Assessing and Adjusting for Response Error Using a Multi-Phase Survey Approach in the Residential Energy Consumption Survey

Marilyn Worthy, Katie Joseph
Energy Information Administration, 1000 Independence Ave SW,
Washington, DC 20585

Abstract

The Residential Energy Consumption Survey (RECS) is a quadrennial survey sponsored by the Energy Information Administration (EIA) which collects energy data on a sample of U.S. households. The RECS questionnaire is technical, challenging respondents' capacity to report information on energy use, equipment, and related expenses. Householders who do not pay their energy bills directly, roughly 10% of the population, are more prone to sources of response error: difficulty understanding an energy question, not having access to information, and inability to produce an accurate response. EIA fills this quality gap by conducting a supplemental rental agent survey (RAS) for these households. We examine 2009 RECS data in households with both a household and a RAS interview to show patterns of response and nonresponse bias. We discuss the impact on key estimates and energy models, the utility of this approach in terms of the survey cost-error tradeoff, and the value of further tailoring the scope of the RAS in future iterations of RECS. Since the RAS is an abbreviated version of the household survey, it offers insights to the nature of response error and the value of using a targeted, multi-phase survey approach.

Key Words: item nonresponse, measurement error, piggyback survey, household survey, multi-phase survey, editing

1. Introduction

Item nonresponse and respondent error are common problems that arise in conducting household surveys. Item nonresponse occurs when the respondent answers "Don't Know" or "Refuse" to a survey question. The following are some reasons for item nonresponse: the respondent lacks the information necessary to answer the question, the respondent refuses to give an answer because they feel it is irrelevant, the interviewer fails to record the answer, or the response is rejected based on an edit failure and subsequently erased (Kalton, 1983). In addition, respondents sometimes are disinterested in the topic and are unwilling to put any effort into recalling accurate information. An answer or an item nonresponse can happen at any component of the response process: comprehension, retrieval, judgment, and response (Tourangeau, 1984, 1987; Tourangeau et al., 2000). Respondents answer "Don't Know" because they do not comprehend the question, cannot retrieve the correct memory, do not use the memory to make an accurate judgment, or cannot map their memory to the response set.

Response error, or more generally measurement error precipitated by the respondent (as opposed to the questionnaire, mode, or interviewer), is caused when respondents report an erroneous answer. These erroneous responses can be systematic, resulting in biased survey results and misguided conclusions. Response bias is the systematic difference between real values for a particular question and the responses adjusted for random errors (Groves, 1989; Saris & Gallhofer, 2007). Response error can occur at any point

throughout the response process as well; a respondent misunderstands the question, they remember a memory inaccurately, they estimate the answer based on an incomplete memory, or they misunderstand the choices in the response set. Respondents give an inaccurate response based on conformity bias, desirability bias, acquiescence bias, and other inadequacies (Tourangeau et al., 2000). Several tools are used by survey researchers in order to decrease response errors. For the comprehension component, definitions are useful in assisting the respondent; for retrieval, questions filled-in with answers to previous questions can be useful; for judgment, interviewer probes can be useful; and for response, interviewers sometimes use show cards with pictures. Other possible options can include changing the question to ask for less detailed information, helping the respondent estimate the answer, or changing the objective (Fowler, 2002).

This paper explores an alternative method that was used to reduce item nonresponse and measurement error: conducting a follow-up survey with a second, more knowledgeable respondent. With this method, the survey researcher is less reliant on the primary respondents' capacity to produce accurate answers. The third-party respondent would be more knowledgeable about the questions difficult to the primary respondent. The term often used for this method is "piggyback," where data is collected solely for the purposes of assisting another survey, in this case data editing and quality control (Tourangeau & Smith, 1985). This method, while useful, carries its own set of inherent risks, such as mapping the piggyback response to the primary response set, the third-party respondent being less knowledgeable than expected, or the third-party incorrectly answering for the wrong primary respondent.

The purpose of this study is to explore the piggyback survey, the *Rental Agent Survey*, and evaluate its effect on the primary survey, the *2009 Residential Energy Consumption (Household) Survey*. The focus will be on the extent to which the piggyback survey reduced measurement error and item nonresponse, the evaluation of its effect on the primary survey's estimates, and suggestions to improve efficiency.

2. Data

The Residential Energy Consumption Survey (RECS) is a quadrennial survey, sponsored by the Energy Information Administration (EIA) within the U.S. Department of Energy, which collects data on energy characteristics and usage in U.S. housing units. The survey's primary purpose is to present an official comprehensive picture of how energy is consumed in U.S. homes. In 2009, a nationally representative sample of 12,083 households responded, yielding an AAPOR response rate of 79%.

Data collection for the RECS consists of two field periods. During the first field period, the *household survey* is conducted. A CAPI interview is conducted with a respondent at the sampled housing unit, and information about the features of the house that affect energy usage, such as type of housing unit, number of people living in the home, and heating fuel, is collected. During the second field period, the *Energy Supplier Survey* is conducted - utilities that supply energy to the sampled home self-report data from their administrative records on the amount of energy used by the sampled house and its associated costs. Data collected from the household survey and Energy Supplier Survey are linked, and used to present the results of the RECS. All analyses presented in this paper focus on the household survey.

A subset of the RECS questions is very highly correlated with a household's energy usage; these variables are used to impute a household's energy usage and costs if the

actual values are not obtained from the Energy Supplier Survey in a complex nonlinear model. They are also important to key data users. Many of those questions are covered under the topic of *end uses*. End uses are purposes for which energy is used in the home, such as space heating, cooling, water heating, and cooking.

In general, renters or tenants who occupy without payment of rent that do not pay their own utility bills (about 10% of the population) 1) have more difficulty accurately answering these questions, evidenced by high rates of edit failures, and 2) have high item nonresponse. In order to compensate for measurement error and item nonresponse for these respondents, the *Rental Agent Survey (RAS)* was created. For factual questions, the optimal answer is that which would be reported if the survey sponsor had direct access to the information (Fowler, 2002). The closest entity to this information is the landlord or agent who pays the utility bills for the sampled housing unit. They are able to observe relationships between energy usage and costs, and therefore have a better understanding of the energy characteristics of that unit.

The RAS is conducted in the same field period as the household survey. If a household respondent said that any of their bills were paid for by a landlord or a third party, a RAS case was spawned. The interviewer then attempted to conduct a rental agent interview, via CAPI, immediately after completing the household interview if the rental agent was at the same location (i.e. at a rental office in an apartment building). If an interviewer was not able to contact the rental agent while still on-site or if the rental agent was not located on-site, the case was transferred to a centralized field operation and the interview was conducted via computer assisted telephone interview (CATI); about 68% of the cases were completed via CATI.

The RAS is a much shorter version of the household survey. It consists of less than 50 questions--the exact number depends on skip patterns—and takes about 18 minutes to administer, compared to the household survey which has hundreds of questions and takes approximately 52 minutes to administer. The RAS respondent can be any rental agent, ranging from a landlord of a large apartment complex or an individual who rents their house to the selected unit. In 2009, out of 12,083 completed household interviews, 942 RAS cases were spawned and 584 cases (62%) were completed.

Response rates varied greatly by type of housing unit. Apartment units overall had higher response rates than single family or mobile homes, 64% vs. 48%. Mobile homes and single family detached homes had the lowest response rates of 36% and 31%, respectively. Single family attached homes were an anomaly with a higher response rates on par with apartments, about 71%. Apartments in buildings with 2-4 units and apartments in buildings with 5 or more units had response rates of 52%, and 68%, respectively.

3. Methods

3.1 Questionnaire

As previously mentioned, the questions asked in the RAS correspond to only a subset of questions asked in the household survey. There are six sections: building/apartment structure, heating, water heating, cooling, cooking, and who pays the bills. Question structure and language in the RAS is comparable to the household survey, and the response sets for parallel questions are identical. The following are sample questions from the RAS:

Building/apartment structure: What is the total floorspace in Unit A? Please consider only unit A. Your best estimate is fine.

Heating: In 2009, what was the main fuel used for home space heating in Unit A?

Water Heating: Approximately how old was the main water heater?

Cooling: In 2009, did Unit A use any air-conditioning?

Cooking: What fuel did the stove use?

Who pays the bills: In 2009, was the {insert heat fuel} used for heating in Unit A paid for by the tenant, included in the rent or condominium fee, or paid some other way?

3.2 Reducing Item Nonresponse

At the conclusion of data collection, the data from the household survey were compared to data from the RAS. For cases where the respondent for the household survey did not know or refused to answer a question, the response from the matching question in the RAS was filled in. This relationship was not reciprocal; item nonresponse in the RAS was never corrected or filled in by the household survey because the RAS had no intrinsic value. The sole purpose of the RAS is to "clean" household survey data. There were a few rare situations in which RAS data was not used to fill in household data. These were usually instances where RAS data lacked internal consistency, indicating the validity of the data was questionable. Reducing item nonresponse with the RAS was limited to a subset of the variables. Other variables that were compared between the two surveys were only used for editing purposes.

3.3 Data Reconciliation and Measurement Error

For situations in which there was both a household survey and a RAS response for a particular question, the two responses were compared. In instances where they were the same, no changes needed to be made; the precision of the response in the household survey was confirmed. A similar situation is if the two were not the same but consistent. For example, the household gave a response as a range, and the rental agent gave a point estimate in that range; the household variable was filled in with the more precise point estimate. In instances where the two responses were not the same, they had to be reconciled. Figure 1 below shows the general rules used in reconciling the two data sources.

When the household response was different than the RAS response, the standard rule was that RAS data overruled household data. The exceptions to this rule were outlined as a specific, limited, documented set of exceptions. The exceptions were similar to those described above as reasons household missing data was not filled in by RAS data; either the RAS lacked internal consistency, the RAS respondent could not recall the answers to other related questions, or the household respondent could be more knowledgeable. An example of the third situation would be if the RAS respondent said the housing unit had no heating and the household respondent said they used portable electric space heaters; the rental agent may not be aware that portable heaters were being used in the unit.

	Household Survey		
RAS	Response A	Response B	Don't Know
Response A	No change	Most often RAS override	RAS override
Response B	Most often RAS override	No change	RAS override
Don't Know	No change	No change	Imputed

Figure 1: Editing Rules for Reconciling the RAS and the Household Survey

Interviewer debriefings suggest that many of the differences between the RAS and household responses were due to the household respondent's unfamiliarity with the topics and trouble retrieving factual information. Interviewers recalled instances in which household survey respondents appeared to guess answers, particularly for more technical questions.

4. Results

4.1 Data Editing Effort

For the 584 RECS household cases with a completed RAS, the responses to the edited household dataset were compared to responses in the raw, unedited dataset to assess the overall editing effort. The editing effort was measured by the number of cases with different responses in the two datasets - either a "Don't Know" or "Refuse" response in the household survey was filled in with a RAS response, or a response from the household survey was changed to a different response given by the Rental Agent. Table 1 summarizes the editing effort for the questionnaire items considered in this analysis. The water heating items were the most edited: age of water heating equipment was changed in 64% of the cases and water heating fuel was changed in 60% of the cases due to the RAS. The space heating items also had a high percentage of edited responses: 59% of the cases had the age of heating equipment responses changed and 36% had space heating fuel responses changed due to the RAS.

The "cooking" questionnaire items, such as cooking equipment and cooking fuel, were edited at a much lower rate. Only 7% of the cases had their main cooking fuel changed, and less than 1% had their cooking equipment changed due to the RAS.

Table 1: Summary of Editing Effort by Questionnaire Item

Questionnaire Item	# of Cases Changed	# of Cases Not Changed	% of Cases Changed
Age of water heater	373	211	64%
Water heating fuel	349	235	60%
Space heating equipment age	345	239	59%
Year housing unit was built	261	323	45%
Age of refrigerator	215	369	37%
Space heating fuel	208	376	36%
Type of space heating equipment	184	400	32%
Age of central air conditioning unit	115	469	20%
Main cooking fuel	43	541	7%
Type of air conditioning equipment	40	544	7%
Number of separate ovens in the household	2	582	0%
Number of separate stove tops in the household	2	582	0%
Number of combination stove/oven appliances in the household	1	583	0%

The editing effort varied considerably by housing type. Apartments in buildings with 5 or more units had the most cases edited due to the RAS for every variable, and the most RAS cases by far (417). Apartments in buildings with 2-4 units had the second highest number of cases edited for every variable (102). Mobile homes, single family detached homes, and single family attached homes had very few cases edited due to the RAS, as seen in Table 2 below.

Table 2: Summary of Editing Effort by Questionnaire Item and Type of Housing Unit

	Type of Housing Unit				
Questionnaire Item	**Mobile Homes (n=5)**	**Single Family Attached Homes (n=21)**	**Single Family Detached Homes (n=39)**	**Apartments, Buildings with 2-4 Units (n=102)**	**Apartments, Building with 5+ Units (n=417)**
Age of water heating equipment	2	12	23	63	273
Water heating fuel	1	3	14	47	284
Space heating equipment age	1	10	20	62	252
Year housing unit was built	4	5	10	56	186
Age of refrigerator	3	13	8	49	142
Space heating fuel	0	3	14	31	160
Type of space heating equipment	0	2	8	35	139
Age of central air conditioning unit	0	5	9	17	84
Main cooking fuel	1	3	7	12	20
Type of air conditioning equipment	0	0	3	6	31
Number of separate ovens in the household	0	0	0	0	2
Number of separate stove tops in the household	0	0	0	0	2
Number of combination stove/oven appliances in the household	0	0	0	0	1

4.2 Effect on Item Nonresponse

The RAS greatly reduced item nonresponse in five variables: water heating fuel, age of water heating equipment, year the housing unit was built, age of space heating equipment, and space heating fuel. “Don’t Know” or “Refuse” responses given by the householder were replaced by a RAS response, and therefore did not have to be imputed, in 288 of the cases for water heating fuel, 260 of the cases for age of water heating

equipment, 188 of the cases for year the housing unit was built, 178 of the cases for age of space heating equipment, and 90 of the cases for space heating fuel. The reduction of item nonresponse was minimal in each of the remaining items compared, less than 10% of the 584 total cases.

Natural gas was the type of fuel most often replaced by a "Don't Know" or "Refuse" answer for both space heating and water heating. Of the 90 cases where the RAS filled in a value for space heating fuel, 61 (67%) were changed to natural gas; of the 288 cases where the RAS filled in a value for water heating fuel, 222 (77%) were changed to natural gas. This suggests a systematic underestimation of natural gas as a fuel for these end uses by respondents, which could potentially lead to bias in the estimates if not adjusted for, either with the RAS or another method.

The RAS effect on item nonresponse varies by type of housing unit, with the greatest effect on the apartments in buildings with 5 or more units. Table 3 shows the number of cases by type of housing unit and questionnaire item that were changed from "Don't Know" or "Refuse" to a value due to the RAS.

Table 3: RAS Effect on Item Nonresponse by Type of Housing Unit

	Type of Housing Unit				
Questionnaire Item	**Mobile Homes (n=5)**	**Single Family Detached Homes (n=21)**	**Single Family Attached Homes (n=39)**	**Apartments, Buildings with 2-4 Units (n=102)**	**Apartments, Building with 5+ Units (n=417)**
Water heating fuel	0	0	9	30	249
Age of water heating equipment	0	2	8	31	219
Year housing unit was built	3	0	6	41	138
Age of space heating equipment	1	2	8	29	138
Space heating fuel	0	1	2	8	79
Age of central air conditioning unit	0	1	2	5	46
Age of refrigerator	0	0	0	9	41
Type of space heating equipment	0	0	1	4	28
Main cooking fuel	0	0	0	1	0
Type of air conditioning equipment	0	0	0	0	0
Number of separate ovens in the household	0	0	0	0	0
Number of separate stove tops in the household	0	0	0	0	0
Number of combination stove/oven appliances in the household	0	0	0	0	0

4.3 Reduction of Measurement Error

The RAS reduced measurement error incurred during the response process by replacing the original householder response with a different RAS response in:

- 167 cases for age of heating equipment
- 156 cases for year the housing unit was built
- 151 cases for type of space heating equipment
- 118 cases for space heating fuel
- 113 cases for water heating equipment age
- 61 cases for water heating fuel

Household respondents overestimated electricity and underestimated natural gas as their heating and water heating fuel. Electricity is a more widely known and used fuel than natural gas, and perhaps respondents tended to select a response option more familiar to them. Out of the 118 cases where heating fuel was changed, 66% were changed from another fuel (mostly electricity) to natural gas. Out of the 61 cases where water heating fuel was changed, 65% were changed from another fuel (again, mostly electricity) to natural gas.

As was the case with item nonresponse, apartments (in 2-4 unit buildings and 5 or more unit buildings) were more prone to measurement error than mobile homes and single family homes; 87% of the changes to space heating fuel and 61% of the changes to water heating fuel were in apartments.

4.4 Effect on Weighted Estimates

Weighted estimates from the final edited household dataset were compared to weighted estimates from the raw dataset to determine the RAS impact on key estimates published for the RECS survey. The difference for all comparisons is less than 1%, indicating that the RAS had very little effect on overall nationwide estimates. Water heating fuel proportions are impacted the most, but the differences are not practically significant. See Figure 2 below.

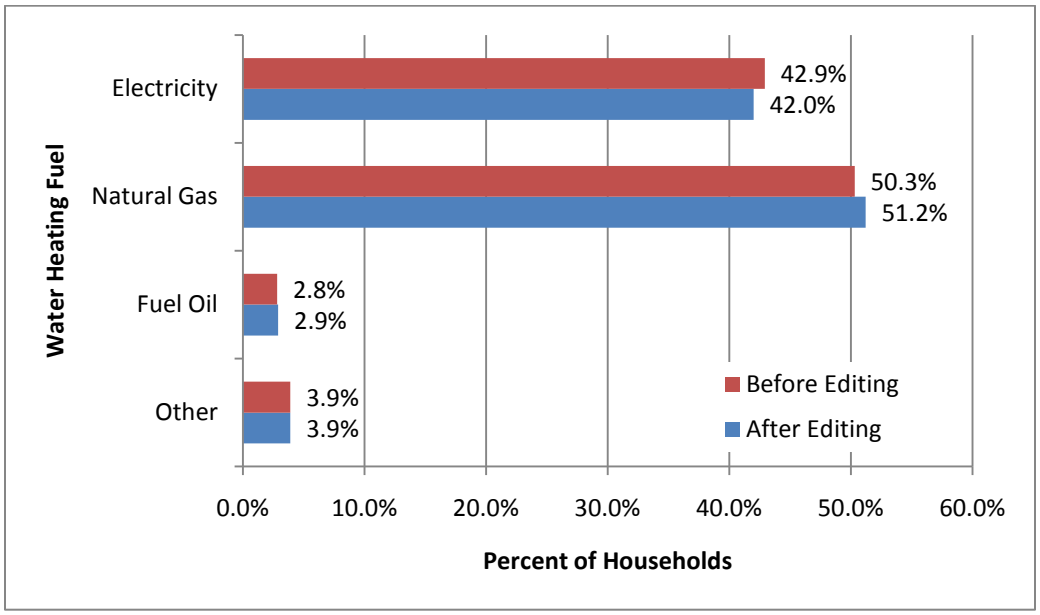

Figure 2: Water Heating Fuel Estimates Before and After RAS Editing

However, when the estimates are broken down by type of housing unit, the differences are significant for some housing types. In mobile homes, single family detached homes, and single family attached homes, the space heating fuel and water heating fuel proportion differences were less than .5%. Changes in the estimates for apartments in 2-4 unit buildings and apartments in 5 or more unit buildings were notably larger. Table 4 shows the estimates for space heating fuel and water heating fuel before and after RAS editing. The percentage of apartments in buildings with 5 or more units that use natural gas for water heating increased 7.9% due to the RAS; the percentage that use electricity decreased 9.5%.

Table 4: Comparison of Space Heating and Water Heating Fuel Estimates Before and After Editing in Apartments

	Apartments in Buildings with 2-4 Units			Apartments in Buildings with 5+ Units		
	% Before RAS Editing	% After RAS Editing	Difference	% Before RAS Editing	% After RAS Editing	Difference
Space Heating Fuel						
Electricity	41.4	38.7	-2.7	56.4	52.0	-4.4
Natural Gas	51.1	53.0	2.0	37.2	41.1	4.0
Fuel Oil	6.5	6.7	0.2	4.7	5.3	0.6
Other	1.0	1.6	0.5	1.7	1.6	-0.1
Water Heating Fuel						
Electricity	40.6	38.4	-2.2	60.4	50.9	-9.5
Natural Gas	55.2	57.1	1.9	38.1	45.9	7.9
Fuel Oil	2.8	2.9	0.1	1.1	2.1	1.0
Other	1.4	1.6	0.2	0.4	1.1	0.7

Differences in weighted estimates by other important domains due to RAS editing were examined, and most were found to be insignificant. Two notable exceptions were water heating fuel estimates for *all* renters (RAS eligible households, and households that pay their own energy bills), and space heating and water heating fuel estimates for households in the New England states. The RAS increased natural gas for water heating fuel for all renters 3.3%. In the New England states, the percentage of households that use natural gas increased 4.6% and 6.0% for space heating and water heating, respectively.

5. Discussion and Recommendations

The RAS was effective in reducing item nonresponse and measurement error in households that do not pay their energy bills directly in the 2009 RECS, especially in apartments. While the overall nationwide weighted estimates were not affected by the RAS, the estimates for some domains – apartments, all renters, and the New England states – would have been significantly different for some variables. Specifically, natural gas for space heating and water heating fuel would have been underestimated and electricity overestimated for these end uses.

Considering how resource intensive the RAS is in both field operations and editing effort, it is in EIA's interest to make it as efficient as possible. The results of this study suggest

that perhaps a RAS interview should be attempted for apartments only. Apartment renters were more prone to item nonresponse and measurement error, and the estimates for apartments were significantly affected by editing with the RAS data. The estimates for mobile homes and single family homes were unchanged by the RAS. Response rates are higher for apartments than mobile homes and single family detached homes, and RAS interviews for apartments are more efficient and cost-effective because the interviewer attempts to do a CAPI interview immediately following the household interview if the rental agent is on-site. Many large apartment buildings have a leasing office in the same building or a neighboring building of the sampled unit. Rental agents of mobile homes and single family detached homes are usually not on-site, and are sometimes in different states. These households require more contacts and field work, increasing field costs.

Another suggestion to make the RAS more efficient is to remove the "cooking" questions. Nonresponse and measurement error rates were very low for these items and the estimates were unchanged by the RAS editing. There is a high fixed cost to getting a RAS respondent to agree to an interview and the marginal cost to ask these questions is minimal, but it could trim several minutes off the RAS interview time.

Respondents overestimate electricity as their fuel for heating and/or water heating likely because electricity is the most commonly used fuel in U.S. households, and therefore, the fuel respondents are most familiar with. If they do not know what the fuel is, they may pick electricity out of familiarity. Patterns of data edit failures suggest that this source of error may not be limited to renters who do not pay their own utility bills, but all renters in general. A possible method to test this assumption in future rounds of RECS would be to subsample renters who do pay their utility bills and attempt a RAS interview for these cases.

There are plenty of other avenues for future research on this topic. The imputation models could be applied to the RAS cases to determine how accurately they impute a missing response given by the household. If the imputation model performs well, it could possibly be used to impute variables for RAS cases in place of pursuing a RAS interview in a subset of cases. A nonresponse bias analysis of the RAS would be helpful in determining if nonresponding RAS cases are different than responding cases. If they are fundamentally different, nonresponse adjustments can be explored.

Other survey organizations that do "piggyback" surveys can apply this study's methodology to evaluate how well they reduce item nonresponse and measurement error, and to potentially suggest improvements, such as fielding the piggyback survey only to respondents where item nonresponse and/or measurement error by the respondent could bias the estimates, or eliminating questionnaire items where data quality is not improved by asking the second respondent. Researchers interested in collecting residential energy data from the householders should be aware that people who do not pay their own energy bills, specifically apartment dwellers, tend to overestimate electricity and underestimate natural gas as their space heating and water heating fuels.

References

Fowler, F.J.Jr. (2002). *Survey Research Methods*. 3rd Ed. Thousand Oaks, CA: Sage Publications.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

Groves, R. M. & Couper M. P. (1998). *Nonresponse in Household Interview Surveys*. Hoboken, New Jersey: John Wiley & Sons.

Kalton, G. (1983). *Compensating for Missing Survey Data*. Ann Arbor, MI: The University of Michigan.

Saris, W. & Gallhofer, I. (2007). *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, New Jersey: John Wiley & Sons.

Tourangeau, R., & Smith, A. W. (1985). "Finding subgroups for surveys." *Public Opinion Quarterly*, 49, 351-365.

Tourangeau, R. & Rips, L. J. & Rasinski, K. (2000). *The Psychology of Survey Response*. U.K.: Cambridge University Press.