 1,301 were completed by CAPI.

3. Analysis

This study used one-to-one matching on a range of respondent characteristics to identify comparable groups of sample members who completed the survey either on the telephone or by face-to-face since random assignment was not possible after the fact. While this method does not guarantee comparability between the groups, matching individual cases on important characteristics is widely accepted as a quasi-experimental alternative to random assignment. However, observed differences in the two groups may be the result of self-selection or other factors unrelated to the mode of data collection. As such, the results should be viewed as preliminary and will hopefully prompt more rigorous investigation.

For purposes of this analysis, cases completed by proxies (n=1,999) and cases that were started in one mode but completed in the other (n=105) were first removed from the analysis file. This left 4,616 cases completed by the beneficiary by CATI and 906 cases completed by the beneficiary by CAPI. To match cases that were completed in the two modes, each case that was completed by face-to-face was matched to a case that was completed on the telephone if an exact match could be made on specific major sample frame variables. Cases were matched by six characteristics: age at interview, race, ethnicity, sex, benefit type (SSI, SSDI, or both SSI and SSDI), and SSA impairment type from administrative records. These characteristics were chosen because they are all related to mode of interview. If a face-to-face case matched more than one telephone case, one matching telephone case was randomly selected. This process yielded a data set

with 772 observations comprised of 386 cases completed in CATI and 386 cases completed in CAPI.

3.1 Measures

We examined several indicators of data quality including comparisons of item non-response, proportion of socially desirable responses, amount of non-differentiation, and acquiescence. Because research generally indicates that differences in modes are most evident when respondents are required to answer difficult or complex questions where burden is higher, we attempted to examine key questions that were more subjective, sensitive, vague, or that could be construed as cognitively demanding questions. Finally, we looked only at items that were administered to all respondents.¹

3.1.1 Item Non-Response

Item non-response occurs when a respondent is unable or unwilling to answer a question. To estimate the impact of data collection mode on item non-response, we created dummy variables for key items that had more than two percent nonresponse overall to indicate that the response was valid or was coded as either Don't Know or Refused. These items included age that the beneficiary first became limited, work goals ("You see yourself continuing {to work/working} for pay in the next year", and "You see yourself continuing to work/working} for pay in the next five years." with response options ranging from strongly agree to strongly disagree), household income in 2004, and certain demographic questions (race, father's education, mother's education, and weight). In addition, where several questions related to the same topic, we created a dummy variable that indicated whether the responses to any question in a series of questions was missing. Five such series were investigated. The first included five items asking about awareness of particular SSA programs (ever heard of Blind Work Expenses, ever heard of Expedited Reinstatement, ever heard of Benefits Specialists, and ever heard of the Ticket to Work Program). The second was a series of four items asking about use of employment services (ever received employment services, job training, medical services, or

¹ An item that is sometimes on and sometimes off path cannot usually be statistically evaluated for mode effects. The routing creates a subset of the matched dataset, and this subsetting can destroy the randomness of the matching. This routing issue exists for experiments as well as for matching and is a common problem in entirely different fields of study such as clinical trials (Pierzchala et al, 2005) The implication is that only variables that are asked to all beneficiaries are used for this analysis.

mental health services to help get a job or help live independently). The third was a series of four health and functional status items (how much physical health problems limited physical activities, the degree of difficulty doing daily activities, how much emotional problems limited social activities, and how much emotional problems limited daily activities asked on a 5-point Likert scale ranging from “Not/None At All” to “Could Not Do Physical Activity/Daily Work/Social Activity”). The fourth was a series of four items asking about health insurance (currently covered by Medicare, Medicaid, military care, or private insurance) and the fifth was a series of items asking whether income was received from any of eight sources in the last month. Chi-square tests were performed to determine whether there was an association between mode and nonresponse for the individual and series of items described above. We expected more item non-response in CATI than in CAPI.

3.1.2 Social Desirability

The social desirability effect occurs when people are unwilling to admit holding undesirable opinions or to report undesirable behaviors. We selected several items that could be perceived as sensitive and compared the estimates across modes. These items included the work goals items mentioned above in addition to the item “Do your personal goals include {getting a job/moving up in a job}, or learning new job skills?”, alcohol use (“In the past 12 months, have you ever felt you ought to cut down on your drinking?”), drug use (“During the 12 months have you used drugs on your own more than 5 times?”), and household income in 2004. Chi-square tests were performed to determine whether there was an association between mode and response for these sensitive items. In addition, a t-test was performed to determine if the mean reported income differed between the two modes. Since our review of the literature on social desirability by mode revealed no consistent trend, given our population, we expected more socially desirable responses in CATI than in CAPI since in person interviewing provides more opportunity to build rapport and trust.

3.1.3 Non-Differentiation

Non-differentiation occurs when respondents fail to distinguish between different questions and select the same answer choice on a scale for all, or almost all, similar questions. For analysis comparing the amount of non-differentiation, we examined response patterns to the series of four health and functional status items mentioned above. A dummy variable was created that indicated whether a respondent choose the same response option for all four items or not. A chi-square

test was performed to determine whether there was an association between mode and the likelihood of choosing the same response for all four items. We expected more non-differentiation in CATI than in CAPI.

3.1.4 Acquiescence

To measure the amount of acquiescence we counted the number of “yes” responses to the series of questions about awareness of SSA programs (listed above). These items were chosen because they were asked in the same series and because it was reasonable to assume that respondents who were inclined to acquiesce might say that they had heard of these programs without making the effort to determine if, in fact, they had. A chi-square test was performed to determine whether there was a significant association between mode and whether or not a respondent answered “yes” to each of these four items. We also created a dummy variable which indicated whether respondents answered “yes” to all four items or not. This was used to perform a chi-square test comparing mode and the tendency to respond “yes” to all four items. We expected more acquiescence in CATI than CAPI.

All analyses were done using SAS (Version 8). Cross tabulations were developed to test the hypothesis of no association between mode for categorical variables. In cases where expected counts in one or more cells in the contingency table was less than five, Fisher’s Exact Test was used rather than the chi-square statistic. Means were calculated to test the hypothesis of no difference between mode for continuous variables. A t-test was used to determine whether a significant difference existed between modes for these items.

4. Results

4.1 Item Non-Response

Non-response was significantly associated with mode for three of seven items examined at the $p < .05$ level. CATI non-response was higher than CAPI non-response on the following items: race, father’s education and mother’s education (see Table 1).

Table 1: Item Non-Response by Mode
AAPOR - ASA Section on Survey Research Methods

<i>Item</i>	<i>CATI</i> <i>Non-Response</i>	<i>CAPI</i> <i>Non-Response</i>	<i>Total</i> <i>Non-Response</i>	<i>Chi-Square Statistic,</i> <i>p-value</i>
See Working for Pay Next Year	2.1% (n=8)	4.04% (n=17)	3.2% (n=25)	$\chi^2=3.35, p=.067$
See Working for Pay Next Five Years	4.9% (n=19)	4.7% (n=18)	4.8% (n=37)	$\chi^2=0.28, p=.866$
Race	4.2% (n=16)	0.78% n=(3)	2.5% (n=19)	$p=.0042^2$
Father's Education	40.0% (n=152)	23.8% (n=92)	31.6% (n=244)	$\chi^2=21.57, p<.0001$
Mother's Education	25.4% (n=98)	17.4% (n=67)	21.4% (n=165)	$\chi^2=7.41, p=.007$
Respondent's Weight	4.2% (n=16)	2.6% (n=10)	3.4% (n=26)	$\chi^2=1.43, p=.231$
Household Income	36.0% (n=139)	33.4% (n=129)	34.7% (n=268)	$\chi^2=.572, p=.450$

Among the five series of items analyzed, only two series showed a significant association between mode and whether at least one item in the series had a substantive response or not. For the health and functional status series, nine percent of the CATI respondents had at least one Don't Know or Refusal response in the series compared to four percent of CAPI respondents ($\chi^2=9.26, p=.002$). For the health insurance series of items, six percent of the CATI respondents had at least one Don't Know or Refusal response compared to three percent of CAPI respondents ($\chi^2=4.43, p=.035$). No significant association was found between mode for the series of questions on awareness of SSA programs, use of employment services, or income sources.

4.2 Social Desirability

There was a significant association between mode and whether the respondent answered "yes" to having goals to move up in a job or learn new skills, having goals for working in the next year, and having goals for working in the next five years (See Table 2). The percentage of respondents who gave the socially

reported an average \$15,696 and CAPI respondents reported an average of \$11,382 ($t=2.39, p=.017$) earned in the previous year.

4.3 Non-Differentiation

There was no significant association between mode and each of the series of four health related items that we tested. Only a small proportion of respondents in either mode (21% in CATI and 19% in CAPI) gave the same answer to all four of these items ($\chi^2=.399, p=.527$).

4.4 Acquiescence

As shown in Table 3, there was a significant association between mode and those who provided an affirmative response for three of the four awareness items tested. Additionally, CATI respondents were more likely to give a "yes" response to all four items than CAPI respondents were (5% of CATI respondents compared to less than 1% of CAPI respondents; $\chi^2=11.98, p=.001$).

Table 2. Social Desirability by Mode

<i>Item</i>	<i>CATI</i>	<i>CAPI</i>	<i>Chi-Square Statistic,</i> <i>p-value</i>
Goals Include Moving Up (Yes)	59.3% (n=223)	45.8% (n=175)	$\chi^2=13.84, p=.000$
See Working For Pay Next Year (Strongly Agree Or Agree)	38.3% (n=148)	28.0% (n=108)	$\chi^2=9.35, p=.002$
See Working For Pay Next Five Years (Strongly Agree Or Agree)	50.8% (n=196)	42.5% (n=164)	$\chi^2=5.33, p=.021$
Felt Need To Cut Down On Drinking (No)	8.1% (n=31)	8.9% (n=34)	$\chi^2=.16, p=.689$
Used Drugs In Last 12 Months (No)	5.2% (n=20)	3.7% (n=14)	$\chi^2=1.08, p=.300$

desirable answer, "yes" to these three items was consistently higher among CATI respondents than among CAPI respondents. However, neither of the two questions about problem behavior (use of alcohol or drugs) yielded significant associations with respondent mode. There was a significant association between the amount of total household income reported and mode of interview. CATI respondents

5. Discussion

5.1 Item Non-Response

For the individual items we tested, four out of seven showed a significant association between mode of interview and item non-response. For these four times, there was less item non-response in CAPI. Given that some of these questions may be more cognitively challenging than others in the main body of

² Fisher's Exact Test since expected cell counts were >5.

Table 3. Acquiescence by Mode
 AAPOR - ASA Section on Survey Research Methods

<i>Item</i>	<i>CATI</i>	<i>CAPI</i>	<i>Chi-Square Statistic, p-value</i>
Heard Of Blind Work Expense	13.95% (n=53)	6.01% (n=23)	$\chi^2=13.42, p=.000$
Heard Of Expedited Reinstatement	22.11% (n=84)	11.55% (n=44)	$\chi^2=15.15, p=.000$
Heard Of Benefits Specialist	22.22% (n=84)	7.96% (n=30)	$\chi^2=29.96, p=.000$
Heard Of TTW	40.63% (n=156)	38.06% (n=145)	$\chi^2=.53, p=.467$

the questionnaire (e.g. father's education), the face-to-face approach may cause respondents to make more of an effort to recall the information (or retrieve it) than in the CATI mode. Similarly, we surmise that a non-response may seem futile when the interviewer is able to observe situations or behaviors for him or herself, such as with race. We were surprised to find no strong evidence of differences in willingness to offer sensitive information, such as income, since a face-to-face interaction may offer more opportunity to establish trust and rapport than in the CATI mode.

Only two of the five series of items showed a significant association between mode of interview and non-response in a series of items. The finding that a higher percentage of CATI respondents provided a Don't Know or Refused response to at least one item in the series of health and functional status but not for program use and income suggests that the mode effect was more evident for subjective items. It is unclear why CATI respondents would have more missing data on the health insurance items, although it is possible that CAPI respondents were more likely to look up such information with someone physically present in their home.

5.2 Social Desirability

We also observed some evidence of more socially desirable responses being offered in CATI than CAPI. Respondents were significantly more likely to report having work goals and to have higher expectation of future work activity in CATI rather than CAPI. We also found that CATI respondents reported higher mean household incomes. This is consistent with the Holbrook et al. (2003) finding that respondents are more likely to report socially desirable behavior in CATI than in the CAPI mode. This may be because the ability to establish trust and rapport with respondents with higher incidence of cognitive or physical disability is more important than the anonymity that CATI provides. However, the difference may also be an artifact of underlying differences in those respondents who responded by CATI rather than CAPI.

We were surprised to find no effect for questions related to drug use and drinking since these items are generally considered quite sensitive. However, cell sizes for these items after matching were small which may have contributed to the lack of effect. These findings should be further explored in a more rigorous randomized study.

5.3 Non-Differentiation

We found no significant association between the amount of non-differentiation displayed between modes for the items we tested. These items may not have provided the best means to test this however. While they are part of a series of like items, items using similar but slightly different response scales were interspersed. Interviewers are therefore instructed to read the response options for each item as it appears. This may have had the effect of heightening attention to the items and could have minimized satisficing in both modes.

5.4 Acquiescence

We did find that respondents in the CATI mode were more likely to agree that they had heard of various SSA work incentives. While this supports the hypothesis that respondents to CATI are more likely to exhibit satisficing behaviors than CAPI respondents, it is possible that these findings reflect real differences in the two populations. Given that the CATI respondents generally have higher educations and income, it is likely that they are also better informed about SSA programs. This group is also easier to contact which may make it easier for SSA to disseminate information to them.

6. Conclusion

In general, this study found some evidence that mode of data collection may impact data quality for this population. Data collected via CAPI tended to have lower item non-response and showed less evidence of socially desirable responses and acquiescence. These findings are generally consistent with the literature on mode effects.

It is important to note, however, that the evidence does not entirely support the conclusion that data collected by CAPI is consistently of better quality than that collected by CATI. We found no difference between the modes when we investigated non-differentiation. In addition, the results for the three other measures, while generally consistent with the assumption that data quality would be better in CAPI than CATI, did not completely support this conclusion as several tests were non-significant.

While our results are somewhat mixed, it appears that items that were vague or demanded more attention and cognitive processing showed the greatest differences in data quality. For example, questions about the future showed both higher item non-response and socially desirable responses when collected by CATI. On the other hand, factual questions about behavior appear to show the fewest mode effects.

These results should not be interpreted to mean that data collected by CATI is of poor quality. To the contrary, most of the data we looked at showed small differences in these measures, if any. Rather, it suggests that researchers should carefully consider the interplay of question content or complexity and mode of data collection in the design phase and that consider how mode effects may be influenced by the population being studied. Specifically, in studies of people with mental and physical disabilities with high prevalence of cognitive limitations, in person interviews may provide better quality data on items that are more complex or challenging.

Perhaps our chief conclusion is that the underpinning issues in this study should be examined with a more rigorous design that includes random assignment. Some of the conclusions we reached could well be explained by differences that we did not control for in our one-to-one matched comparison group design.

References

- de Leeuw, E. (2005). To mix or not to mix data collection modes in surveys. Journal of Official Statistics, 21, 233-255.
- de Leeuw, E. & van der Zouwen, J. (1988). Data quality in telephone and face to face surveys: A comparative meta-analysis. In R. M. Groves, P. P. Biemer, L. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.). Telephone Survey Methodology (pp 283-299). New York: Wiley.
- Holbrook, A. L., Green, M. C., Krosnick, J. A. (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, 79-125.
- Jordan, L. A., Marcus, A. C., & Reeder, L. G. (1980). Response styles in telephone and household interviewing: A field experiment. Public Opinion Quarterly, 44, 210-222.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology, 5, 213-236.
- Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., Kopp, R. J., Mitchell, R. C., Presser, S., Ruud, P. A., Smith, V. K., Moddy, W. R., Green, M. C., Conaway, M. (2002). The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? Public Opinion Quarterly, 66, 371-403.
- Locander, W. B., & Burton, J. P. (1976). The effect of question form on gathering income data by telephone. Journal of Marketing Research, 13, 189-192.
- Lyberg, L. & Kasprzk, D. (1991). Data collection methods and measurement error : An overview. In P. Biemer, R.M. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman (Eds.) Measurement errors in surveys (pp. 237-257). New York: Wiley.
- Pierchala, M., Wright, D., Wilson, C., and Guerino, P. (2005). Telephone Collection as Part of a Multi-Mode Survey. Paper presented at the Second International Conference on Telephone Survey Methodology, Miami, 2006.
- Sykes, W., & Collins, M. (1988). Effects of mode of interview: Experiments in the U.K.. In R. M. Groves, P. P. Biemer, L. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.). Telephone Survey Methodology (pp 301-320). New York: Wiley.
- Voogt, R. J. & Saris, W. E. (2005). Mixed mode designs : Finding the balance between nonresponse bias and mode effects. Journal of Official Statistics, 21, 367-387.