

The Pennsylvania State University

The Graduate School

Department of Sociology

DON'T KNOW RESPONSES IN SURVEY RESEARCH

A Dissertation in

Sociology and Demography

by

Rebekah Young

© 2012 Rebekah Young

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2012

The dissertation of Rebekah Young was reviewed and approved* by the following:

David R. Johnson
Professor of Sociology, Demography, and Human Development and Family Studies
Dissertation Advisor
Chair of Committee

Paul Amato
Arnold and Bette Hoffman Professor of Family Sociology and Demography

Alan Booth
Distinguished Professor of Sociology, Human Development, and Demography

Melissa Hardy
Distinguished Professor of Sociology, Demography, and Human Development and Family Studies

Eric Plutzer
Academic Director, Survey Research Center
Professor of Political Science and Sociology

Barrett Lee
Professor of Sociology and Demography
Sociology Graduate Officer

* Signatures are on file in the Graduate School

Abstract

Survey respondents are occasionally unable to generate the type of response researchers hope to record. Respondents may, instead, offer non-substantive answers such as saying they are unsure, cannot recall, or don't know. Understanding these responses is fundamental for optimal survey design and for identifying sensible methods to handle these responses in data analysis.

Unfortunately, most research about non-substantive responses has been limited to cross-sectional data analysis from non-representative samples and has examined political topics only. This is a serious limitation because researchers are making decisions about survey design and data analysis based on incomplete information. In a series of three papers, this dissertation generalizes what is known about "don't know" (DK) responses, explores the meaning of these responses by analyzing longitudinal data, and evaluates methods for treating these responses in data analysis.

Since the majority of research on DK responses has occurred in the context of questions related to politics, extending this literature to general survey data is an important contribution. Here, twelve nationally representative datasets were combined to analyze individual and question characteristics that predict DK responses. A total of 334 questions were analyzed from questions asked of 24,915 combined respondents. Men, people with greater education, and those who identified as being white were less likely than other respondents to give DK as a response. Challenging the conventional belief that women are always more likely to say DK, men, in fact, were more likely than women to say DK to questions with topics relating to family and friends, health and healthcare, or religion and spirituality.

Looking at DK responses over time, I found that respondents increased their use of DK responses within a single survey, consistent with the belief that some DK responses are the result of survey satisficing. The tendency for a respondent to offer DK replies was fairly consistent between surveys, evidence that these responses could also reflect a stable personality trait. Forty percent of participants who said DK to a question in the first survey said DK again when asked the same question 9-10 years later, providing a compelling case that at least some DK responses reflect genuine lack of knowledge. Finally, DK responses during the first survey were highly predictive of panel attrition, even when controlling other characteristics known to predict survey drop-out. These findings suggest that including a DK or 'no opinion' option in a survey design may be worthwhile since identifying likely future study dropouts is an important step in improving panel data response rates.

A number of methods are available for handling DK responses, but extant techniques have only had minimal empirical evaluation. I evaluated several common strategies for handling DK responses in scale items. These methods included: complete case analysis, coding the DKs as "neutral" or "middle" categories, coding all DK responses in one direction (for example, as all "no" responses), available item averaging, and multiple imputation. These methods were tested under four simulated meanings underlying DK responses: survey satisficing, passive refusal, valid response, and a random answer style. When DKs were the result of satisficing or a random answer style, and even when they were valid responses, multiple imputation was found to be the best solution. DKs that were passive refusals were best treated with directional coding, although this method required bold assumptions that are largely untestable with observed data.

Table of Contents

List of Figures	vi
List of Tables	vii
Acknowledgements	viii
Chapter 1 Don't Know Responses and Survey Research	1
The Historical Context of Don't Know Responses	3
The Present Study	7
Chapter 2 Men Don't Know Everything: Gender Differences in Don't Know Responses	8
Background.....	9
Individual Demographics and DK Responses	9
Question Characteristics and DK Responses	12
Data and Method	16
Results	22
Conclusion	27
Chapter 3 "I Still Don't Know": Non-substantive Responses in Longitudinal Data	33
Background	33
Data and Method	37
Results	42
Conclusion	46
Chapter 4 Strategies for Analyzing Data from Respondents who "Don't Know"	55
Background	56
Methods for Treating DK Responses to Scale Items	56
Complete Case Analysis	57
Neutral/Middle Category Coding	57
Directional Coding	58
Available Item Averaging	58
Dummy Variable Adjustment	59
Modern Methods	59
Theories about the Underlying Meaning of DK Responses	60
Response Style	61
Satisficing	61
Valid Response	62
Passive Refusal	62

Data and Method	63
Simulated DK Responses	64
Response Style	64
Satisficing	65
Valid Response	65
Passive Refusal	66
Combined Scale Items	66
Results	67
When DKs were a Response Style	69
When DKs were Satisficing	70
When DKs were a Valid Response	72
When DKs were a Passive Refusal	73
Conclusion	74
Chapter 5 What We Now Know About Don't Know Responses	82
References	85
Appendix A Question Wording for Individual-level Characteristics	93
Appendix B Question Wording from 12 Surveys	97

List of Figure

Figure 1 Pattern of DK Responses to 40 Questions Asked 9-10 Years Apart.....	52
--	----

List of Tables

Table 1 Data Source Descriptions	29
Table 2 Number of Questions by Topic and Survey Source	30
Table 3 Question Characteristics and Gender Differences in Don't Know Response Rates	31
Table 4 Random-Intercept Logistic Regression Models Predicting Don't Know Responses	32
Table 5 Linear Regression Model Predicting Proportion of DK Responses	50
Table 6 Linear Regression Model Predicting Proportion of DK Responses at Second Survey ...	51
Table 7 Logistic Regression Model Using DK Response Tendency to Predict Panel Attrition ...	53
Table 8 Logistic Regression Model Using Refusals to Predict Panel Attrition	54
Table 9 Observed Scale Item Frequencies	76
Table 10 Observed Scale Items Combined Using 6 Strategies for Handling DK Responses	77
Table 11 Simulated Scale Item Frequencies	78
Table 12 Simulated Scale Items Combined Using 6 Strategies for Handling DK Responses	79
Table 13 Linear Regression Results for Models Using Simulated Scale Items	81

Acknowledgments

I owe this dissertation to a curious person who once asked me if it made sense to impute “don’t know” responses. Although I had been working to become an expert on missing data, it was not clear to me whether these responses were “missing” at all. It seemed likely that don’t know replies to survey questions could sometimes be legitimate and meaningful answers; after all, I could only answer this question with uncertainty myself. Since I could not impute an answer, I wrote a dissertation in search of one. Who was this influential person – the person who sparked my obsession with don’t know responses and instigated thousands of hours of research? I don’t know. I no longer remember. Whoever you are, I thank you for asking me that question because I have had an enormous amount of fun trying to answer it.

It would have been impossible for me to complete this dissertation without the support of many colleagues, students, teachers, and friends who assisted, advised, and critiqued my research and writing efforts over the years. Especially, I want to express gratitude and deep appreciation to my adviser, David R. Johnson. He has been an outstanding mentor, friend, and role model. I would have never survived graduate school without his constant encouragement and unwavering confidence that I would succeed (despite giving me the only “B” I ever received on an exam). David gave me the opportunity to write the most exciting dissertation I could imagine, even though he eventually confessed that he first thought it was a “weird little topic.” David is one of the smartest and funniest people I know. I have enjoyed every minute of working with him and I could not have asked for a better advisor. I also am indebted to Laurie Scheuble for letting David work on my projects instead of hers, for many hours of listening to missing data talk (even on vacation), and for getting really excited about my TCRM award.

My dissertation committee was a marvelous group. Thank you to Paul Amato for embracing my weird little topic with such genuine interest and enthusiasm. I aspire to be as responsive and encouraging to my students as he was to me. He set the standard quite high for responding to emails quickly; I’ll never forget how much it motivated me to see thoughtful comments return from him within minutes. Alan Booth is one of the most caring people I have ever met and I am so thrilled that he agreed to join my committee. I appreciate Alan’s abundant moral support and life advice. One of my favorite moments with Alan occurred when I nervously sent out a tentative draft of my dissertation proposal. His email back to me read simply, “Your proposal is fine. I will read it later.” A copy of his email hangs above my desk today as a reminder that if such a brilliant man can have such faith in my work, then I should have confidence in myself. I thank Melissa Hardy, for all the advice, and especially the laughter, that guided me through this process. Thank you to Eric Plutzer for the tough questions, the clever ideas, and the fantastic and constructive criticism that improved the quality of my work.

Eric Silver deserves special thanks for serving as my secret mentor all these years. Everybody needs a rational person in their life who they can approach in confidence during their craziest moments. Eric was that person for me and his door was always open. From having a suicidal student in my class, to my panicked realization that I was working on 17 projects at once, he heard my troubles without judgment or criticism. With patience and clarity, he helped me solve problems that would have otherwise prevented me from having the energy to write this dissertation.

Thank you to everyone who supported my passion for missing data. I appreciate Michelle Frisco and Valarie King, who allowed me to write missing data papers in their health disparities and family sociology courses. I thank Jeremy Staff for being an excellent role model who always reminded me to stay productive. Thank you to all of my students for doing the ‘missing data wave’ with so much enthusiasm (even Jacob and Brett, eventually). Thank you to all my fellow graduate students who tolerated, frequently against their will, endless missing data presentations in class and many loud conversations about it in the lab.

So many of my friends kept me sane during this process and it is a pleasure to thank them here. Patty Gwartney was an extraordinary mentor who understood where I came from and helped me learn to assimilate in the ivory tower. Claire Altman spent thousands of hours reading my work, meeting weekly to keep my writing on schedule, and stopping by to check on me. I thank Shannon Fleishman for keeping me grounded in the real world and for her many hours spent pursuing our CATA bus epiphany. Thank you to Li-Chen Lin for reminding me about life’s priorities, for teaching me about her culture, and for her heartfelt encouragement. I thank Seung-won Choi for making me feel smart, for our chats in the car, and for being one of the sweetest women in the world. Thank you to Cat Tucker for being on my side no matter what the issue was. I appreciate Adriana Reyes for having a positive outlook on life and for keeping me young. Thank you to Veronica Roth for being a survey methodologist with an unbridled zest for Stata graphing and “celebrity stalking” at AAPOR. Thank you to Laura Rosell for teaching me that a word can be as pretty as a number and that it is possible to type fearlessly. I especially appreciate Laura’s comments on many drafts of this dissertation. Jennifer Pearce-Morris really deserves a whole chapter of thanks here as we spent nearly every day of graduate school together. This process would have been unmanageable without her love, acceptance, and tireless support.

Thanks are also due to my loving and supportive family who let me move 3,000 miles away from home to study “something that isn’t there.” I missed birthdays, holidays, and special events, all to pursue becoming “Dr. Pickles.” I’m so thankful to have such a wonderful mother who values education as much as I do. I did not always love school, and there were many times I tried to give it up. As a kindergartener, I tried to convince my mom that I could live a productive life without ever knowing how to read. In second grade, I was sure that rounding numbers would never be useful and was a waste of energy to learn. Mom, you were right, I am glad that I learned to read and round. (Ironically, David and I wrote a paper on methods for rounding imputed values.) My sister Sarah is probably my biggest fan. Thanks for taking red-eye flights to see me every Thanksgiving, for laughing at academic norms, and for listening when I needed to blow off steam. My youngest sister, Deborah, will probably skip straight to reading this line. Thank you for being hilarious, for thinking that I know more than Google, and for letting me miss your graduation. Thanks for visiting me during the horrible February weather and for spending such a special summer with me in Washington, D.C.

Last, but never least, I thank Tim Sanders – a man who gives a “don’t know” response nine out of ten times I ask him a question. Tim, I’m sorry to tell you that my dissertation concludes that imputing “don’t know” responses is not always appropriate; therefore, you can expect me to continue pestering you to give something other than a ‘don’t know’ reply for many years to come. With all my love and thanks, this dissertation is dedicated to you.

Chapter 1

Don't Know Responses and Survey Research

Gathering data from people is a foundation of social science research. Social scientists ask people questions and use their answers to draw conclusions about the state of society.

Occasionally, respondents are unable to generate the type of response survey researchers hope to record. Many surveys include a “don't know” (DK) option to ensure answers to all questions.

Those survey participants who offer DK replies are the subject of this dissertation. Who these people are and how we handle their responses has enormous implications for the conclusions we draw from their data and how we should collect information from them in the future. Increased attention should be given to developing a better understanding of DK responses in multiple topic areas, across surveys, and over time.

The first article of this dissertation explores characteristics of people who provide DK responses and the types of questions that tend to elicit more of these responses. Prior research on this topic has been limited by data that were non-probability samples or contained only politically-related topic domains. As a result, these findings may be inconsistent with other types of survey data, particularly with respect to respondent gender. For example, survey researchers believe that women offer DK replies more frequently than men. While this is known for questions about politics, women could actually be less likely to say DK to questions about other topics, such as family, friends, or religious views. Understanding how individual and question factors interact, and how this relates to DK responses, is an important step in toward improving survey questions and optimizing the information that respondents provide.

The second article of this dissertation examines DK responses in longitudinal data. The stability and predictive power of DK responses from one survey to the next may have important

implications for survey design, methods for reducing item nonresponse, and predicting future panel attrition. Uncorrelated DK responses over time may indicate that contextual survey factors tend to elicit these responses, rather than reflecting a stable response style that people carry from one survey to the next. DK responses correlated over time may suggest the presence of a stable group of non-responders, a sign that manipulation of the survey design is unlikely to eliminate this type of item nonresponse. Still another reason to explore DK responses in longitudinal data is that DK responses in the initial wave of a survey may be predictive of later panel attrition. If this is the case, including a DK or ‘no opinion’ option in a survey design may be worthwhile since identifying likely future study dropouts is an important step in improving response rates in panel data.

The final article of this dissertation is intended primarily for researchers who use secondary data and find DK responses to be a dilemma in statistical analysis. Many researchers analyze data that has already been collected and therefore have no control over whether a DK response option was offered to study participants. When DKs are present in secondary data, few guidelines exist to help researchers decide how to analyze these responses. Unfortunately, the most common practices for handling DK responses have received little empirical scrutiny. For example, DKs are routinely recoded to represent a “middle” or “neutral” category, although this method has never been advocated by methodologists. In this article I evaluate the following six common methods for handling DK responses: complete case analysis, neutral/middle category coding, directional recodes, available item averaging, dummy variable adjustment, and multiple imputation. Each of the six methods is evaluated under four simulated causes of DK responses; the results suggest that some DKs are relatively easy to handle in statistical analysis, but that others can be quite problematic and researchers should be wary of these.

THE HISTORICAL CONTEXT OF DON'T KNOW RESPONSES

'Don't know' responses are a small component of the broader field of survey methodology. Survey research generally adheres to five fundamental principles: define the population of interest, make a sampling frame of all population members, use probability methods to select the sample, carefully design and pre-test the questionnaire to be used with an identified sample, and attempt to minimize error (Couper & Nicholls, 1998). Reducing survey error, the last principle, has led many researchers to view DK responses as problematic for survey data. DK responses are often regarded as instances of two types of error: either as invalid responses or as nonresponse. A historical perspective on how DK responses have come to be perceived as a problem is useful when thinking about whether or not they are a part of survey error, and if so, what can be done about them.

Nearly half of a century ago, Converse (1964) wrote a book chapter that raised concern about whether many of the attitudes measured by researchers are actually "non-attitudes." He postulated that while many people may not have thought through the nuances of an issue, they will provide opinions to survey researchers nonetheless. Rather than browbeating respondents into offering nonexistent opinions, Converse (1970) suggested, researchers should instead invite them to state that they either have no opinion or don't know the answer to a question. This advice provoked a series of debates among survey practitioners about how to design survey response categories that best reflect a respondent's condition.

Following Converse, many researchers demonstrated that not only did people provide opinions when they did not have one, people also expressed thoughts about entirely fictitious issues. Schuman and Presser (1980) found about 30% of the population gave opinions on a highly obscure bill before Congress when it was unlikely these respondents had heard of the bill.

To combat this problem, the authors suggested the inclusion of a “filter question” to screen the respondent out of substantive categories. For example, an interviewer might ask a respondent if s/he has ever thought about an issue prior to asking for an opinion. Another way to filter non-attitudes is the inclusion of an explicit ‘no opinion’ option among the given response categories. This strategy is advantageous because it does not require asking additional questions. Schuman and Presser (1980) found 10% of people offered opinions about the fictitious Congressional bill when a DK option was included explicitly in the answer choices. The result of this research was widespread inclusion of filter questions and explicit DK options in questionnaires and interviews.

Not long after survey researchers concluded that this issue had been resolved by asking respondents whether they had an opinion or not, caution was called for once again. Methodologists warned about the downside to minimizing nonattitude reporting; giving people a DK option could turn them into lazy participants. The belief was that survey respondents are less interested in providing responses than researchers are in receiving them. In fact, “satisficing” occurs when respondents settle for a satisfactory response rather than searching for an optimal response (Krosnick, 1991). Explicit inclusion of a DK option may encourage a respondent to give an unusable answer rather than search for a true answer (Krosnick et al., 2002). Of course, people do not carry fixed attitudes toward every issue, so DK may on occasion be the “correct” response. Many researchers have attempted to show forced-choice questions recording nonexistent attitudes will not affect the distribution of real responses (Krosnick, 2002; Mondak & Davis, 2001). This fashion of research prompted the removal of DK options from questionnaires and less frequent use of filter questions.

The field of survey research traveled full circle in its beliefs about best practices for allowing or discouraging DK responses. What has been learned from this journey is that regardless of the methods used to record or filter out non-attitudes, some type of error is certain to occur, even when DKs are legitimate responses for some respondents. There are primarily two ways that survey methodologists have tried to interpret the error from DK responses: as measurement error or as nonresponse error.

Measurement error in survey research typically pertains to deviation from answers to a given survey question and the underlying attribute being measured (Groves et al., 2004a). Two types of measurement error occur for DK responses. False-positives occur when the respondent has no opinion but expresses one anyway. Converse (1964) raised this issue; researchers now know if the DK option is not accepted or not made explicit, some non-attitudes will be incorrectly measured as attitudes. False-negatives occur when a respondent holds an opinion but declines to express it. Although these respondents may have many reasons for wishing to withhold an answer, satisficing (or laziness) appears to be survey researchers' main concern. Since measurement errors with DK responses take two forms, the way that survey design treats the DK option changes the risk of each error type's occurrence (Giljam & Granberg, 1993).

Regarding DK responses as measurement error implies the solution to the problem lies in the way questionnaires are designed and responses are recorded. This poses a serious dilemma. From asking respondents directly if they have no opinion, to not allowing DK responses whatsoever, survey researchers have rigorously tested questionnaire design options. This exhaustive body of research warrants one conclusion – no option is perfect. When DK responses are not allowed, and substantive answers are recorded in their place (false-positives), researchers cannot identify such occurrences with any certainty. These “incorrect” responses are identical to

accurate measurements. Unless wildly inconsistent with other answers given by a particular individual, such errors will go unnoticed and uncorrected. Rather than seeking only to reduce measurement error, then, perhaps we should also consider the extent to which we are able to identify error when it is inescapable. When allowed, it is difficult to know if DK answers are a correct response, a passive refusal, or the result of satisficing, but at least the troublesome answers can be identified. Recording the DK responses allows interpretation of the error as item-level missing data, or nonresponse error. Missing data are a familiar problem; false-positive measurement errors are a problem that cannot be identified.

The second way researchers have interpreted DK responses is as item nonresponse error. Item nonresponse is a type of missing data where many individual items for a respondent were measured successfully, but some were not (Groves, et al., 2004a). For example, income is a well-researched type of sensitive question eliciting higher than average levels of nonresponse, even from respondents who were otherwise happy to participate in the survey (Riphahn & Serfling, 2002). Under some circumstances, DK responses are like all item nonresponse where the respondent initiated a decision to withhold judgment about a particular question. DK responses are a traditional type of missing data when the true value is unknown but must be one of the given response categories (Rubin et al., 1995). For instance, many insured people are unaware of specific costs and procedures covered by their health insurance policy. When asked by an interviewer if a specific procedure is covered, some respondents will truly not know the best answer. A “true” answer, however, does exist for the respondent. An advantage of treating DK responses as a form of general item nonresponse is that modern techniques for handling missing data are widespread. A simultaneous disadvantage is that many of the missing data techniques used to treat item nonresponse have been given little empirical attention regarding DK responses

specifically. In particular, what are the consequences of treating DK responses as missing data when a true value does not necessarily exist for a respondent?

THE PRESENT STUDY

There are many questions left unanswered by the literature on DK responses. Survey practitioners are concerned about the best way to measure attitudes and have consequently engaged in much debate about whether an explicit DK option should be provided to survey respondents. Survey users are concerned with the best way to handle DK responses when they are recorded in data. Unfortunately, these decisions are being made with incomplete information. The majority of what is known about DK responses comes from data on political topics. These findings may not be consistent with other survey domains.

In this dissertation I address the following three research questions related to DK responses:

1. Who says DK, and under what circumstances?
2. What can we learn about the meaning of DK responses by analyzing panel data?
3. How should DK responses be handled in statistical analysis?

Answering these questions will begin to generalize current scholarship on DK responses, help survey creators optimize use of the DK option, and suggest solutions for analyzing these nebulous responses.

Chapter 2

Men Don't Know Everything: Gender Differences in Don't Know Responses

From survey designers debating the inclusion of filter questions at the design stage, to secondary data users grappling with non-substantive responses in statistical analysis, the “don't know” (DK) response option is a survey research dilemma. Addressing this problem requires an understanding of the relationship between non-substantive response and individual demographics, question characteristics, and survey context. Unfortunately, much of what is known about DK responses has not been generalized across topical domains from nationally representative data. Because almost all research has dealt with political topics, findings may not be easily generalized to other types of survey data, particularly with respect to gender differences. Furthermore, those datasets from which researchers have analyzed DK responses outside of political contexts have tended not to be nationally representative. These limitations suggest that what we “know” about DK responses may not be as well understood as the literature suggests.

Understanding who gives DK responses, and in what context, is fundamental for optimal survey design and for identifying sensible methods to handle these responses in data analysis. In statistical analysis, for example, we need to understand these responses before we know if they represent “missing data” or whether the DK is a legitimate response category. From a survey design perspective, a substantive understanding of DK responses is important for deciding whether these responses should be allowed, and, if so, whether the option should be made explicit. Exploring who offers DK responses and to what types of questions will help researchers make sense of non-substantive responses and greatly inform future methodology.

This paper expands research on DK responses by analyzing 334 questions from 12 sources of nationally representative data. Multilevel models were used to identify individual-level and question-level predictors of DK responses. An exploration of individual and question factors is consistent with recent research which highlights the importance of understanding how these factors interact in the assurance of survey quality (Couper, 2008; Wagner, 2010). This paper specifically focuses on how a respondent's gender interacts with question characteristics in predicting the probability of a DK response. Our findings challenge the conventional belief that women are always more likely to offer DK responses to survey questions. With a better understanding of who uses the DK option, and under what circumstances, survey researchers will be able to optimize the use of this option in questionnaire design and analysis.

BACKGROUND

DK responses are known to be related to individual characteristics and to features of the survey context, particularly to the specific questions asked. Regarding demographic characteristics of individuals, gender, education, race, and age are all strong predictors of DK responses. Question characteristics, such as sensitivity, difficulty, and topic are also known to influence the occurrence of DKs. The interaction between individual and question characteristics in producing DK responses, however, is a less-understood relationship. This paper helps disentangle these complex interactions by focusing specifically on gender's influence across a variety of question characteristics.

Individual Demographics and DK Responses

The survey literature largely agrees upon the relationship between a respondent's gender and his or her propensity for a DK response; female respondents' data are more likely to feature item nonresponse, and women are up to twice as likely as men to give DK responses (Francis &

Busch, 1975; Mondak & Anderson, 2004; Van Es et al., 1996). These findings are limited though because the majority of research on DK responses is based on political research; conclusions about DK responses to political questions may not be easily generalized to other types of data. Similarly, before relying on this literature to make gendered inferences about DK responses, it is important to recognize that politics has historically been a male endeavor and therefore may be a less compelling topic for women. With other topics, in fact, such as nutrition, females have been found to be no more likely than males to give a DK reply (Durand & Lambert, 1988). It bears noting, however, that these samples were non-representative, so the question remains: do women have higher DK response rates for all types of survey questions? To answer this question, it is necessary to account for other demographic characteristics related to DK responses.

Education, in particular, appears to have a consistent, negative relationship with the likelihood of DK responses (Jean M. Converse, 1976; Durand & Lambert, 1988; Riphahn & Serfling, 2002). Two complementary theories explain this trend. People with lower levels of education may be more likely to indicate DK because they are less informed and have less knowledge in general (Francis & Busch, 1975). At the same time, highly educated respondents may resist admitting to a lack of knowledge and therefore refrain from using DK as an option (Ziller & Long, 1965). The latter type of individual may perceive a DK response as indicative of personal shortcomings; to maintain status, therefore, he or she might deliver a substantive response, however meaningless or incorrect (Ziller & Long, 1965). What remains unclear in the literature is whether the effect of education on DK responses is consistent across genders.

Race may be related to DK response propensity as well, independent of race-related educational differences. Research on response styles has shown that different racial groups have distinct patterns of answering questions. For example, Bachman and O'Malley (1984) found

Black adolescents to be twice as likely as white adolescents to respond ‘no opinion’ to a question about dishonesty in the presidency and administration. Similarly, although studies in the U.S. context have not fully explored whether Latino ethnicity uniquely influences DK response patterns, some research from abroad suggests that this might be possible as well. For instance, Sicinski (1970), using nationally representative data for each country included, uncovered national-origin differences in a European study; in Sicinski’s study, French respondents, compared to Polish respondents, were both less likely to answer “DK” in a series of knowledge questions and to answer questions incorrectly. This finding suggests that culture might influence an individual’s willingness to acknowledge lack of opinion or lack of awareness on a given topic, at least within the context of a survey. Aside from DK responses, we do know that outright refusals are strongly predicted by ethnicity in the U.S. context, although the relationship with general item nonresponse (including DKs) is unclear (Peytchev et al., 2009). Here again, however, just as with education, it is still unknown whether race and ethnicity interact with gender in influencing an individual’s likelihood of declining to provide a substantive answer.

Age is a robust and positive predictor of DK responses, although few researchers have hypothesized about why this pattern occurs (DeMaio, 1980; Rapoport, 1985). In examining age, researchers have noted that older respondents are more likely to reply “DK” to survey questions. This is perplexing, on the surface, since older respondents have more life experience than younger ones and could, presumably, possess a broader wealth of knowledge. Indeed, if the relationship between age and DK responses is linear, it may indicate that DK responses are substitutes for passive refusals among older respondents. On the other hand, however, Bogart (1967) pointed out that opinions invariably transcend knowledge, so younger respondents might believe they know a great deal, while older respondents assess their own knowledge more

realistically. Ziller and Long (1965) found an interaction between age and gender wherein younger females and older males were the most likely to give a DK response, but the researchers offered no explanation for this finding. If withholding the DK response is a status-defense mechanism, it seems possible that younger females and older males feel less threatened by not knowing. No other research reports the existence of an interaction between sex and age on DK responses. The paucity of such literature may indicate that such interactions were not tested or, if tested, did not emerge as significant.

Question Characteristics and DK Responses

Survey researchers undoubtedly believe that question characteristics will influence the type of answers received, as well as that interviewers might give respondents clues as to the acceptability of a DK answer. Whether or not DK responses are actually accepted by researchers is a crucial element in how uncertainty is recorded. DK responses may be handled in any of the following ways: an option to say DK may be offered explicitly, a DK response may be accepted if the respondent decides to offer it, interviewers may probe if a DK is given, or a DK response may not be allowed at all (Beatty & Herrmann, 2002). A large body of research has explored these options, investigating which are preferable for survey research (Philip E. Converse, 1970; Giljam & Granberg, 1993; Sanchez & Morchio, 1992; Schuman & Presser, 1981).

When DK responses are allowed, many additional features of the questions asked will further influence a respondent's willingness and ability to generate a substantive response. Perhaps the most important of these question characteristics is the topic itself. Question content affects survey nonresponse in general and is correlated with the frequency of DK responses for questions on various topics (Durand & Lambert, 1988; Groves et al., 2004b; Schuman & Presser, 1980). If a topic is particularly remote or obscure, respondents truly may lack the knowledge

necessary for answering questions in an otherwise meaningful way (Feick, 1989). When respondents lack interest in an issue, find the subject to be of low personal importance, or have little affective involvement with a topic, nonresponse is higher as well (Krosnick, et al., 2002).

Gender also influences a respondent's engagement with a given topic; research has shown clear gender differences for a wide range of interests, activities, values, and personal priorities. For example, in surveys, men tend to display more interest than women in politics, sex, and sporting events, and less interest in topics such as health and nutrition (Ares & Gambaro, 2007; Baldwin & Baldwin, 1997; Mondak & Anderson, 2004; Wenner, 1998). This paper does not address how topic-related gender differences might arise or why differences occur, but does bring to light a surprising gap in our understanding of gender's potential to influence DK response patterns.

As with gender-topic differentials, question characteristics may not have the same patterns of DK responses for men and women. It is widely recognized that question sensitivity, in general, is especially important in predicting nonresponse; unsurprisingly, sensitive questions elicit higher levels of item nonresponse than those of a less personal nature (Berinsky, 1999). Although opinions differ on what constitutes a sensitive topic, content area clearly plays a role in determining the social difficulty with which any given question might be answered honestly. One common hypothesis for DK responses is that participants withhold sensitive information intentionally, rather than because they honestly do not know it (Beatty & Herrmann, 2002). This is consistent with the observation that DK responses are generally more frequent for questions about HIV status, sexual activity, income, and receipt of government assistance (Tourangeau & Smith, 1996; Wright et al., 1998). Respondents might also believe that their attitudes are socially unacceptable and wish to hide their true answer without going so far as to give a false one

(DeMaio, 1980). A DK response may mask, for example, unpopular political opinions or racist ideology (Berinsky, 1999; DeMaio, 1980). And, of course, gender plays a role here again; researchers largely agree that women have higher item nonresponse rates to sensitive questions (Catania et al., 1996; Tourangeau & Smith, 1996). Thus, a control for question sensitivity is important when analyzing the influence of question topic.

Recording responses in the same metric is a necessary step for comparability between respondents and consistency with other surveys. Many surveys include pre-defined answer categories rather than allowing open-ended answers. One downside to this approach is that some respondents may find the response categories too restrictive, particularly when few categories are provided. Converse (1976) showed that respondents have difficulty selecting between dichotomous response options (e.g., yes/no, agree/disagree) and that 'no opinion' responses are higher for dichotomous choice responses than for polytomous ones. Finer-graded measurement is generally believed to elicit lower levels of item nonresponse because it gives participants more options for revealing the information they believe is most accurate; alternatively, it provides them with an opportunity to describe neutral opinions, rather than hemming them into reporting a lack thereof (Jean M. Converse, 1976). Whether men and women have similar patterns of DK responses based on the response metric is unknown.

Questions that require a long view into the future or predictions about future events are generally believed to elicit higher nonresponse also (Jean M. Converse, 1976; Feick, 1989). Of course, no respondent actually knows what will happen in the future, but some people are probably more comfortable conjecturing than others. For example, Shmotkin (1992) analyzed a question about how well respondents expected their subjective well-being to be 5 years in the future. The item nonresponse rate was less than 1% for people aged 20-59 but nearly 20% for

those aged 60 or older (Shmotkin, 1992). Since these data came from a convenience sample, the finding is limited, but such a large differential warrants further investigation into this question characteristic.

The type of information that must be generated to answer a question influences its cognitive demand. Researchers generally believe that this is why questions requiring particular knowledge or factual information receive more DK responses than those which solicit an attitude or opinion (Beatty & Herrmann, 2002; Jean M. Converse, 1976). This question characteristic has been investigated extensively because researchers worry that respondents may legitimately not know an answer and should not be forced into giving one (Poe et al., 1988; Sanchez & Morchio, 1992). Empirical evidence has demonstrated that respondents are willing to offer attitudes and opinions that they do not hold (on obscure topics, for instance), but this answering strategy may be less likely—if not also more difficult—with knowledge/factual questions. Gender, again, makes a difference; when answering knowledge questions, women generally have higher DK response rates than men and are more likely believe they know less (Linn et al., 1987). All in all, when controlling for question topic, it is important to account for whether a given question demands objective or subjective information.

Questions that occur later in the survey are generally believed to elicit higher item nonresponse than earlier items (Culpepper et al., 1992). As a survey drags on, respondents lose the motivation to draw on their memories in the formulation of careful answers. We know that respondents can change answering styles within a single survey and that increasing the frequency of DK responses may be an example of respondent satisficing (Beatty & Herrmann, 2002; Hui & Triandis, 1985). Though the length of an interview is not always proportional to its burden, fatigue may be measured by counting the position of each question (Groves & Couper, 1998). It

is relatively common to avoid asking certain questions, such as those about income and demographics, until the end of the survey, so it is necessary to account for this pattern when exploring DKs by topic.

DATA AND METHOD

This study combined 12 publicly available data sources which were designed to be nationally representative probability samples. A limitation of previous work on DK responses is that the data covered a limited range of questions or were drawn from non-representative samples. The choice of data sources for this study was guided by the desire to obtain a wide variety of question topics and to ensure that findings could be generalized beyond those resulting from any one particular survey. Table 1 contains a list of survey sources used, each selected according to three criteria: data collection occurred after 1998, the design employed probability-based sampling methods, and the survey contained questions from a minimum of four topic domains. All sources are available through public archives; information on sampling strategies and response rates can be found in the methodology reports provided by the data collectors. Data were analyzed from a combined 24,915 respondents.

While the information researchers collected from respondents varied widely across data sources, the collection of general demographic information was fairly consistent. All data sources reported the respondent's education level, race and Hispanic ethnicity, age, and gender. Each source's exact wording for these demographic questions is available in Appendix A. When all sources were combined, 56% of the respondents were female, and average respondent age was 49 years old. The level of measurement for age and gender was consistent across sources, but there were some problems with comparability of measures for race and education.

For education, some data sources collected finely-graded categories of educational attainment while others used rather crude measures. Take, for example, the “some college” category; some researchers collected an exact number of years of college attendance or noted the receipt of an Associate’s degree, while others asked only if the respondent had ever attended college. Furthermore, there was little consistency in how respondents with a GED were categorized with respect to high school graduates, as well as a general lack of standardization for measurements beyond receipt of a Bachelor’s degree. The most comparable categories for education were: less than high school (8.5% of respondents), high school graduate (27.2%), some college (30.5%), college degree (20.5%), and graduate education (13.3%). Some measurement error was unavoidable for this category. Depending on the data source, respondents with GEDs could have been categorized as either “less than high school” or “high school graduate.” Likewise, there was no way to distinguish whether respondents in the “graduate education” category had completed a graduate degree, so those who were ever enrolled in a post-graduate program were sometimes classified as having “graduate education” and other times as, simply, “college graduates.” The error for these categories did not appear to be systematic across data sources, so we believe that the error produced conservative estimates of the effect of education.

Race and ethnicity were collected quite differently from source to source. Some surveys marked “Hispanic” as a race category, while others treated it as a separate question about ethnicity. Race was occasionally measured in an open-ended response format or using a “check all that apply” format, producing finely-graded response categories. At the crudest level of measurement, some surveys collected only the categories Asian, Black, Hispanic, other, and White, making it impossible to identify other important groups. When data sources were

combined for this study, to reduce inappropriate re-classification, respondents who had chosen any category other than Asian, White, Black or Hispanic were coded as “Other/Mixed race.” This resulted in a substantial proportion of respondents falling into the latter category; exactly who this group of people represents is unclear. Nonetheless, this seemed preferable to other coding options because it ensured that the people in each remaining race/ethnic category had, in fact, self-identified with that group. When the data sources were combined, the race/ethnic category representation was 2.5% Asian, 9.9% Black, 8.4% Hispanic, 6.2% Other/Mixed race, and 73% White.

A total of 334 questions from across the 12 data sources were analyzed. Exact wording for all questions is available in Appendix B. Six question characteristics were identified for each of the 334 questions: question sensitivity, whether the response options were dichotomous or not, the time-orientation of the question (past/present vs. future), subjectivity (fact/knowledge vs. attitude/opinion), the question’s position in the survey, and the question topic. Questions were selected based on the criteria that at least two explicit DK responses had been recorded in the dataset and that the question was asked of the majority of respondents. When the data sources were combined, 23.9% of the response options were dichotomous choice, 42.1% were about facts or knowledge, and 5.5% asked for a judgment about the future.

Criteria for deeming survey questions as “sensitive” have not been clearly delineated in the literature. Tourangeau and Smith (1996) suggested that topics which are considered private, those which are considered “taboo,” and those that may elicit negative views or consequences may be viewed as sensitive. In general, questions herein identified as being sensitive dealt with drug or alcohol addiction, specific financial information, homosexuality, or prejudiced views of racial or ethnic groups. For example, the National Survey of Midlife Development in the United

States asks, “Was there ever a time in your life when you regularly had at least one drink three or more days a week?” (Brim et al., 1996a). Such a question was considered sensitive, on the grounds that it potentially captured an addiction.

Not all questions about a potentially sensitive topic, however, were considered sensitive in and of themselves. For example, the CBS News Monthly Poll asks, “Do you have any major credit cards?” (CBS News, 2009). Although the subject of this question deals with finances, a topic known to be sensitive, this particular question seemed benignly general and was not coded as sensitive. By contrast, if the question had asked for specific amounts owed on credit cards, the question would have been coded as sensitive. It is important to bear in mind that all data sources were publically available, so extremely sensitive topics, such as HIV status, were not recorded in these data. Overall, 8.3% of the questions were identified as sensitive.

A question’s position within a survey is another characteristic that may influence DK responses. Identifying position is typically straightforward but requires some explanation. One way to consider question position is to identify where each item occurred in a sequential count of all questions asked. Question position as a sequential count, however, is complicated when combining surveys of different lengths; the fiftieth question could mark the end of the interview in one survey but could appear in the introductory phase of another. Assuming that respondents begin a survey with an idea of how long it will take (such as described in the introduction for most of the surveys reviewed for this study), perhaps a more meaningful measure of question position is where the item occurred relative to the length of the entire survey. The surveys used here were divided into approximately equal fifths, and question position was identified by the quintile in which it occurred. When data sources were combined, 17.2% of questions occurred in

the very beginning, 16.1% occurred near the beginning, 21.5% occurred in the middle, 23.2% occurred toward the end, and 22% occurred at the very end of the survey.

Eleven mutually exclusive question topics were coded for all questions: activities and behaviors, demographics and employment, family and friends, finances and economics, health and healthcare, personality and feelings, politics and government, race and ethnic relations, religion and spirituality, specific policies, and war and terrorism. This scheme emerged from an iterative process based on two criteria: any potential data source must contain questions from at least four question topics, but each question topic must be covered by at least four data sources. This process was an important step toward findings that could be generalized beyond any single data collection effort. Many survey questions were impossible to categorize in a meaningful way and thus were not used. The National Firearms Survey, for example, asked the question, “On a different topic, some people believe they have seen UFOs or aliens. Have you personally ever been in contact with aliens from another planet?” (Hemenway, 1999). It is unclear whether this question is best categorized as an “activity or behavior” or some other category. Table 2 shows question topics present in each data source and the number of questions coded from each. An average of 28 questions was taken from each data source.

To combine the data sources in a manner that enabled analysis of whether a respondent offered DK responses to multiple questions, and multiple respondents from different surveys (containing different questions), the file was constructed with a person-question as the unit of analysis ($N = 680,734$). This strategy is consistent with previous research on DK responses and facilitates analysis with a multilevel model (Jean M. Converse, 1976; Rabe-Hesketh & Skrondal, 2008). Each person-question record contained the respondent’s answer to applicable questions, respondent characteristics, and question characteristics. The dependent variable was a

dichotomous variable that indicated whether the response was a DK or some other response (either a substantive response or a refusal; inapplicable responses were excluded). Stata 11.1 was used to fit a three-level random-intercept logistic regression model for DK responses from each gender with questions i nested within respondents j who are nested in surveys k . Using the latent-response formulation, the model can be written as

$$\begin{aligned}
Y_{ijk} = & \beta_1 + \beta_{2jk}(\text{Less than high school}) + \beta_{3jk}(\text{High school}) + \beta_{4jk}(\text{Some college}) + \\
& \beta_{5jk}(\text{Graduate degree}) + \beta_{6jk}(\text{Asian}) + \beta_{7jk}(\text{Black}) + \beta_{8jk}(\text{Hispanic}) + \\
& \beta_{9jk}(\text{Other/mixed race}) + \beta_{10jk}(\text{Age}) + \beta_{11jk}(\text{Age}^2) + \beta_{12ijk}(\text{QPosition2}) + \\
& \beta_{13ijk}(\text{Qposition3}) + \beta_{14ijk}(\text{Qposition4}) + \beta_{15ijk}(\text{Qposition5}) + \beta_{16ijk}(\text{Qsensitive}) + \\
& \beta_{17ij}(\text{QDichotomous}) + \beta_{18ij}(\text{Qfuture}) + \beta_{19ij}(\text{QAttitude/opinion}) + \beta_{20ikj}(\text{Activities} \\
& \text{\& behaviors}) + \beta_{21ijk}(\text{Demographic \& household}) + \beta_{22ijk}(\text{Family \& friends}) + \\
& \beta_{23ijk}(\text{Finances \& economics}) + \beta_{24ijk}(\text{Health \& healthcare}) + \beta_{25ijk}(\text{Personality \&} \\
& \text{feelings}) + \beta_{26ijk}(\text{Specific politics}) + \beta_{27ijk}(\text{Race/Ethnic relations}) + \beta_{28ijk}(\text{Religion} \\
& \text{\& spirituality}) + \beta_{29ijk}(\text{War \& terrorism}) + \zeta_{jk} + \zeta_k + \varepsilon_{ijk}
\end{aligned}$$

where ζ_{jk} is the random intercept for the individual-level and ζ_k is the random intercept for the survey level. The random intercept for individuals can be thought of as the combined effect of omitted respondent-specific covariates that cause some people to be more prone to DK responses than others, while the random intercept for the survey adjusts for the clustering by data source.

The survey sources used here contained a variety of survey weights to adjust for differential nonresponse, for oversampling of certain demographic groups, and for unequal probabilities of selection. Analytic uses of survey variables that ignore survey design features are vulnerable to misspecification and inference errors; there is widespread consensus that using survey weights is necessary for descriptive inference (i.e., percentages, means) if the findings are

to be generalized to the population from which the sample was drawn. There is less agreement though on when and whether weights should be used with multivariate methods, such as with the multilevel logistic models used here (Gelman, 2007). Instead of combining the survey weights across data sources, we use a model-based approach to incorporate survey design features into the multivariate model. Because the survey weights were largely a function of the independent variables in the model (e.g., race and education), the unweighted regression models provide unbiased estimates (Gelman, 2007; Winship & Radbill, 1994).

RESULTS

Descriptive statistics for DK response rates by question characteristic and gender are shown in Table 3. Overall, women had a higher DK response rate than men on most measures, with several noteworthy exceptions. For questions related to family and friends, health and healthcare, personality and feelings, or religion and spirituality, men were more likely than women to report not knowing. Sensitive questions, and those asking about politics and government, elicited the highest rates of DKs for both women and men. The lowest DK response rates were observed for questions on activities and behaviors and for demographic or employment information. Consistent with previous findings on DK responses, questions with a dichotomous choice option had more DKs than those with more finely-graded answer options. Questions asking about the future received more DK responses than those asking about the past or present. Both men and women gave the most DK responses to questions that occurred in the middle fifth of the survey.

The results shown in Table 4 are from a multilevel logistic model that used individual and question characteristics to predict DK responses. The models were estimated separately for men and women because individual and question characteristics could exert different influences

on DK responses by gender. The significance ($p < .001$) of these differences was tested and is indicated in the table by the presence of a dagger symbol (†). Rather than using a set of dichotomous indicators for question topic, which would only have allowed comparisons relative to an omitted category, effect coding was used for these indicators (Hardy, 1993). The odds ratios for question topic represent the estimate for each question topic, relative to the unweighted overall mean of all topics. This strategy is advantageous because the comparison is not influenced by the different number of questions that were asked of each topic, and each question topic can be compared to all others (Hardy, 1993).

Prior research found that DK responses were less frequent among highly educated respondents. We observed this pattern, but also found important gender differences with respect to education. Gender differences became more salient as educational attainment increased; women who did not graduate high school were less different from men, when compared to their peers who held a 4-year degree. In fact, men with less than a high school degree were more likely than similarly educated women to give a DK response. Compared to men with a Bachelor's degree, men with less than a high school education were 2.10 times more likely to give a DK response. For men, each successive educational category prior to a 4-year degree reduced DK responses. Unlike men, women with some college were no different from their degree-holding peers. Graduate school had no effect on further reducing the likelihood of a DK response for men. Compared to men with a Bachelor's degree, those with graduate education were neither more nor less likely to say DK. For women, graduate school lowered the incidence of DK responses; women with a graduate education were 20% less likely to say DK than women with a Bachelor's degree.

Consistent with research on general item non-response, we found that non-white race and ethnic groups were the most likely to have given a DK reply. Gendered patterns of DK responses were inconsistent across racial and ethnic groups. Men who identified as Asian and men who were coded in the other/mixed race category, were .65 times more likely to offer DK than white men. Comparatively, Asian women were .44 times more likely to say DK than white women, but those coded in the other/mixed race category were .20 times more likely. The difference between Hispanic and white men was smaller than the difference between Hispanic and White women. Prior research on Black adolescents (Bachman & O'Malley, 1984) suggested we might find Black respondents to have higher DK response rates than white respondents, but this was not observed in this sample of adults; neither Black men nor Black women were more likely to offer DK than their white counterparts. There also appears to be an unexplained interaction between race and gender that may have important implications for survey design and for imputation procedures when DKs are treated as item nonresponse (von Hippel, 2009). We urge readers to interpret the race results with due caution because of the limitations imposed by combining 12 data sources; as noted, the race measure is rather crude, and the other/mixed race category is disproportionately large. Nonetheless, our findings (and current limitations) suggest important avenues for future research.

Regarding respondent age, early research on DK responses hypothesized that younger women and older men were the most likely to give this response (Ziller & Long, 1965). Although the age-gender interactions for DK responses were unremarkable here, the effect of age was clearly non-linear. This echoes the perplexing relationship noted in the literature whereby older respondents are more likely to report not knowing, despite the reality that they probably know more than younger respondents do. Treating age as a non-linear term showed that the

presence of DK responses decreased from ages 18 to 48 and increased thereafter. The oldest respondents in the sample were the most likely to say DK, a relationship that intuitively makes more sense than the notion of age as a positive linear effect. At the oldest ages, cognitive functioning may play a role in DK responses, but there is another important factor; note that all but one of the 12 combined surveys were conducted by telephone. Some older respondents may have had trouble hearing the questions and found it easier to give DK replies than to continually ask that questions be repeated. This could have been the case, even when respondents knew the answers and would have otherwise been motivated to give them.

In addition to respondent characteristics influencing DK responses, question characteristics were also influential, and their effects demonstrated important differences and similarities by gender. Unsurprisingly, sensitivity was the question characteristic with the largest observed effect. When asked a sensitive question, men were 11.5 times more likely to give a DK response, while women were 7 times more likely. Though gender difference here is large, this finding is not without limitations. A large body of research has noted the influence of an interviewer-respondent gender interaction in various types of nonresponse to sensitive questions (Catania, et al., 1996). Unfortunately, it was not possible to control for interviewer gender here, and the observed gender differential could be a spurious relationship, due to survey call centers' having a large proportion of female interviewers. As shown by the descriptive statistics in Table 3, however, men were less likely to have said DK when other variables were unadjusted. This gave us confidence that the relationship observed in the multivariate model was not likely to have been solely an interviewer effect.

Although dichotomous choice options did receive more DK responses than finer-graded categories, the multivariate results showed the opposite effect when other respondent and

question characteristics were adjusted. All else being equal, questions with a dichotomous response option were actually less likely to receive a DK response than questions with more response options. This suggests that the distinction between dichotomous and polytomous responses may not be as important as other characteristics, such as question sensitivity or topic.

Questions requiring speculation about the future received slightly (.02 times) more DK responses from women. Comparatively, men were .30 times more likely to give a DK response to questions that asked about the future than to questions about the past or present. Again, this relationship was undetected in descriptive statistics that did not control for other respondent and question characteristics.

Where a question was positioned within a survey did influence the likelihood of a DK response. Compared to the earliest questions asked (first quintile), those in the second and third quintiles were more likely to receive DK responses. For both men and women, the most DK responses were present in the middle quintile of the survey. Gender differences emerged, however, in the final quintile; men were more likely to give a DK response to questions asked in the final quintile than they were in the beginning of the survey, while women were less likely.

The topic of a question mattered enormously for gendered patterns of DK responses. Men were more likely than women to offer DK to questions about family and friends, health and healthcare, personality and feelings, and religion and spirituality. Women were more likely than men to indicate DK to questions about finances and economics, politics and government, specific policies, race and ethnic relations, and war and terrorism. Both genders were equally likely to offer DK to questions about activities and behaviors, and to those asking for basic demographic or employment information. These gendered patterns were present in both the descriptive

statistics and the multivariate analysis, suggesting that this is a robust and important pattern of DK responses.

CONCLUSION

In several ways, this paper generalizes the current state of knowledge about DK responses. First, the use of 12 datasets in this study's analyses ensures support of the conclusions beyond any single data collection effort. Next, most of the results were consistent with those offered in the literature. Regarding individuals, people with more education; men; and those who identified as being white were less likely than other participants to indicate DK as a response. Regarding survey items, sensitive questions; those with dichotomous response categories; and questions that asked about the future were more likely to receive DK responses. Third, all data in these datasets were collected within the past 15 years, so the current study, in as far as its results align with findings on DK responses from half a century ago, confirms that the latter are still applicable today. Other results indicated relationships that deserve further attention, particularly regarding respondent gender and question topic.

The pre-existing literature on DK responses suggests that women are more likely to report not knowing the answers to survey questions. While this is not incorrect, our data showed that, when it comes to survey questions, men don't know everything either. Men were more likely than women to offer DK responses to questions relating to family and friends, health and healthcare, personality and feelings, or religion and spirituality. That is, men answer sometimes that they "don't know" as well, with some topics making them more uncertain than others and more likely than women, in particular topic domains, to offer a DK. These important gender differences relating to question characteristics have previously gone unnoticed or untested in nationally representative data.

Evidence of a gender-and-topic interaction for DK responses has important implications for understanding the meaning behind these nebulous responses. A leading theory of DK responses suggests that they result from survey satisficing, when unmotivated respondents take shortcuts to complete the survey faster and more easily (Bishop et al., 1986; Krosnick, et al., 2002). One implication of this theory is that DK response options should be excluded from questionnaires (Krosnick, et al., 2002; Mondak & Davis, 2001). An alternative explanation for DK responses is that some respondents truly lack attitudes or opinions about particular topics, and we should probably listen when they tell us so (Sanchez & Morchio, 1992). Because gender and topic differentials emerged net of other respondent and question characteristics in the current, more generalizable study, it seems likely that respondents were expressing lack of interest or opinions, even if satisficing was also partially to blame.

Although it was beyond the scope of this paper to make a judgment as to whether DK response options should be included in questionnaires, our findings suggest that DKs may have a deeper meaning than mere satisficing. It is possible, in fact, that some DK responses could be interesting in their own substantive right. For example, what are the implications for preventative health behaviors when men do not know how many of their family members have had cancer or heart disease, while women do? Overall, this paper reaches the conclusion that the current state of knowledge about DK responses is generally correct, but there is still much to be explained and discovered.

Table 1. Data Source Descriptions

Study	Acronym	Field Period	# of Persons	Survey Mode
ABC News Nukes Poll	Abcnuke	08/2005 - 08/2005	1,001	Telephone (RDD)
ABC News / Washington Post Monthly Poll	Abcpoll	12/2009 - 12/2009	1,002	Telephone (RDD)
CBS News Monthly Poll	Cbspoll	08/2009 - 08/2009	1,096	Telephone (RDD)
21st Century Americanism: Nationally Representative Survey of the United States Population	Century	07/2004 - 10/2004	2,799	Telephone (RDD)
United States Citizenship, Involvement, and Democracy Survey	Cid	05/2005 - 08/2006	1,000	Face-to Face Interview
Public Opinion on the Courts in the United States	Courts	03/2000 - 05/2000	1,563	Telephone (RDD)
National Firearms Survey	Firearm	03/1999 - 07/1999	2,520	Telephone (RDD)
Health Information National Trends Survey	Hints	02/2005 - 08/2005	5,585	Telephone (RDD)
National Survey of Midlife Development in the United States	Midus	01/1995 - 01/1996	3,033	Telephone (RDD)
Measuring Perceptions of Appropriate Prison Sentences in the United States	Prisons	05/2000 - 08/2000	1,299	Telephone (RDD)
Baylor Religion Survey	Religion	10/2005 - 11/2005	1,720	Mixed-mode (RDD & SAQ)
Social Side of the Internet Survey	Social	11/2010 - 12/2010	2,302	Telephone (RDD)

Note – Response rates and calculations varied by source and were unavailable in some public release files

Table 2. Number of Questions by Topic and Survey Source

Question Topic	Data source ^a											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Activities & behaviors	4						4	10	6	2		8
Demographic & employment			3	1	2	1	2	2	3	1		2
Family & friends			1		5		1	2	5		2	
Finances & economics	1	6	11	2			1	2	1	1		
Health & healthcare		1	8				1	6	21			
Personality & feelings		3	2	5	3	1		6	16			3
Politics & government	2	3	9	5	9	1						3
Race/ethnic relations				14	17	4	1			1		
Religion & spirituality		2	1	1	1			1			4	1
Specific policies		12	21		3	4	4			5	2	
War & terrorism	10	6	8		9						2	
Total questions	17	33	64	28	49	11	14	29	52	10	10	17
Total topics	4	7	9	6	8	5	7	7	6	5	4	5

Note – Data sources and acronyms described in Table 1.

^a(1) Abcnuke; (2) Abcpoll; (3) Cbspoll; (4) Century; (5) Cid; (6) Courts; (7) Firearm; (8) Hints; (9) Midus; (10) Prisons; (11) Religion; (12) Social.

Table 3. Question Characteristics and Gender Differences in Don't Know Response Rates

Question Characteristics (<i>N</i> = 334)	"Don't Know" Response Rate (Percent)	
	Female (<i>N</i> =14,113)	Male (<i>N</i> =10,802)
Sensitivity		
Sensitive	11.05	9.80
Not sensitive	3.22	2.84
Response options		
Dichotomous choice	4.49	3.66
Finer-graded categories	3.64	3.34
Time orientation		
Future	5.03	4.51
Past or present	3.79	3.33
Subjectivity		
Fact or knowledge	4.00	3.77
Attitude or opinion	3.72	3.18
Position in Survey		
Very beginning	2.21	1.81
Beginning	3.80	3.57
Middle	7.57	6.54
End	2.17	2.02
Very end	3.49	2.75
Topic		
Activities & behaviors	0.61	0.56
Demographic & employment	0.83	0.62
Family & friends	2.47	3.21
Finances & economics	6.58	3.87
Health & healthcare	4.62	5.29
Personality & feelings	1.00	1.15
Politics & government	10.33	7.10
Race/Ethnic relations	4.14	2.94
Religion & spirituality	3.13	3.33
Specific policies	8.72	6.75
War & terrorism	6.23	4.51

Table 4. Random-Intercept Logistic Regression Models Predicting “Don’t Know” Responses

Variables	Females		Males	
	OR	SE	OR	SE
Education (college grad omitted)				
Less than high school	1.87*	(0.05)	2.10*	(0.07) [†]
High school graduate	1.30*	(0.04)	1.50*	(0.05) [†]
Some college	1.08	(0.04)	1.21*	(0.05) [†]
Graduate degree	0.80*	(0.05)	0.83	(0.06) [†]
Race (White omitted)				
Asian	1.44*	(0.09)	1.65*	(0.10) [†]
Black	1.03	(0.05)	1.07	(0.06)
Hispanic	1.29*	(0.05)	1.18*	(0.06) [†]
Other / mixed race	1.20*	(0.06)	1.65*	(0.01) [†]
Age	0.97*	(0.00)	0.96*	(0.01) [†]
Age-squared	100.05*	(0.04)	100.05*	(0.05) [†]
Position in survey (very early omitted)				
Early	1.35*	(0.04)	1.32*	(0.05) [†]
Middle	2.47*	(0.04)	2.52	(0.04) [†]
Later	0.81*	(0.04)	0.88	(0.05) [†]
End	0.96*	(0.04)	1.15*	(0.05) [†]
Sensitive topic	7.71*	(0.04)	11.45*	(0.04) [†]
Dichotomous response option	0.95*	(0.02)	0.97*	(0.03)
Future oriented	1.02*	(0.04)	1.30*	(0.05) [†]
Attitude/opinion	0.53*	(0.03)	0.57*	(0.04)
Topic				
Activities & behaviors	0.08*	(0.09)	0.07*	(0.07)
Demographic & household	0.32*	(0.07)	0.30*	(0.10)
Family & friends	0.91	(0.05)	1.78*	(0.06) [†]
Finances & economics	0.53*	(0.04)	0.28*	(0.05) [†]
Health & healthcare	1.67*	(0.03)	2.50*	(0.03) [†]
Personality & feelings	0.62*	(0.04)	0.94	(0.05) [†]
Politics & government	1.61*	(0.06)	0.71*	(0.06) [†]
Race/ethnic relations	1.38*	(0.04)	1.04	(0.05) [†]
Religion and spirituality	1.41*	(0.05)	1.76*	(0.06) [†]
Specific policies	3.70*	(0.03)	3.45*	(0.05) [†]
War & terrorism	3.14*	(0.04)	1.94	(0.05) [†]
Random effects (estimated residual standard deviation)				
Individual-level	0.95	(0.02)	1.04	(0.02)
Survey-level	0.57	(0.12)	0.51	(0.11)
Wald chi-squared	13,038.65		9,767.02	

Note—Cells contain odds ratios; age-squared odds ratio is multiplied by 100 and the SE is multiplied by 1000 to show significant digits; question level $N = 334$, individual-level $N = 24,915$, survey-level $N = 12$; * $p < .001$ (two-tailed); [†] Difference between females and males is significant at $p < .001$ (two-tailed)

Chapter 3

“I still don’t know”: Non-substantive Responses in Longitudinal Data

The presence of item non-response in longitudinal data is typically regarded as a nuisance that may bias statistical inference. Researchers have rarely considered the possibility that some types of non-substantive response in fact may offer informative missing data in a longitudinal framework. The stability and predictive power of DK responses from one survey wave to the next may have important implications for survey design features, methods for reducing item non-response, and predicting future panel attrition. Uncorrelated DK responses over time may indicate contextual survey factors are to blame for these responses, rather than indicative of a trait that respondents carry from one survey to the next. DK responses correlated over time may suggest the presence of a stable group of non-responders. In the initial wave of a survey, DK responses also may be predictive of later panel attrition. If this is the case, including a DK or ‘no opinion’ option in a survey design may be worthwhile since identifying likely future study dropouts is an important step in improving response rates in panel data.

BACKGROUND

Non-substantive responses occur when respondents indicate that they are unsure, undecided, cannot recall, don’t know, or have no idea. Understanding non-substantive responses is fundamental for optimal survey design and for identifying sensible methods to handle these responses in data analysis. In statistical analysis, for example, we need to understand these responses before we know if they are “missing data” or whether DKs can legitimately be recoded as a middle category, as many researchers do. When designing surveys, understanding the DKs is important for deciding whether these responses should be allowed, and if so, whether the option should be explicit. Because understanding these responses has such enormous

implications, a great deal of research has been dedicated to figuring out what respondents actually mean when they report to not know an answer (Beatty & Herrmann, 2002; Rapoport, 1979; Smyth et al., 2006). Based on suspected meanings, best practices in survey research cycled between recommending an explicit DK option always be included in questionnaires to arguing instead that a forced-choice option is best (Giljam & Granberg, 1993; Poe, et al., 1988; Sanchez & Morchio, 1992; Schuman & Presser, 1980). Which option is ideal for a survey, however, is largely dependent on the meaning attributed to DK replies.

Primarily four theories about the underlying meaning of DK responses have been demonstrated in the literature on nonsubstantive responses. A common hypothesis is that DK responses are an expression of survey satisficing, where unmotivated respondents take shortcuts to complete the survey faster and more easily (Krosnick, 1991, 2002). According to this theory, providing DK responses may be less demanding than remembering and applying relevant information to the response task (Beatty & Herrmann, 2002). For example, some people have trouble remembering their wedding anniversary; rather than looking the information up or considering the date at length, the satisficing theory suggests that respondents learn throughout the survey to circumvent optimal responding – which is more work – and instead say DK. Offering an explicit DK option, or even allowing DK responses, may encourage satisficing and unnecessarily increase these uninformative responses. For questionnaire design, the satisficing theory implies that a forced-choice response option is preferable.

A second theory about the underlying cause of DK responses is that they are an expression of a general response style. Like straight-lining or agreeing with all questions, certain people could be prone to offering DK regardless of the question content (Beatty & Herrmann, 2002). Bachman and O'Malley (1984) showed that respondents have a rather stable tendency to

use, or to avoid, extreme response categories, supporting the idea that individuals have consistent answer styles over time. A general tendency to offer DK replies is not exclusive of the satisficing theory as satisficing itself could also be a distinct response style. This theory suggests that response style is primarily a manifestation of individual differences or personality traits, rather than a being a function of the survey context. Again, the implication of this theory for questionnaire design is that DK replies are not reflecting the respondent's true state and should probably not be allowed.

A third suspected cause of DK responses is that some people claim to not know an answer because they do not want to reveal the true answer to a question. This passive refusal theory suggests when an individual does not want to reveal the true answer to a question, saying DK is a gentle way to refuse (Berinsky, 1999; Rubin, et al., 1995). For example, people may wish to hide publically unfavorable attitudes (such as racism), or unacceptable behavior (such as child abuse) from an interviewer because it is too socially uncomfortable to discuss (Shoemaker et al., 2000). Sensitive questions, in general, are known to generate higher levels of item nonresponse (Berinsky, 1999), evidence supporting the belief that at least some DKs are passive refusals. Whether excluding the DK option from questionnaires is the best way to reduce passive refusals is unclear. On the one hand, a forced-choice option could nudge respondents into offering a meaningful response. On the other hand, when not give the option to say DK, respondents could still offer a hard refusal, or perhaps worse, intentionally give an incorrect response. Bullying a respondent to give an answer hardly seems prudent in this circumstance; a reduction in non-substantive response is a poor strategy if the data are meaningless (Sanchez & Morchio, 1992).

A final theory of DK responses is that, although an underlying meaning to DK responses may exist, surely there are times when the information is unknown and cannot be estimated. According to this theory, respondents may truly lack attitudes or opinions about particular topics, and we should probably listen when they tell us so (Jean M. Converse, 1976; Sanchez & Morchio, 1992). For example, someone who has never taken a blood sugar test could find it impossible to estimate his or her glucose level. Probing DK responses, in this case, could improve the number of numeric replies but would largely be the product of guesswork; these guesses would often be implausible, and therefore unusable, anyway. Additionally, DK responses may be interesting in their own substantive right (De Leeuw, 2001). Staff et al. (2010), for example, found that adolescents who reported not to know about their future career aspirations had lower socio-economic attainment at age 26 when compared to their peers. The theory that DK really means DK implies that surveys should routinely allow DK responses.

Even after decades of research on nonsubstantive responses, theories about the underlying meaning of DK responses continue to suggest conflicting strategies for questionnaire design. Survey researchers recognize that these theories overlap considerably and that DK responses for each reason are likely to be present within single question, but that recognition has done little to solve the DK dilemma. One limitation of prior research on DK responses, with a few notable exceptions (Duncan & Stenbeck, 1988; Durand et al., 1983; Rapoport, 1985), is that the majority of research has been done with cross-sectional data only. As with almost all answers to survey questions, fully understanding the meaning of DK responses is difficult to do without observing their patterns over time.

A major aim of this paper is to contribute to existing literature on DK responses by asking the following questions about these responses in longitudinal data: (1) Do DK responses

early in the survey predict later use of this option within the same interview? (2) Does the tendency to say DK to the first survey predict the tendency to say DK in a follow up survey? (3) Do respondents repeatedly offer DK to the same questions? (4) Does the tendency to say DK in the first survey predict which respondents do not participate in the second survey? An additional aim is to discuss how these findings relate to theories about DK responses and what this might imply for questionnaire design.

DATA AND METHOD

To examine DK responses in panel data we used two waves from the National Survey of Midlife Development in the United States (MIDUS). The first wave of data was collected from 1995-1996 and the second wave was collected from 2004-2006. Participants were recruited by random digit dialing (RDD). All respondents were invited to participate in a phone interview of approximately 30 minutes in length and to complete two self-administered questionnaires (SAQs), to be mailed back upon completion. Only telephone interview data will be used in this study because the SAQs did not contain an explicit DK option. The sample size for the main respondent who completed the telephone interview was 3,487 for the first survey and 2,257 for the second survey. The original RDD response rate was approximately 70 percent (Brim et al., 1996b). Adjusting for mortality, the longitudinal retention rating was approximately 71% (Brim, et al., 1996b).

The MIDUS survey content was primarily about health-related questions. For example, the third question in the survey asked “In general, compared to most (men/women) your age, would you say your health is much better, somewhat better, about the same, somewhat worse, or much worse?” There were modules in the survey about health incidence, such as cancer and heart disease, and questions about health behaviors, such as cigarette smoking and alcohol use.

Additional questions were largely demographic, asking about education, occupation, marital status, and completing a household roster. The content of the survey is an important feature because question topics are likely to influence DK responses. For example, women are widely cited to have higher DK response rates than men, but this trend largely comes from literature on data about political topics (Mondak & Davis, 2001). When questions about vitamins and nutrition were asked about, Durand and Lambert (1988) found that men actually had higher DK response rates than women. Different demographic groups may have varying interest in health questions, so it seems likely that DK responses in the MIDUS data may be different from those observed in other types of data (e.g., political polls).

DK responses were allowed and recorded when offered, although the MIDUS design never explicitly encouraged DKs by the choice of response options. This strategy is likely to have reduced the overall number of DK responses and to have discouraged some respondents from satisficing. Additionally, the interviewers were instructed to probe the DK responses for selected questions. For example, the fourth question in the survey asks, “In the past 30 days, how many days were you totally unable to go to work or carry out your normal household work activities because of your physical health or mental health?” When a respondent indicated uncertainty, the interviewer asked, “What’s your best guess?” It is important to keep these strategies in mind because the DK responses that were recorded in the MIDUS data may be reflective of only the most assertive and insistent respondents. The behavior of DK responses in other types of panel data, where different DK strategies were used in the questionnaire, could be different from those reported here. Nonetheless, the way DK responses were handled in the MIDUS survey is fairly typical, so the patterns we observe here may be present in many other data sources.

The first survey contained 213 questions analyzed here. The average respondent received 107 questions; the fewest questions asked were 75 and the most questions asked were 160. Of 3,487 respondents, 1,030 (30%) never gave a DK reply. The average respondent said DK to 2 questions, and the most DKs recorded for one person was 38. On average, 1.9% of all answers were DK replies with a standard deviation of about 2%. A minority of respondents answered DK to a rather large proportion of questions; 29 respondents offered DK to more than 10% of questions they were asked. To ensure that our results were not unduly influenced by these “outliers,” all analyses were run with and without these 29 cases. The results of this sensitivity analysis (not shown) yielded identical substantive conclusions to those presented here and are available upon request from the first author.

The second survey, conducted nine to ten years after the first survey, contained 288 questions analyzed here. The average respondent received 139 questions, a greater number than were asked in the first interview. The fewest questions asked were 94 and the most asked were 207. Of the 2,257 respondents, 609 (27%) never gave a DK reply. The average respondent said DK to 2.4 questions, and the most DKs recorded for one person was 27. For all applicable questions, 1.8% of all answers were DK replies, with a standard deviation of about 2%; this was a nearly identical distribution of DK responses to those observed in the first survey. Again, a few respondents had a particularly high DK rate, offering this response to more than 10% of questions. We tested our models with and without these responses and found their exclusion did not change the results in a meaningful way.

Our first research question asked about within-survey use of DK responses. Specifically, we wondered if the presence of a DK response early in the survey could predict the proportion of all later responses that were non-substantive. To answer this question we used as a dependent

variable the percentage of all responses after the first five questions that were DKs. This was a straightforward measure where the number of DK responses was divided by the total number of questions the respondent received, excluding the first five questions. We predicted this proportion in a linear regression model with a dichotomous indicator of whether a DK response occurred within the first five questions. Our use of the first five questions was rather arbitrary; our goal was to capture early survey behavior but of course there was nothing magical about the first five questions. We also tested models with the first three and first ten questions and the results were indistinguishable from those reported here. Of 3,487 respondents, 89 (2.6%) offered a DK reply to one of the first five questions.

The second research question we asked was whether the tendency to say DK in the first survey would predict the tendency to say DK in the second survey. The dependent variable in this model was the percentage of all responses in the second survey that were DKs. The dependent variable was predicted in a linear regression model by the percentage of responses at the first wave that were observed to be DKs. Because not all people were present for the second survey, and our fourth hypothesis was that DK responses are related to survey dropout, we used inverse probability weighting to adjust this model for unit nonresponse. The attrition model included marital status, self-rated health, education, and gender – all factors that were significant predictors of non-contact or refusal to participate at the second survey. Experimenting with different attrition models, and even failing to adjust for attrition, did not change the substantive findings we present here. Again, results from this sensitivity analysis are available from the first author.

Our third question asked whether people repeatedly say DK to the same questions. We analyzed 40 questions with the exact same wording that were asked 9-10 years apart. These 40

questions were selected because they were asked of all respondents at both waves. Many of the questions in the MIDUS data were inapplicable for many respondents. For example, only female respondents were asked questions about cervical cancer history and treatment. Selecting only those questions asked of all respondents could have eliminated some of the more sensitive questions asked in the MIDUS data because the most specific questions were not used. Again, this is important to bear in mind when reviewing the patterns observed in this study since question sensitivity is believed to elicit DK responses. We reshaped the 40 questions into a person-question level file, resulting in a sample size of 121,842 and used descriptive statistics to illustrate the patterns of repeated DK responses.

The final question asked was whether DK response tendency in the first survey predicted panel attrition. The dependent variable for this model was whether or not a survey was obtained for the respondent at the second wave. Unfortunately, the reason that a survey was not obtained at the second wave was unavailable; it would be interesting to know if early DKs were predicting later refusals, non-contacts, or both. Completion of the second survey was predicted in a logistic regression model using the proportion of DK responses from the first survey as the predictor of interest. Although we do not show the results here, we also tested whether a DK response in the first three, five, or ten survey questions was predictive of panel attrition and they were not.

For all the multivariate models reported here, we show both unadjusted coefficients and those adjusted for respondent gender, age and education. These are important adjustment variables because the survey data were analyzed unweighted, and people are known to have differential unit and item nonresponse based on these characteristics. Although survey weights are necessary for descriptive analysis, there is less agreement on when and whether weights should be used with multivariate methods, such as with the logistic models used here (Gelman,

2007). Instead of using weighted regression, we use a model-based approach to incorporate survey design features into the multivariate model. Because the survey weights were largely a function of the independent variables in the model (e.g., age and gender), the unweighted regression models provide unbiased estimates of the coefficients and more efficient standard errors (Gelman, 2007; Winship & Radbill, 1994). Additionally, age, gender, and education are consistently shown to be related to DK responses (Beatty & Herrmann, 2002; Linn, et al., 1987). A quadratic term for age was used because this was a significant non-linear effect in all models.

RESULTS

Table 5 shows a linear regression model that used the presence of a DK response occurring in the first five questions of the survey to predict subsequent DK replies. As shown by both the adjusted and unadjusted b-coefficients, early use of this non-substantive response option was strong predictor of overall DK response tendency. In the unadjusted model, a DK response having occurred in the first five questions explained 12.6% of the variance in the overall proportion of DKs. In the adjusted model, respondents who had offered DK to one of the first five questions had around 4.7% more DK responses throughout the rest of the survey, compared to those respondents who gave substantive answers to the first five questions.

This pattern is consistent with what the survey satisficing and response style theories suggest. Early DK responses predicting later responses may show that disinterested respondents learn early on in the survey that a DK reply is a simple way to move through the survey and continue to use this option, consistent with the satisficing hypothesis. This pattern also shows that, satisficing or not, certain people are prone to greater use of DK replies, supporting the theory that DK responses may be a general response style. Since the first five questions of the survey seemed relatively benign in nature, we think it is unlikely that this measure captured

passive refusals to sensitive questions. It is possible that the first five questions captured legitimate DK responses, where participants were unable to estimate the information being asked for, such as the number of days that work activities were limited. Overall, looking at DKs over time within a single survey provided little additional insight into what lies behind these responses because multiple theories were supported. Knowing that early DK responders are likely to keep offering nonsubstantive replies, however, is still useful information when considering questionnaire design.

The satisficing theory suggests that survey cues teach respondents about the acceptability of DK responses – either because the choice is offered, or because interviewers accept the response. Respondents come to believe that this strategy is faster and less cognitively demanding than processing the question and choosing a substantive response. Since the first five questions appear to set the tone for the rest of the survey, it may be possible to utilize these questions to discourage satisficing, but still allow DK responses for other questions in the survey. For example, respondents who offered early DKs could be given a battery of questions asking why the information was not known, what would be useful in estimating the information, and so forth. Using this type of questionnaire design, respondents prone to satisficing might come to believe that giving a substantive response is faster or easier, and therefore avoid the DK option when the information is really known. Respondents who legitimately did not know, however, might still be likely to continue saying so in later questions, and this answer could be useful to record. Combining what we would expect theoretically with the knowledge that DK responses early in the survey predict their occurrence in the rest of the survey suggests that the dilemma of whether or not to include a DK option on a survey may be more nuanced than a yes or no decision. Perhaps satisficing, or even avoiding expression of a DK response style, could be

reduced by carefully planned questions early on in a survey. If this worked, DK responses could then be accepted and interviewers would not be required to probe participants into meaningless responses throughout rest of the interview.

Table 6 shows a linear regression model using the proportion of DK responses from the first survey to predict a respondent's proportion of DK responses in the second survey. A respondent's tendency to say DK during the first wave was quite predictive of his or her tendency to say DK almost a decade later. Unadjusted, DK responses to the first survey explained 15.7% of the variance in DK responses to the second survey. In the adjusted model, each percentage greater in DK responses observed at the first wave was associated with around two-fifths of a percent more DKs at the second survey.

There are several possible explanations for respondents' consistent tendency to offer DK responses over time. A substantial proportion of questions asked during the first survey were repeated in the second survey. Respondents may have continually found the same questions to be boring, or too demanding, and satisfied to the same questions. If DK responses are a general response style, the same questions could have triggered identical reactions towards ambivalence from year to year. Perhaps an even simpler explanation is that respondents were legitimately trying to communicate, "I still don't know," and this was the answer category that best reflected their condition.

Our third research question asked whether respondents were offering DK responses to repeated questions. As shown in Figure 2.1, of those who gave a substantive response in the first survey, 63% gave a substantive response to the same question at the second survey, 36% were not present for the second survey, and 1% gave a DK response. Of those who offered a DK

response during the first survey, 24% gave a substantive response to the same question during the second survey, 36% were not present, and 40% said DK again.

Prior research by Bishop et al. (1984) showed that respondents try and be consistent within themselves to an interviewer. For example, respondents might think “What must my interest in politics be if I just told you I don’t know what my congressman has done for the district or how he has voted on any legislation?” before they answer the question (Bishop, et al., 1984). Although it is possible that some participants remembered having answered DK to some questions from the prior survey, and repeated a DK response to be consistent, this is an unlikely explanation given the span of time between the surveys. Rather, the simplest explanation for repeated DK responses is that this answer reflects a true and persistent state (e.g., continuing to have no interest in a topic, remaining undecided, or still not remembering). This suggests that trying to solve the dilemma of including or excluding the DK option in a survey may not be the primary issue relevant to analyzing data from this group of people. A more informative task for researchers may be to ask how well people are informed about the question topics generally, why uncertainty or ambivalence is higher for some people than for others, or why respondents were uninterested in particular questions but agreed to the survey while others did not.

The final question we asked about DK responses in longitudinal data was whether a respondent’s tendency to use DK responses in the first survey would predict panel attrition. As shown in Table 7, the DK response propensity in the first survey was strongly predictive of dropping out by the second wave. Gender, age, and education held equal, respondents who never said DK in the first wave had a 31% predicted probability of not responding to the second survey nine to ten years later. Respondents with 5% DK responses in the first survey had a 41% predicted probability of dropping out of the second survey, slightly higher than the overall

dropout probability of 35%. The respondents who gave the most DK responses during the survey were more likely than not to dropout; of those who had given 10% or more DK responses, more than half were predicted to not be present at the second survey.

Whatever the reason people give non-substantive answers, the presence of DK responses is clearly a bad omen for future data collection efforts. Although identifying who is most likely to drop out of a survey is an important step in reducing future unit non-response, showing that DK responses predict panel attrition may not be reason enough to collect them. Item nonresponse from other causes is widely recognized to predict survey dropout, for example, so collecting DKs in addition to refusals might not add any information. To see if refusals to questions in the MIDUS data were as predictive of survey dropouts as DK responses, we used tendency to refuse answers in the first survey (% of all applicable questions that were refusals) to predict attrition. As shown by the results in Table 8, refusals were not predictive of panel attrition for the MIDUS data. Keep in mind that since a DK option was allowed, although actively discouraged through probing, for some questions, the refusals in these data may be different from refusals coming from a data collection effort where DKs were not allowed at all. We believe this would be a productive avenue for an experiment in future data collection.

CONCLUSION

DK responses to the first five questions predicted later use of the DK option; this suggests that, consistent with the satisficing hypothesis, respondents may learn that saying DK is an easy way out. If respondents are using the DK option as a satisficing mechanism, the implications for survey design are clear. An explicit DK option could be removed from paper or web questionnaires and never mentioned as an option during phone or face-to-face interviews. When DKs are offered, interviewers could be instructed not to allow them, or to try probing

rather than easily accepting these responses. Of course, if the DKs result from other causes – and these findings suggest that they do – this could be a risky strategy because respondents could be tempted to make up data or become annoyed and end the survey.

Looking at the stability of DK response tendency over time, there was evidence that saying DK may be a generalized pattern of responding. DK response propensity was fairly stable over time, consistent with the hypothesis that DKs result from a generalized response style. For example, we saw implications that some people may be generally uncertain or ambivalent. Satisficing cannot be ruled out by this measure, because it is also possible that satisficing itself a type of personality trait or response style that respondents carry from one survey to the next. We also showed that people who offered a DK reply the first time were quite likely to offer another DK when asked the same question 9-10 years later. This suggests that if DKs result from a response style, this style is expressed repeatedly for particular types of questions. Alternatively, respondents may be telling researchers something more obvious – they still “don’t know.”

Survey methodologists have long recognized the pros and cons of including an explicit DK option in the survey design. One negative consequence of having an explicit option is that it may encourage respondents to satisfice, making their answers unnecessarily unusable. The evidence presented here, however, suggests that satisficing is only one of many reasons that respondents choose to use the DK option. A large proportion, 40%, of respondents who said DK in the first wave still did not know when asked the same question nine to ten years later. This suggests that while some survey satisficing probably does occur in surveys, a significant number of respondents are not being lazy and genuinely do not know the answers to the questions. For survey design, this presents a good case for permitting the DK option on surveys rather than

having forced-choice answers only. DK responses may also be meaningful in their own right, particularly for those who moved from uncertainty to certainty, or vice versa.

Another advantage of including the DK response option on surveys is that these responses in the first wave were highly predictive of panel attrition, above and beyond basic demographic characteristics that are known to be related to item nonresponse. Knowing this information could help researchers immediately identify which respondents have a high risk for dropping out of the study. Investigators could then employ targeted strategies to prevent attrition. If the period of time between data collection waves was extensive, for example, contacting respondents between collection periods to continue building rapport and update contact information is known to reduce panel attrition (Taris, 2000). Such an endeavor is costly for studies with a large number of participants. One efficient strategy might be to target follow-up toward only the respondents who are the most likely to drop out, particularly those with a large proportion of DK responses.

Although it was beyond the scope of this paper to test our ideas for questionnaire manipulation, it would be worthwhile to explore these alternatives in future questionnaire experiments. It would be nice to know if the use of satisficing-reducing techniques in the first few questions of a survey could pave the way for a less problematic DK option thereafter. For example, when a respondent offered DK, interviewers could launch into a series of open-ended follow-up questions, asking respondents why they don't know or what additional information would be useful in making a decision or retrieving the information. In this way, respondents would be discouraged from saying DK as a way to move the survey along. Another interesting design option would be to give some respondents only a refusal option while others could offer both DKs and refusals. A researcher could then tell if refusals only predicted panel attrition while

refusals with a DK option did not, allowing an improved assessment of how useful DK responses are in predicting future dropout.

Table 5. Linear Regression Model Predicting Proportion of DK Responses ($n= 3,487$)

	Unadjusted		Adjusted	
	B	SE	B	SE
DK occurred in first five questions (=1)	4.740***	0.211	4.147***	0.206
< High school			1.479***	0.114
High school graduate			0.308***	0.082
Some college (reference)				
College graduate			-0.278**	0.094
Graduate school +			-0.392**	0.119
Age in years			-0.048**	0.016
Age in years squared			0.061***	0.016
Female			-0.074	0.064
Intercept	1.754***	0.038	2.419***	0.376
R-squared	0.126		0.200	

Note – Age-squared is multiplied by 100 to show non-zero digits.

* $p < .05$, ** $p < .01$, *** $p < .001$ (all two-tailed t-tests)

Table 6. Linear Regression Model Predicting Proportion of DK Responses at Second Survey ($n = 2,257$)

	Unadjusted		Adjusted	
	B	SE	B	SE
DK proportion at first wave (9-10 years prior)	0.437***	0.037	0.387***	0.037
< High school			0.611**	0.196
High school graduate			0.155	0.094
Some college (reference)				
College graduate			-0.162	0.095
Graduate school +			-0.190	0.132
Age in years			-0.611**	0.024
Age in years squared			0.087***	0.026
Female			0.153*	0.074
Intercept	0.989***	0.060	1.846***	0.523
R-squared	0.157		0.198	

Note – Age-squared is multiplied by 100 to show non-zero digits; sample is weighted by the inverse probability of dropout to adjust for panel attrition

* $p < .05$, ** $p < .01$, *** $p < .001$ (all two-tailed t-tests)

Figure 1. Pattern of DK Responses to 40 Questions Asked 9-10 Years Later ($n = 3,487$)

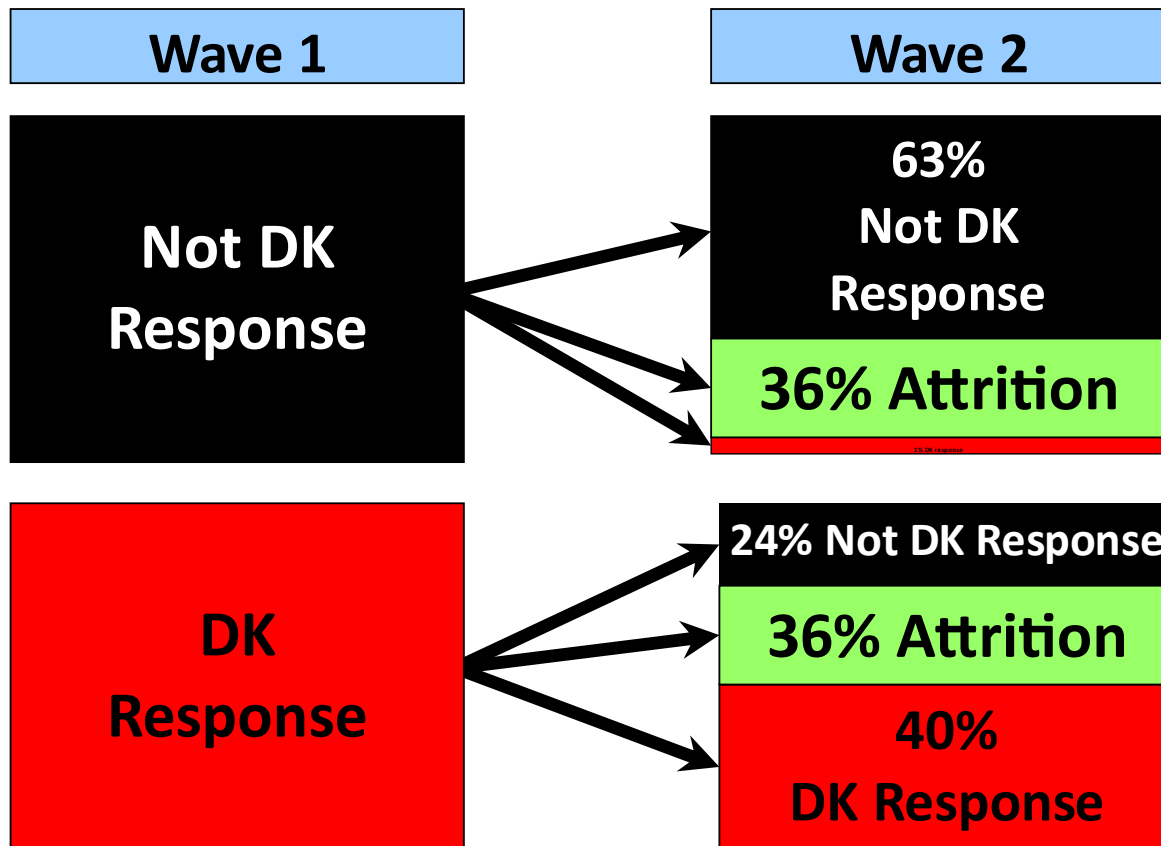


Table 7. Logistic Regression Model Using DK Response Tendency to Predict Panel Attrition ($n = 3,487$)

	Unadjusted		Adjusted	
	OR	SE	OR	SE
DK proportion at first wave (9-10 years prior)	1.145***	0.021	1.095***	0.021
< High school			1.752***	0.222
High school graduate			1.020	0.093
Some college (reference)				
College graduate			0.559***	0.062
Graduate school +			0.489***	0.073
Age in years			0.909***	0.017
Age in years squared			1.001***	0.000
Female			0.762***	0.056
AIC	1.282		1.248	

Note –Cells contain odds ratios.

* $p < .05$, ** $p < .01$, *** $p < .001$ (all two-tailed t-tests)

Table 8. Logistic Regression Model Using Refusals to Predict Panel Attrition ($n = 3,487$)

	Unadjusted		Adjusted	
	OR	SE	OR	SE
Refusal proportion at first wave (9-10 years prior)	1.042	0.067	1.017	0.070
< High school			2.029***	0.249
High school graduate			1.047	0.095
Some college (reference)				
College graduate			0.546***	0.061
Graduate school +			0.470***	0.070
Age in years			0.905***	0.168
Age in years squared			1.001***	0.000
Female			0.758***	0.055
AIC	1.299		1.255	

Note – Cells contain odds ratios; using a chi-square test of the likelihood ratio, the unadjusted model was not significant at the .05 level.

* $p < .05$, ** $p < .01$, *** $p < .001$ (all two-tailed t-tests)

Chapter 4

Strategies for Analyzing Data from Respondents who “Don’t Know”

Social science research is founded upon gathering data from people. Social scientists ask people questions and use the answers to draw conclusions about the state of society. Many surveys include “don’t know” or “no opinion” (hereafter referred to as DK) options to ensure an answer for each question. For researchers analyzing secondary data, dealing with DK responses is necessary if these responses were collected in the survey. Unfortunately, no guidelines exist for how to handle this type of response in data analysis. We believe a guide would be useful because DK responses are present in many family-related datasets such as the National Longitudinal Study of Adolescent Health Add Health), National Survey of Families and Households (NSFH), and the Fragile Families and Child Wellbeing Study. Researchers make varying decisions about handling the DK responses while little attention is given to the possibility that research findings may be conditional on the meaning attributed to and the handling of these replies. An important contribution of this article is an explicit discussion of a theoretical framework for DK responses and the development of appropriate procedures when “don’t know” responses are observed in survey data.

A number of methods are available for handling DK responses, but extant techniques have only had minimal empirical evaluation. In this article we compare several common strategies for handling DK responses in scale items. These methods include: complete case analysis, coding the DKs as “neutral” or “middle” categories, coding all DK responses in one direction (for example, as all “no” responses), available item averaging, and multiple imputation. We show that different methods for handling the DK responses can influence the conclusions drawn from multivariate analysis. Our goal is to help researchers identify the optimal strategies

for analyzing data that contain DK responses by identifying which techniques are the least biased, most efficient, and most practical for family researchers to implement.

BACKGROUND

“Don’t know” responses plague social science data. We illustrate this problem with a 4-item scale, shown in Table 9. The scale items measure values on marriage and family (Rossi, 2001) and come from the National Survey of Midlife Development in the United States (MIDUS) collected from 2004-2006. The four questions ask whether men and women can have happy and fulfilling lives without marrying and without having children. Responses are coded on a scale of 1 (*Strongly Agree*) to 6 (*Strongly Disagree*); higher scores reflect greater importance of marriage and children for life satisfaction. Each of the scale items has a triangular frequency distribution, with large proportions of respondents agreeing with the question and few respondents disagreeing. “Don’t know” responses were recorded for each item and ranged from 12% to 14% for each item. The scale is constructed by taking the average of the four items. This scenario leaves researchers with the choice of dropping the cases entirely when they have DK responses to some questions, or, alternatively, using techniques that require accurate assumptions about the underlying meaning of DK responses.

Methods for Treating DK Responses to Scale Items

There are six strategies commonly employed by family researchers to construct scales in the presence of DK responses: complete case analysis, middle/neutral category coding, directional coding, available item averaging, dummy variable adjustment, and multiple imputation (MI). Other options are available, such as full-information maximum likelihood (Enders, 2010). This strategy is not reviewed here because FIML and MI produce equivalent results when the models are the same and sufficient imputations are used (Collins et al., 2001)

therefore, an exploration of these techniques would be redundant. Item response theory is another potential option, but our review of family literature suggests that it is infrequently used in this context, and may indeed be inappropriate in the absence of a theoretical model suggesting how the DKs are ordered relative to the other response categories. We implemented each commonly used technique to average six versions of the marriage and family values scale, shown in Table 10. Below, we give a brief discussion of these techniques, along with comparisons of the scale properties observed with each technique.

Complete case analysis. A simple technique for dealing with scale-item DK responses is to exclude the cases from further analysis. Complete case analysis is widely recognized as an inefficient procedure because excluding all cases with missing data may substantially reduce the number of cases available for analysis (Acock, 2005; Allison, 2001). Additionally, if the DK responses did not occur completely at random, excluding the cases from the analysis may bias the resulting estimates (Acock, 2005; Allison, 2001). In the marriage and family values scale, as shown in Table 10, excluding any item containing a DK response from the summated scale resulted in substantial data reduction; 643 cases, or 21% of the entire sample, were lost. Using only complete cases, the scale mean was 2.22 with a standard deviation of 1.27.

Neutral/middle category coding. To avoid discarding cases, researchers frequently handle DK responses by coding them as a “middle” or “neutral” category. The origins of this method are unknown, but the idea behind it is intuitive; since respondents commit neither to agreeing nor to disagreeing, DK responses are assumed to be indicators of neutrality or ambivalence. To capture this neutrality, the DK responses are coded to the middle category of the response options. In our example, we coded the DK responses to equal 3.5, the middle value of our six response categories. Although widely accepted as a valid method, the consequences of coding

DKs as a neutral category have been given little empirical scrutiny in many of the situations where this technique is commonly employed. When implemented in the marriage and family values scale, middle category coding produced a mean of 2.43 and a standard deviation of 1.25.

Directional coding. Another common method of recoding the DKs to avoid discarding cases is to assume that the DKs “really” fall into one category. It is common for yes/no format questions, for example, to count only the number of “Yes” responses as indicators of the measure while DK responses are treated as a certain “No” (e.g., Maume, 2006). Although many people would find neutral/middle category coding to be an intuitive solution for the marriage and family values scale, it is also possible that researchers would use directional coding if the belief was that respondents said “DK” to “hide” their true feelings. Since few people who responded to our scale strongly disagreed with the importance of marriage and childbearing, we know such answers to be unpopular. Some might be ashamed to hold this unconventional belief and say “DK” to avoid revealing private information. In this case, the DKs may be best represented if coded as “*Strongly Disagree*”. Directional coding produced the highest observed mean (2.72) and largest standard deviation (1.59) for the summated scale of all strategies.

Available item averaging. Available item averaging avoids the problem of assuming that the DKs really fall into a particular category; it makes use of the known responses only. In this technique, the scale value assigned for each respondent disregards DK responses and averages only those items relevant and available for each case (Schafer & Graham, 2002). This method also has a nice intuitive appeal; since we would expect the scale items to be highly correlated, using only the known responses seems less likely to bias each respondent’s scale average in an unreasonable direction. In our example, available item averaging produced a scale mean of 2.33 and a standard deviation of 1.35. A limitation of this technique, as can be seen in the sample size,

is that respondents must have answered at least one of the scale items. Otherwise, the entire scale remains missing.

Dummy variable adjustment. Certainly there are times when respondents lack attitudes or opinions about the survey questions. To legitimize DKs as a response category, some researchers opt for dummy variable adjustment. In this technique, a dichotomous indicator for whether or not a DK occurred is added to a multivariate model, and DK responses on the scale items are replaced with the item mean. In this way, researchers argue that the DK responses have been “controlled for”. In the marriage and family values scale, replacing the DKs with the average scale score produced a scale mean of 2.28 and a standard deviation of 1.18, which was the lowest standard deviation among the techniques.

Modern methods. Modern missing data techniques are increasingly implemented when dealing with DK responses in scale items. Full-information maximum likelihood (FIML) and multiple imputation (MI) are equally attractive options (and the results are equivalent), although MI may offer more flexibility in its compatibility with a variety of statistical models at present (Johnson & Young, 2011). MI has been routinely tested as a technique for dealing with general, item-level nonresponse to scale items (e.g., Roth et al., 1999). Exploring MI in relation to DK responses is a worthwhile addition to this literature because the performance of any imputation technique is dependent on the type of missingness, and DK responses may be different from other kinds of nonresponse (Duncan & Stenbeck, 1988; Little & Rubin, 2002). In our example, we multiply impute the DK responses using the Stata ICE software program. Estimates from twenty-five datasets were pooled with the MIM prefix and produced a scale mean of 2.32 and standard deviation of 1.23

The different strategies for dealing with DK responses produced similar, though not identical scale means and reliability coefficients. The differences in the distributions of the scales, particularly at the disagreement end of the scale, were notable. For example, if we were trying to estimate the proportion of the population who disagreed or strongly disagreed, the estimates would range from 7% to 19% based on the method selected! Clearly, the method researchers select for treating DK responses is important and can influence the results they obtain. But which method is optimal? The most appropriate strategy for analyzing DK responses depends upon an understanding of their causes. Since DK responses occur in different contexts, it is worthwhile to consider these responses in a general and systematic way.

Theories about the Underlying Meaning of DK Responses

Under some circumstances, it is reasonable to treat DK responses like any other type of item nonresponse. When a survey question requires the respondent to report a fact or express knowledge on a particular matter, a “true” value exists for the respondent. For example, all people have an age, weight, and blood pressure; a factual answer exists for every person, regardless of whether a respondent knows the answer or not. A true answer also exists for knowledge questions; if a participant is asked in what year Abraham Lincoln was shot by John Wilkes Booth, the answer is 1865, even if the respondent does not know this. Provided that the research question itself has nothing to do with uncertainty regarding a given item, DK responses are ordinary missing data when a true value exists and do not require special consideration beyond the typical attention that must be given to nonresponse (Rubin, et al., 1995).

DK responses are a special case, however, and different from general missing data, when a survey question concerns an attitude or opinion. It is not at all clear that a true value exists under such a circumstance; surely there are times when people do not have attitudes or opinions

about the questions we ask (Sanchez & Morchio, 1992). In fact, this is why “don’t know” or “no opinion” responses are collected in the first place (Philip E. Converse, 1964; Schuman & Presser, 1981). This circumstance, like the one shown in the example scale from the MIDUS data, requires researchers to guess about the underlying meaning of the DK response before the data can be analyzed. There are four leading theories about what DK responses mean, with different implications for the appropriate approach to analyzing “don’t know” data.

Response style. The generalized response-style hypothesis suggests that, like ‘yea-saying’ or straight-lining, saying “don’t know” may be a generalized style of answering questions (Bachman & O’Malley, 1984). In this case, some people have a tendency to say “DK” to questionnaire items regardless of content or any other personal characteristic (Durand, et al., 1983). Although this hypothesis seems the least likely to explain all of the DK responses, if true, it would have important consequences for the treatment of these answers because it postulates that the DK responses are missing completely at random (MCAR). Additionally, this hypothesis appears to be the implicit assumption that many researchers make when handling DK responses (e.g., Anderson, 2008; Goettems et al., 2010).

Satisficing. A second popular hypothesis suggests that DK responses are the result of survey satisficing, which occurs when respondents settle for a satisfactory response that will appease the interviewer, rather than searching for an optimal response (Krosnick, 1991; Krosnick, et al., 2002). Some respondents are unmotivated to give the best answer possible, and they take shortcuts to complete the survey faster and more easily. A DK response is often less work than remembering and applying relevant information (Beatty & Herrmann, 2002). For example, if some people have difficulty remembering their wedding anniversary, rather than look the information up or suffer the cognitive demand of deducing the information, they might

simply say that they do not recall. The satisficing hypothesis suggests that DK responses are erroneous indicators of existing states – a true value exists, but the respondent is unwilling to provide it because of question characteristics (e.g., length, difficulty) or because of having a satisficing personality (Krosnick, 2002). In this case, a DK response can be described as missing at random (MAR), because the probability of missingness is likely related to known characteristics and unrelated to the true answer (Allison, 2001; Little & Rubin, 2002).

Valid response. A third hypothesis about the underlying meaning of DK responses is that they indicate an actual lack of attitude or opinion. In this circumstance, a DK response is the true value, and it would make little sense to treat this response as missing data because it is a legitimate response option (De Leeuw, 2001). DK responses, in this case, may be meaningful when analyzed as a separate category. Staff et al. (2010), for example, found that adolescents who were unsure of their career aspirations in high school (i.e., said “don’t know”) had lower socio-economic attainment than their peers during early adulthood. Legitimate DKs may cause the most trouble for scale construction because it is not obvious (at least to us) how one should model uncertainty for some scale items when more concrete responses were obtained for others. For example, when a respondent answers the first three questions, then answers “DK” to the last question, how can we incorporate information about the items we know and the item we don’t know in one model?

Passive refusal. The final hypothesis about the underlying meaning of DK responses suggests that a true value exists, but that the respondent intentionally wishes to hide it from the researcher. Respondents may be unwilling to answer a question but not want to refuse outright; here the DK is best thought of as a passive refusal. There are a number of reasons why participants may consider their information too sensitive or private to reveal, particularly when

their answers could be socially uncomfortable to discuss. For example, someone may wish to hide publically unfavorable attitudes such as racist ideology or suicidal thoughts. Passive refusals are a dangerous type of DK; if actual responses are being disguised, then the DKs are missing not at random (MNAR), as the probability of missingness is related to the true value (Allison, 2001; Little & Rubin, 2002). For example, the probability of a DK response to a question about drug use would be related to having used drugs.

Each theory about the underlying meaning of DK responses has different implications for how DK responses should be handled. The theories are useful insofar as they can be correctly applied to a particular situation, and the best methods for handling DKs in that situation are known. In addition to figuring out the optimal methods for analyzing data with different types of DK responses, an equally important question concerns the consequences of applying the wrong theory and adjusting the DKs with the incorrect technique. To help answer these questions, we proceeded with a simulation model.

DATA AND METHOD

To gain insights into how DK responses to scale items should be handled, we simulated four DK response scenarios, applied six methodological approaches for scale construction, and assessed the accuracy of each of the six methods. As we constructed the simulated datasets, our goal was to create only realistic situations. Great care was taken to mimic the relationships found in the MIDUS data, and we continually asked the question, “Would this situation really occur?” For all four scenarios, we generated 10,000 cases on one dependent variable (Y), four independent variables (X scale items), and in some cases used a correlated “auxiliary” variable (Z).

Like the items found in the marriage and family values scale, our simulated items had a triangular distribution and ranged from 1-6, with 6 being the least likely response. Scale items were generated to have a reliability coefficient of .89. The descriptive statistics associated with the scale items were similar to those observed in the MIDUS data, as shown by the population (“true”) data prior to assigning DK responses in Table 11. The scale mean was 2.17 with a standard deviation of 1.14. We checked the correlations between items and their ability to predict one another in regression models to make sure the results were similar to what was observed in the MIDUS data. Once we were confident that the simulated dataset was a close match to real-world data, the next step was to assign DK responses in patterns that corresponded to four theoretical meanings: response style, satisficing, valid responses, and passive refusal.

Simulated DK Responses

Response style. In our first dataset, we assigned DK responses almost completely at random to simulate the theory that DKs are the result of a generalized response style – a random process where some people are simply more prone to using this response option. We gave each observation in each of the four scale items a number between 0 and 1, randomly selected from a uniform distribution wherein each number had an almost equal probability of being assigned. The probability of a DK response was unrelated to the outcome Y and unrelated to the X-item responses. Since it is highly unlikely that purely MCAR data occur in the real world (outside of an intentionally designed skip pattern), we made the probability of a DK response for each item correlated ($r=.01 - .10$) with the probability of a DK response to the other three items. The proportion of DK responses to each scale item ranged from 10% to 23%. Of the 10,000 responses, 5,168 were DK.

Satisficing. In the next dataset, DKs were assigned to simulate satisficing. Here, the probability of a DK response had a .4 correlation with an auxiliary variable Z. Satisficing is known to be related to question difficulty, question sensitivity, and to certain characteristics of a respondent's personality, for example. In the simulated data, Z could hypothetically represent any of these things. We gave each observation in each of the four scale items a number between 0 and 1, randomly selected from a uniform distribution that was correlated with Z. The correlations of the probability of a DK response and auxiliary variable Z ranged from .19 to .31 for each item. Assigning the DK responses in this way created a MAR situation. We tested models where Z was both related and unrelated to Y. The results had no practical differences, so we report here the model where Z is unrelated to Y (this also simplified the regression models we tested later). A correlation matrix from the MIDUS data was used to replicate the relationship between DK responses and observed responses to other items. These relationships were modest ($r=.03 - .16$), but important for ensuring a realistic scenario. The proportion of DK responses to each scale item again ranged from 10% to 23% and 3,736 of the 10,000 cases contained at least one DK to one scale item.

Valid responses. In the third simulation we modeled a situation where the DK was a legitimate, or valid, response category. The "true model" in this scenario is that the DK responses have a unique relationship with the outcome Y. This represents a situation where the presence of uncertainty itself is substantively interesting. The probability of a DK response to each scale item was correlated with the outcome Y ($r=.2$) as well as being correlated to the probability of a DK to other scale items ($r=.4-.6$). Assigning the DK responses in this way created a MAR situation but allowed the DK responses to have a relationship with Y that was

distinguishable from any other response category. The proportion of DK responses to each scale item ranged from 10% to 20%. Of the 10,000 cases, 3,726 were DKs.

Passive refusal. In the last dataset, DK responses were assigned in a MNAR pattern. In this scenario, the probability of a DK response was related to the true value of that response. For each scale item, a differential probability of saying “DK” was assigned according to the known response category. Cases with a true answer of “1” had a .02 probability of saying “DK,” whereas cases with a true answer of “6” had a .60 probability of saying “DK.” This mimics a situation where many, but certainly not all, participants said “DK” to a question as a way of hiding their socially undesirable but true response. To be realistic, it was important that only some of the respondents had a true value of “6” and the DK assignment was not deterministic. Even if the passive refusal hypothesis explained every DK response, we would still expect people to hide a range of answers for varying reasons. The overall proportion of DKs to each item was set to be much lower in this scenario than in the previous three models, ranging from 2% to 10%. Overall, 8,458 of the 10,000 responses had no DKs. We did test a model with same level of DKs that occurred in the first three scenarios, but the data structure was very unrealistic, and the results from combining the scale techniques were outrageously bad. We also prefer the model with few DK responses because it illustrates that even relatively small amounts of missing data are not necessarily negligible.

Combined Scale Items

For each of the four simulated datasets (response style, satisficing, valid response, and passive refusal), we treated the DK responses with six options for creating the averaged scale. In the complete cases method, only cases with non-DK responses to all four scale items were summed, and then divided by four. In the second method, DK responses were recoded to equal

3.5, representing a neutral or middle category. DK responses in the directional coding method were recoded to equal 6. In the available item average technique, scale items without a DK response are averaged and the DKs are ignored. For example, if a respondent said “DK” to the first and second items, but gave dis-/agreement responses for the third and fourth items, only items three and four would be averaged to compute the scale score.

For the fifth method, dummy variable adjustment, DK responses were replaced with the mean for each item. A dichotomous indicator variable was created to identify whether any of the items had a DK response. An alternative strategy, particularly if the research question aims to capture the degree of uncertainty, might be to create an indicator for the number of DKs. We did not test this unconventional strategy here, but it may be one approach for modeling valid responses to some items while DKs were given to others.

In the MI approach, we imputed the DKs for individual scale items and then averaged the scale. The auxiliary variable Z, and the dependent variable Y, contained no missing values and were included in all imputation models. Imputations were done in Stata ICE; 25 datasets were generated under an ordinal logistic model (to accommodate the triangular distribution of the X variables). An alternative strategy to imputing each scale item and then creating the scale would be to impute the combined scale itself. This strategy could be useful in a large imputation model where including individual items is not a feasible strategy (Graham, 2009). It is ideal, however, to impute each item and then combine the scale (Graham, 2009).

RESULTS

We evaluated the six different methods for combining scale items in the presence of DK responses in two ways. First, in Table 12, we show the scale properties resulting from each of the computation methods, under the four conditions of DK responses. The “true” model shows the

scores that existed in the dataset with no missing values. Methods which perform well for combining the scale should closely match the true scale. For the dataset simulating DKs as a valid response, no “true” scale is listed. It is not entirely clear what the true model is, in this situation, because categorical responses were obtained for some of the items but DKs were received for others. Should respondents who ever said “DK” be moved into a DK category and excluded from the true scale? If so, then the complete case model is the “true” scale. One DK carrying more weight than three valid responses, however, seems problematic because the respondent was more certain about the entire scale than not. Perhaps the “true” model, then, is the available item average, which is calculated from all answers that were not DKs.

The second strategy we used to evaluate the performance of the six methods was to test the scale as a dependent variable in a linear regression model. For each of the four simulated datasets, and each of the six scale computation methods, we regressed a dependent variable Y on the combined scale X. For ease of interpretation, we set the “true” b-coefficient to always equal 1.0; a coefficient larger than 1.0 indicates a positive bias, and anything smaller than 1.0 indicates a negative bias. The results of these models are shown in Table 13. The true model when DKs are a valid response category is estimated by two regression models. In the first model, the scale effect is estimated by the equation:

$$y = a + b_1x_{complete\ case\ scale}$$

where the complete case scale excludes any respondent to express DK to any of the items. In the second model, the effect of DK responses compared to all others is estimated by the equation:

$$y = a + b_2x_{DK\ indicator}$$

where the DK indicator is a dichotomous variable for presence of a DK (1) versus all other responses (0). In reality, we would need to be able to use both measures in a regression model if

any other independent variables were to be estimated, which is a more complicated situation than we simulate here. Our model concerns two primary questions. First, what are the consequences of ignoring the fact that DKs are a legitimate category when they are a distinct group? Second, does dummy variable adjustment “account for” the DK group, as many researchers claim? Our model should identify which method produces the most correct b-coefficient for the scale, but ignores the point that we could also really need a b-coefficient for the DK group. Methods for simultaneously estimating dis-/agree responses and DK responses are not evaluated here.

When DKs were a Response Style

Perhaps the best case scenario for DK responses is that they result largely from a random generalized response style. In this case, DK responses would be MCAR, and could be handled with precision. As shown in Table 12, complete case analysis and MI produced scale properties that were closely matched to the true scale. MI produced the correct mean, standard deviation, and Chronbach’s alpha. Available item averaging was also an acceptable strategy in this situation and produced results that were comparable to the true scale. Complete case analysis and MI, however, were slightly more accurate.

Neutral/middle category coding, directional coding, available item averaging, and using mean substitution (the DK indicator) were poor methods. Neutral/middle category coding, where the DKs were all coded as 3.5, produced incorrect scale properties. Far too few respondents were estimated to ‘*Strongly Agree*’, the scale mean was overestimated, the standard deviation was underestimated, and Chronbach’s alpha was too small. Directional coding, where the DKs were all coded as 6, severely distorted the scale distribution. For example, 15.75% of respondents were estimated to ‘*Strongly Agree*’, compared to the actual 30.93%. Neutral/middle category coding produced too high of a scale mean (2.82 rather than 2.17) and too large of a

standard deviation (1.28 rather than 1.15). Chronbach's alpha for this strategy was .59 – quite different from the true alpha of .84. Mean substitution, naturally, produced the correct mean. The scale distribution, the standard deviation, and Chronbach's alpha, however, were unacceptable.

Table 13 shows how the scale performed as a predictor of Y; recall that the true coefficient is 1.0. There was little bias in the b-coefficients when only the complete items were analyzed (.983) or the DKs were recoded to a middle category (1.010). No bias was present in the MI model as the b-coefficient was 1.000. Although the complete-item and neutral-recoding methods were not biased, they were inefficient compared to the MI model. The standard error of the complete case model, for example, was .038, compared to a standard error of .028 when multiple imputation was used. Dummy variable adjustment is widely used, despite a large body of research which has demonstrated that it is a biased method when a true value exists, even when the data are MCAR. Our simulation is consistent with this finding. We were surprised to find that available item averaging is a less-than-optimal method (even under the ideal MCAR situation), since it seems such an intuitive approach. The dummy variable adjustment method produced a biased coefficient (1.106). Directional recoding was a terrible strategy for this type of DK response and resulted in a considerably biased coefficient (.636).

When DKs were Satisficing

The trendy theory of DK responses suggests that satisficing is the primary underlying cause and therefore the DKs are MAR. This scenario also appears to be one that we are well-equipped to deal with. If respondents have a true answer but are just too bored or lazy to give it to us, then treating the DKs as missing data and using MI can produce reasonable results. As shown in Table 12, complete case analysis and MI gave good approximations of the scale mean, standard deviation, and Chronbach's alpha, although MI generally performed slightly better. The

MI estimate of the scale mean (2.18) was closer to the true mean (2.17) than was the complete case analysis estimate (2.25). Neither method exactly reproduced the true scale distribution, but both estimates were quite close, with MI being slightly more accurate than complete case analysis. Available item averaging was an acceptable strategy, but not as good as MI.

Neutral/middle category coding distorted the scale distribution by underestimating the number of '*Strongly dis-/agree*' responses. The scale mean, in this case, was too high and the standard deviation was underestimated. Directional coding largely overestimated the number of respondents who said '*Strongly Disagree*', 5.53%, compared to the true scale of 1.48%. This strategy produced a scale mean that was too high, a standard deviation that was too large, and Chronbach's alpha that was too small. Mean substitution did not perform especially rotten, but did distort the scale distribution and underestimate the standard deviation. For example, 30.93% of responses in the true scale fell in the '*Strongly Agree*' range, but only 21.89% of responses did so in the mean substitution model.

The bivariate results, in Table 13, show that MI was a more efficient strategy than complete case analysis, as reflected in the standard errors. For researchers working with small sample sizes, in particular, there is much to be gained by using MI. Neutral-recoding, available item averaging, and dummy variable adjustment methods did not perform especially poorly, but were also not ideal. Enough bias is present in the coefficients from these models to suggest that avoiding these methods in the future is the prudent thing to do. Perhaps the bias is not so severe that past studies to utilize these methods need necessarily be discarded. Directional recoding, on the other hand, is a worrisome approach; the true coefficient of 1.0 comes nowhere near within the confidence interval of the biased, but statistically significant, directional recoding coefficient of .489.

When DKs were a Valid Response

The historic theory of DK responses, and the reason DK was initiated as a response option to begin with, is that people do not have attitudes and opinions about everything. To model this, many researchers prefer to impute the DKs with the item mean and to include a dummy variable for whether or not a DK was ever recorded. The benefit to this approach is that it allows a coefficient for the uncertain respondents to be estimated separately from those who gave a dis-/agree response. This is an important feature because uncertain respondents may be a unique and interesting group of people (De Leeuw, 2001). Unfortunately, however, our inquiries into this type of DK have raised more questions than we have been able to answer. How can a “true” scale be computed when some of the items are dis-/agree responses and others are DKs? We don’t know. What we can say is that conventional strategies do not answer this question either. So, we compare the strategies here, but do not comment on which one is the most correct.

As shown in Table 12, complete case analysis, available item averaging, and MI estimated very similar scale distributions. For example, complete case analysis showed 28.55% of responses in the ‘*Strongly Agree*’ category, available item averaging placed 31.1% there, and MI placed 30.13% in that category. The scale means were similar, though not identical, as were the standard deviations. Chronbach’s alpha was .84 for all three methods. Neutral/middle category coding and directional coding probably suffered from the same problems exposed in simulations for DKs as a result of response style and satisficing. Compared to MI, for example, these strategies produced lower estimates of responses in the ‘*Strongly Agree*’ category.

For the regression models shown in Table 13, we set the “true” b-coefficient to be equal to the complete case analysis model, so deviations from 1.0 are differences between a particular method and the complete case model. Although this does not tell us which strategy is “right,” it

is useful to see that some strategies produce remarkably different coefficients from the complete case model. For example, the b-coefficient for directional recoding was .697, with 1.0 falling nowhere near the confidence interval around this estimate. It is likely that researchers would have reached the same substantive conclusions regardless of whether complete case analysis, neutral/middle category coding, dummy variable adjustment, or MI had been used. Is it possible that the dummy variable adjustment model, which contains the indicator for a DK response, is the most correct? It is possible, although it is probably not the case. The true b-coefficient of Y regressed on the DK indicator was -.325. The b-coefficient for the DK indicator in the dummy variable adjustment model was -.998, a terrible estimate of the true difference between those who said “DK” and those who did not.

When DKs were a Passive Refusal

The worst-case scenario for DK responses is that participants systematically provide this answer with the intention of masking their true response. The data would be MNAR, a sticky situation that even modern methods cannot always tidy up. MI assumes, for example, that the data are MAR. Indeed, MI did not outperform the other methods; complete case analysis, neutral/middle category coding, available item averaging, mean substitution, and MI all produced poor scale approximations. All five methods grossly underestimated the proportion of ‘*Strongly Disagree*’ responses, produced too low of a scale mean, and too low of a standard deviation. Directional coding appeared to be a perverse method in the previous models, but it produced the best estimates of the scale properties. The true mean was 2.17, for example, and the mean estimated with directional coding was 2.19.

The regression model results also show that passive refusals are likely to cause trouble when they occur. Complete case analysis, neutral/middle category coding, available item

averaging, and MI produced biased results. Neutral coding produced the most biased b-coefficient of 1.238. Although MI is often robust to violations of the MAR assumption, in this case it was not, as a coefficient of 1.159 can hardly be described as unbiased. Keep in mind that the level of missing data in this simulation was the lowest of all the models, with DKs to each item ranging from 2% to 10%. This makes it all the more alarming that these commonly used methods, and even the methods we believe should work well, performed so poorly.

Directional coding produced the least biased coefficient in the NMAR situation. Even so, this approach is still a worrisome one. When DK responses were the result of a response style, satisficing, or a valid response, directional coding was an incredibly biased strategy. If a researcher assumed the DKs were passive refusals, but this assumption was incorrect, the consequences for the data analysis would be severe. Although we agree with reviewer comments that it is rather unsatisfying to point out a problem that we do not have a solution for, we believe this was one of the most important parts of the paper because it identified a necessary avenue for future research and can help researchers exercise due caution with DK responses to sensitive questions.

CONCLUSION

The goal of our paper was to convince family researchers that DK responses deserve attention. DK responses are prevalent in observed data, as we showed in Table 9. Alternative strategies for computing a scale in the presence of DK can produce quite different results, which we demonstrated in Table 10. We discussed the leading theories of non-response which suggest that DK responses result from one of four primary causes: a completely random process where DK responses are a generalized response style, survey satisficing, actual lack of information (a true DK), and passive refusal. We constructed four simulated datasets to represent each of these

situations. Next, we applied family research's six most common methods for treating DK responses to each type of DK response.

There are several strategies we now know for analyzing data from respondents who "don't know." Unless respondents are intentionally trying to hide their answers, MI works well to treat DK responses to scale items, even when DKs are a valid response. Complete case analysis generally produces unbiased coefficients but is an inefficient strategy. Directional coding can cause some serious problems unless the majority of the respondents are hiding their "true" response in the direction the researcher selects. If a researcher is supremely confident that most respondents are masking their true responses behind a passive refusal, directional coding may be an option worth exploring. Neutral/middle category recoding, available item averaging, and dummy variable adjustment should be avoided, because in all cases they were either outperformed by or no different from MI estimates.

Table 9. Observed Scale Item Frequencies

<i>Response category</i>	<i>Scale item (%)</i>			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
(1) Strongly agree	37.1	33.4	32.8	31.4
(2) Agree	29.6	26.9	27.9	27.0
(3) Somewhat agree	8.3	8.8	10.0	10.2
(4) Somewhat disagree	3.3	5.0	5.6	5.3
(5) Disagree	6.1	7.8	5.9	6.5
(6) Strongly disagree	3.6	6.0	4.7	5.4
(8) Don't know	12.0	12.2	13.2	14.2
<i>Total</i>	<i>100.00</i>	<i>100.00</i>	<i>100.00</i>	<i>100.00</i>

Table 10. Observed Scale Items Combined Using 6 Strategies for Handling DK Responses

<i>Scale properties</i>	<i>Complete case analysis</i>	<i>Neutral/middle category</i>	<i>Directional coding</i>	<i>Available item average</i>	<i>DK indicator</i>	<i>Multiple imputation</i>
Frequencies (%)						
1 - 1.4 (<i>Strongly agree</i>)	32.22	25.63	25.63	29.82	25.98	29.81
1.5 - 2.4	33.07	29.75	26.56	31.99	38.7	31.93
2.5 - 3.4	16.07	19.68	14.83	17.22	17.99	17.70
3.5 - 4.4	10.23	17.05	13.83	11.27	10.34	11.52
4.5 - 5.4	5.20	5.32	9.47	5.88	4.43	5.80
5.5 - 6 (<i>Strongly disagree</i>)	3.21	2.56	9.67	3.82	2.55	3.25
Mean	2.22	2.72	2.43	2.33	2.28	2.32
Standard deviation	1.27	1.59	1.25	1.35	1.18	1.29
Alpha	0.88	0.88	0.88	0.88	0.87	0.88
Sample size	2,365	3,008	3,008	2,857	3,008	3,008

Table 11. Simulated Scale Item Frequencies

<i>Response category</i>	Scale item (%)			
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
(1) Strongly agree	41.92	37.01	53.31	36.32
(2) Agree	33.09	32.72	28.46	27.96
(3) Somewhat agree	9.93	11.39	11.87	10.56
(4) Somewhat disagree	3.98	6.52	2.91	8.72
(5) Disagree	6.85	6.79	1.93	9.29
(6) Strongly disagree	4.23	5.57	1.52	7.15
<i>Total</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>

Table 12. Simulated Scale Items Combined Using 6 Strategies for Handling DK Responses

<i>Source of the DK</i>	<i>True scale</i>	<i>Complete case analysis</i>	<i>Neutral/middle category</i>	<i>Directional coding</i>	<i>Available item average</i>	<i>DK indicator</i>	<i>Multiple imputation</i>
Response style							
Frequency (%)							
1 - 1.4 (<i>Strongly agree</i>)	30.93	30.48	15.75	15.75	33.49	23.41	30.95
1.5 - 2.4	35.26	35.84	42.30	26.18	34.29	45.94	35.93
2.5 - 3.4	18.12	18.34	25.80	24.74	16.54	18.19	17.39
3.5 - 4.4	9.57	9.42	11.29	19.68	9.22	8.56	9.25
4.5 - 5.4	4.64	4.33	4.04	10.13	4.66	3.08	4.88
5.5 - 6 (<i>Strongly disagree</i>)	1.48	1.59	0.82	3.52	1.80	0.82	1.60
Mean	2.17	2.17	2.39	2.82	2.16	2.17	2.17
Standard deviation	1.15	1.14	1.03	1.28	1.18	0.99	1.15
Alpha	0.84	0.84	0.75	0.59	0.83	0.78	0.84
Sample size	10,000	5,168	10,000	10,000	9,944	10,000	10,000
Satisficing							
Frequency (%)							
1 - 1.4 (<i>Strongly agree</i>)	30.93	28.59	17.91	17.91	30.83	21.89	29.02
1.5 - 2.4	35.26	34.98	36.02	25.89	34.95	46.46	36.33
2.5 - 3.4	18.12	19.06	27.10	20.35	17.41	18.16	17.95
3.5 - 4.4	9.57	10.46	13.64	18.49	9.69	8.66	9.89
4.5 - 5.4	4.64	5.25	4.29	11.83	5.24	3.79	5.14
5.5 - 6 (<i>Strongly disagree</i>)	1.48	1.66	1.04	5.53	1.87	1.04	1.67
Mean	2.17	2.25	2.44	2.87	2.21	2.22	2.18
Standard deviation	1.15	1.17	1.07	1.41	1.19	1.03	1.17
Alpha	0.84	0.84	0.78	0.73	0.84	0.80	0.84
Sample size	10,000	6,264	10,000	10,000	9,739	10,000	10,000

Table 12. (Continued)

<i>Source of the DK</i>	<i>True scale</i>	<i>Complete case analysis</i>	<i>Neutral/middle category</i>	<i>Directional coding</i>	<i>Available item average</i>	<i>DK indicator</i>	<i>Multiple imputation</i>
Valid response							
Frequency (%)							
1 - 1.4 (<i>Strongly agree</i>)	--	28.55	17.91	17.91	31.10	21.79	30.13
1.5 - 2.4	--	34.81	36.27	25.72	34.57	46.45	35.94
2.5 - 3.4	--	19.30	26.96	20.91	17.61	18.27	17.83
3.5 - 4.4	--	10.47	13.57	18.23	9.55	8.70	9.62
4.5 - 5.4	--	5.20	4.24	11.60	5.20	3.74	4.90
5.5 - 6 (<i>Strongly disagree</i>)	--	1.67	1.05	5.63	1.96	1.05	1.57
Mean	--	2.25	2.44	2.86	2.22	2.22	2.19
Standard deviation	--	1.17	1.07	1.41	1.20	1.02	1.17
Alpha	--	0.84	0.78	0.72	0.84	0.80	0.84
Sample size	--	6,274	10,000	10,000	9,742	10,000	10,000
Passive refusal							
Frequency (%)							
1 - 1.4 (<i>Strongly agree</i>)	30.93	36.44	30.82	30.82	31.45	30.91	31.13
1.5 - 2.4	35.26	41.26	37.03	34.99	38.79	40.48	37.85
2.5 - 3.4	18.12	17.05	21.86	17.66	19.91	22.13	19.69
3.5 - 4.4	9.57	4.56	9.33	9.69	7.53	5.78	8.28
4.5 - 5.4	4.64	0.63	0.91	5.03	1.99	0.65	2.73
5.5 - 6 (<i>Strongly disagree</i>)	1.48	0.06	0.05	1.81	0.33	0.05	0.32
Mean	2.17	1.83	2.04	2.19	2.03	1.97	2.05
Standard deviation	1.15	0.8	0.93	1.18	0.96	0.84	0.97
Alpha	0.84	0.76	0.81	0.83	0.76	0.75	0.82
Sample size	10,000	8,458	10,000	10,000	9,951	10,000	10,000

Table 13. Linear Regression Results for Models using Simulated Scale Items

<i>Source of the DK</i>	Strategies for handling "Don't Know" responses											
	Complete case analysis		Neutral / middle category		Directional recode		Available item average		Dummy variable adjustment		Multiple imputation	
	<i>B</i>	<i>SE(B)</i>	<i>B</i>	<i>SE(B)</i>	<i>B</i>	<i>SE(B)</i>	<i>B</i>	<i>SE(B)</i>	<i>B</i>	<i>SE(B)</i>	<i>B</i>	<i>SE(B)</i>
Response style												
4-item scale	0.983	0.038	1.010	0.031	0.636	0.025	0.951	0.027	1.106	0.032	1.000	0.028
DK indicator									-0.092	0.063		
Satisficing												
4-item scale	0.997	0.034	0.929	0.030	0.489	0.023	0.931	0.027	1.060	0.031	0.981	0.028
DK indicator									-0.102	0.065		
Valid response												
4-item scale	--	--	0.962	0.067	0.697	0.028	0.917	0.027	1.048	0.031	0.962	0.028
DK indicator									-0.998	0.065		
Passive refusal												
4-item scale	1.187	0.042	1.238	0.034	0.968	0.027	1.147	0.033	1.059	0.039	1.159	0.033
DK indicator									0.521	0.103		

Chapter 5

What We Now Know About Don't Know Responses

This dissertation contributes to the body of knowledge on don't know responses in several ways. The first article confirmed and challenged some of the longstanding assumptions that survey research has made about this type of response. By using 12 nationally representative data sources, these findings could be generalized beyond a single data collection effort. In particular, DK responses were found to be most frequently recorded for sensitive questions, to questions that occurred in the middle of a survey, and to questions which asked about politics and government or finances and economics. Although women offered DK responses more often than men did, men were more likely than women to say DK to particular types of questions, such as those asking about family and friends or health and health care. With an accurate assessment of who offers DK responses, and under what circumstances people are mostly likely to do so, we can begin to disentangle the range of meanings that underlie these responses.

The second article of this dissertation explored DK responses in longitudinal data. This was a contribution to survey research because most prior research has been done on cross-sectional data only. One of the most important findings from analyzing DK responses over time was that many respondents offered DK replies to the same questions asked nine to ten years apart. This is strong inferential evidence that while some DKs may result from survey satisficing, may represent passive refusals, or may reflect a generalized response style, many of these replies also indicate genuine lack of knowledge. Although it was beyond the scope of this research to experiment with questionnaire design, confirmation that many DKs are "legitimate," or the best assessment of a respondent's condition, has serious implications for the steps that should (or should not) be taken to reduce these responses.

If a DK reply is, indeed, the best answer to a question, it seems prudent and perhaps necessary to allow these responses in data collection. At the same time, DK responses can pose a serious challenge to data analysts. Many of the conventional methods researchers have used to analyze these responses have gone without serious empirical scrutiny. The third article of this dissertation explored six common methods for analyzing these responses when different meanings were assumed to underlie them. In general, if researchers are uninterested in analyzing those who said DK as a categorically different type of response, and instead prefer to treat DK responses as missing values, modern missing data methods were able to treat these responses with minimal bias. The exception to this finding occurred when DK responses represented “passive refusals,” where respondent’s intentionally withheld true answers from the researcher. Under this circumstance, no method for analysis emerged without limitations.

Although it is somewhat unsatisfying to conclude that we, as a scientific discipline, have not identified appropriate methods for handling DK responses that reflect passive refusals, it is nonetheless an important point to consider in questionnaire design. If survey respondents are going to continue hiding responses from researchers – a likely scenario as privacy concerns have existed as long as survey research has – allowing participants the option to mask their condition behind a DK response allows their response to be distinguished from other response categories. That is, a DK option allows researchers to identify which responses are problematic, compared to a forced-choice option where outright lies could not be detected.

An important theme that emerged from each article is that DK responses are much more nuanced, meaningful, and context-based than perhaps we recognized in the past. Unlike the responses we hope to record, correct responses in the appropriate metric, DK responses result from a complicated interaction of survey, question, and individual factors. Many meanings

underlie these responses; these meanings vary from one respondent to the next and even within a single participant from one question to the next. This dissertation has shown that there are many exciting aspects of uncertain responses that we still don't know.

References

- ABC News. (2005). ABC News Nukes Poll [computer file]. In T. N. S. I. [producer] (Ed.): Inter-university Consortium for Political and Social Research [distributor].
- Acock, A. C. (2005). Working With Missing Values. *Journal of Marriage and Family*, 67, 1012-1028.
- Allison, P. D. (2001). *Missing Data*. Thousand Oaks: Sage Publications.
- Anderson, K. L. (2008). Is partner violence worse in the context of control? *Journal of Marriage and Family*, 70, 1157-1168.
- Ares, G., & Gambaro, A. (2007). Influence of Gender, Age and Motives Underlying Food Choice on Perceived Healthiness and Willingness to try Functional Foods. *Appetite*, 49, 148-158.
- Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, Nay-saying, and Going to Extremes: Black-White Differences in Response Styles. *Public Opinion Quarterly*, 48, 491-509.
- Baldwin, J. D., & Baldwin, J. I. (1997). Gender Differences in Sexual Interest. *Archives of Sexual Behavior*, 26, 181-209.
- Beatty, P. C., & Herrmann, D. (2002). To Answer or Not to Answer: Decision Processes Related to Survey Item Non-Response. In R. M. Groves (Ed.), *Survey Nonresponse*.
- Berinsky, A. (1999). The Two Faces of Public Opinion. *American Journal of Political Science*, 43, 1209-1230.
- Bishop, G. F., Oldendick, R. W., & Tuchfarber, A. J. (1984). What Must my Interest in Politics be if I just told you "I don't know"? *Public Opinion Quarterly*, 48, 510-179.
- Bishop, G. F., Tuchfarber, A. J., & Oldendick, R. W. (1986). Opinions on Fictitious Issues: The Pressure to Answer Survey Questions. *Public Opinion Quarterly*, 50, 240-250.

- Bogart, L. (1967). No Opinion, Don't Know, and Maybe No Answer. *Public Opinion Quarterly*, 31, 331-345.
- Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., et al. (1996a). National Survey of Midlife Development in the United States (MIDUS), 1995-1996 [computer file]: Inter-university Consortium for Political and Social Research [distributor].
- Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., et al. (1996b). National Survey of Midlife Development in the United States (MIDUS), 1995-1996: MIDUS Sample Descriptions. *Inter-University Consortium for Political and Social Research*
- Catania, J. A., Binson, D., Canchola, J., Pollack, L. M., Hauck, W., & Coates, T. J. (1996). Effects of Interviewer Gender, Interviewer Choice, and Item Wording on Responses to Questions Concerning Sexual Behavior. *Public Opinion Quarterly*, 60, 345-375.
- CBS News. (2009). CBS News Monthly Poll, August 2009 [Computer file]: Inter-university Consortium for Political and Social Research [distributor].
- Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6, 330-351.
- Converse, J. M. (1976). Predicting No Opinion in the Polls. *Public Opinion Quarterly*, 40, 515-530.
- Converse, P. E. (1964). The Nature of Belief Systems in Mass Publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 206-261). London: Free Press of Glencoe.

- Converse, P. E. (1970). Attitudes and Non-Attitudes: Continuation of a Dialogue. In E. Tufte (Ed.), *The Quantitative Analysis of Social Problems* (pp. 168-183). Reading: Addison-Wesley.
- Couper, M. P. (2008). Technology and the Survey Interview/Questionnaire. In M. F. Schober & F. G. Conrad (Eds.), *Envisioning the Survey Interview of the Future*. New York: Wiley.
- Couper, M. P., & Nicholls, W. L. (1998). The History and Development of Computer Assisted Survey Information Collection *Computer Assisted Survey Information Collection*. New York: Wiley
- Culpepper, I. J., Smith, W. R., & Krosnick, J. A. (1992). *The Impact of Question Order on Satisficing in Surveys*. Paper presented at the Midwestern Psychological Association.
- De Leeuw, E. D. (2001). Reducing Missing Data in Surveys: An Overview of Methods. *Quality and Quantity*, 32, 147-160.
- DeMaio, T. J. (1980). Refusals: Who, Where, and Why. *Public Opinion Quarterly*, 44, 223-233.
- Duncan, O., & Stenbeck, M. (1988). No Opinion or Not Sure. *Public Opinion Quarterly*, 52, 513-525.
- Durand, R. M., Guffey, H. J., & Planchon, J. M. (1983). An Examination of the Random Versus Nonrandom Nature of Item Omissions. *Journal of Marketing Research*, 20, 305-313.
- Durand, R. M., & Lambert, Z. V. (1988). Don't Know Responses in Surveys: Analyses and Interpretational Consequences. *Journal of Business Research*, 16, 169-188.
- Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford Press.
- Feick, L. F. (1989). Latent Class Analysis of Survey Questions That Include Don't Know Responses. *Public Opinion Quarterly*, 53, 525-547.

- Francis, J. D., & Busch, L. (1975). What We Now Know About 'I Don't Knows'. *Public Opinion Quarterly*, 39, 207-218.
- Gelman, A. (2007). Struggles with Survey Weights and Regression Modeling. *Statistical Science*, 22, 153-164.
- Giljam, M., & Granberg, D. (1993). Should We Take Don't Know for an Answer? *Public Opinion Quarterly*, 57, 348-357.
- Goettems, M. L., Ardenghi, T. M., Romano, A. R., Demarco, F. F., & Torriani, D. D. (2010). Influence of maternal dental anxiety on oral health-related quality of life preschool children. *Quality of Life Research*, 20, 951-959.
- Graham, J. W. (2009). Missing Data Analysis: Making it Work in the Real World. *Annual Review of Psychology*, 60, 549-576.
- Groves, R. M., & Couper, M. P. (1998). How Survey Design Features Affect Participation *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004a). *Survey Methodology*. Hoboken: John Wiley and Sons.
- Groves, R. M., Presser, S., & Dipko, S. (2004b). The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly*, 68, 2-31.
- Hardy, M. (1993). *Regression with Dummy Variables*. Newbury Park: Sage Publications.
- Hemenway, D. (1999). National Firearms Survey, 1999 [computer file]: Inter-university Consortium for Political and Social Research [distributor].
- Hui, H., & Triandis, H. C. (1985). The Instability of Response Sets. *Public Opinion Quarterly*, 49, 253-260.

- Johnson, D. R., & Young, R. (2011). Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations. *Journal of Marriage and Family*, *Forthcoming*.
- 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. (2002). The Causes of No-opinion Responses to Attitude Measures in Surveys: They are Rarely what they Appear to be. In R. M. Groves, D. A. Dillman, J. L. Elting & R. J. Little (Eds.), *Survey Nonresponse* (pp. 88-100). New York: Wiley.
- Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., Kopp, R. J., et al. (2002). The Impact of "No Opinion" Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice. *Public Opinion Quarterly*, *66*, 371-403.
- Linn, M. C., De Benedictis, T., Delucchi, K., Harris, A., & Stage, E. (1987). Gender differences in national assessment of educational progress in science items: what does "I don't know" really mean? *Journal of Research in Science Teaching*, *24*(3), 267-278.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken: John Wiley & Sons, Inc.
- Maume, D. J. (2006). Gender differences in restricting work efforts because of family responsibilities. *Journal of Marriage and Family*, *68*, 859-869.
- Mondak, J. J., & Anderson, M. R. (2004). The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge. *Journal of Politics*, *66*, 492-512.
- Mondak, J. J., & Davis, B. C. (2001). Asked and Answered: Knowledge Levels when we will not take "don't know" for an Answer. *Political Behavior*, *23*, 199-224.

- Peytchev, A., Baxter, R. K., & Carley-Baxter, L. R. (2009). Not all survey effort is equal: Reduction of nonresponse bias and nonresponse error. *Public Opinion Quarterly*, 73, 785-806.
- Poe, G. S., Seeman, I., McLaughlin, J., Mehl, E., & Dietz, M. (1988). 'Don't Know' Boxes in Factual Questions in a Mail Questionnaire. *Public Opinion Quarterly*, 52, 212-223.
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata: Second Edition*. College Station: Stata Press.
- Rapoport, R. B. (1979). What they don't know can hurt you. *American Journal of Political Science*, 23, 805-815.
- Rapoport, R. B. (1985). Like Mother, Like Daughter: Intergenerational Transmission of DK Response Rates. *Public Opinion Quarterly*, 49, 198-208.
- Riphahn, R. T., & Serfling, O. (2002). Item Non-response on Income and Wealth Questions, *Working paper*. Switzerland: University of Basel.
- Rossi, A. S. (2001). *Caring and doing for others: Social responsibility in the domains of family, work, and community*. Chicago: University of Chicago Press.
- Roth, P. L., Switzer, F. S., & Switzer, D. M. (1999). Missing data in multiple item scales: A monte carlo analysis of missing data techniques. *Organizational Research Methods*, 2, 211-232.
- Rubin, D. B., Stern, H., & Vehovar, V. (1995). Handling 'Don't Know' Survey Responses: The Case of the Slovenian Plebiscite. *Journal of the American Statistical Association*, 90, 822-828.
- Sanchez, M. E., & Morchio, G. (1992). Probing "Don't Know" Answers: Effects on Survey Estimates and Variable Relationships. *Public Opinion Quarterly*, 56(4), 454-465.

- Schafer, J. L., & Graham, J. W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods*, 7(2), 147-177.
- Schuman, H., & Presser, S. (1980). Public Opinion and Public Ignorance: The Fine Line Between Attitudes and Non-Attitudes. *American Journal of Sociology*, 85, 1214-1225.
- Schuman, H., & Presser, S. (1981). *Questions and Answers in Attitude Survey Experiments on Question Form, Wording, and Context*. New York: Academic Press.
- Shmotkin, D. (1992). The Apprehensive Respondent: Failing to Rate Future Life Satisfaction in Older Adults. *Psychology and Aging*(484-486).
- Shoemaker, P. J., Eichholz, M., & Skewes, E. A. (2000). Item nonresponse: distinguishing between don't know and refuse. *International Journal of Public Opinion Research*, 14(2), 193-201.
- Sicinski, A. (1970). Don't Know Answers in Cross-National Surveys. *Public Opinion Quarterly*, 34, 126-129.
- Smyth, J. D., Dillman, D. A., Melani Christian, L., & Stern, M. J. (2006). Comparing Check-All and Forced-Choice Question Formats in Web Surveys. *Public Opinion Quarterly*, 70, 66-77.
- Staff, J., Harris, A., Sabates, R., & Briddell, L. (2010). Uncertainty in early occupational aspirations: Role exploration or aimlessness? *Social Forces*, 89, 1-25.
- Taris, T. W. (2000). *A primer in longitudinal data analysis*. Thousand Oaks: Sage.
- Tourangeau, R., & Smith, T. W. (1996). Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly*, 60, 275-304.

- Van Es, J., Lorence, D., Morgan, G., & Church, J. (1996). Don't Know Responses in Environmental Surveys. *Journal of Environmental Education, 27*, 13-18.
- von Hippel, P. (2009). How to Impute Interactions, Squares, and Other Transformed Variables. *Sociological Methodology, 39*(1), 265-291.
- Wagner, J. (2010). The Fraction of Missing Information as a tool for Monitoring the Quality of Survey Data. *Public Opinion Quarterly, 74*, 1-21.
- Wenner, L. (1998). *Mediasport*. New York: Routledge.
- Winship, C., & Radbill, L. (1994). Sampling Weights and Regression Analysis. *Sociological Methods and Research, 23*(2), 230-257.
- Wright, D. L., Aquillino, W. S., & Supple, A. J. (1998). A Comparison of Computer-Assisted Paper-and-Pencil Self-Administered Questionnaires in a Survey on Smoking, Alcohol, and Drug Use. *Public Opinion Quarterly, 62*, 331-353.
- Ziller, R. C., & Long, B. H. (1965). Some Correlates of the Don't-Know Response in Opinion Questionnaires. *Journal of Social Psychology, 1965*(67).

Appendix A

Question Wording for Individual-level Characteristics

AGE

“What is your age?” (Abcnuke)

“What is your age?” (Abcpoll)

“How old are you?” (Cbcpoll)

“What is your age?” (Century)

“In what year were you born?” (Cid)

“What is your age?” (Courts)

“What is your age?” (Firearm)

“May I please have your age?” (Hints)

Exact wording unavailable: “Stated age was obtained from a household informant during initial phone screening.” (Midus)

“How old were you on your last birthday?” (Prisons)

“What is your age?” (Social)

EDUCATION

“What was the last grade of school you completed?” [If graduated college] “Was that an associate’s degree, a bachelor’s degree, or what?” [8th grade or less, some high school, graduated high school, some college (ask if technical school; if yes, code as high school graduate), graduated college, post graduate] (Abcnuke)

“What was the last grade of school you completed?” [If graduated college] “Was that an associate’s degree, a bachelor’s degree, or what?” [8th grade or less, some high school, graduated high school, some college (ask if technical school; if yes, code as high school graduate), graduated college, post graduate] (Abcpoll)

“What was the last grade of school you completed? Did you complete the degree?” [Not a high school graduate, high school graduate, some college, college graduate, post grad work or degree] (Cbcpoll)

“What is the highest grade of school or year of college you have completed – less than high school diploma, high school grad, trade/vocational school, some college (no degree or

associate's degree), bachelor's degree (BA or BS), some graduate school (no degree), graduate level degree?" (Century)

"Please look at this card and tell me what number represents the highest grade of school or year of college you have completed." [Showcard with responses: none, or grade 1-9; high school incomplete (grades 9-11); high school graduate; GED; business, technical, or vocational school after high school; some college, no 4-year degree; college graduate; post-graduate training or professional schooling after college.] (Cid)

"What is the highest level of education you have completed – less than fifth grade, less than eighth grade, less than high school, high school diploma or GED, some college or technical training school, associate's degree, bachelor's degree, some graduate school, graduate or professional degree?" (Courts)

"What is the highest level of education you have completed – some high school or high school graduate, some college or associate's degree, bachelor's degree, graduate study or degree?" (Firearms)

"What was the highest level of school you completed – never attended school or only attended nursery school/kindergarten; grades 1 through 5 (elementary); grades 6 through 8 (Middle); grades 9 through 12 (some high school but no diploma); high school graduate (high school diploma or equivalent, e.g., GED, foreign equivalent); vocational or trade school graduate; some college but no degree; associate's degree in college; bachelor's degree; master's degree; professional school or doctorate degree (MD, DDS, JD, DVM, Ph.D., EdD, etc.)?" (Hints)

"What is the highest grade of school or year of college you completed – no school/some grade school (1-6); eighth grade/junior high (7-8); some high school (9-10, no diploma, no GED); GED; graduated from high school; 1 to 2 years of college, no degree yet; 3 or more years of college, no degree yet; graduated from a 2-year college or vocational school, or associates degree; graduated from a 4- or 5-year college, or bachelor's degree; some graduate school; master's degree; PhD, ED.D., MD, DDS, LLB, LLD, JD, or other professional degree?" (Midus)

"What is the highest grade or year that you completed in school or college – less than 12 years, 12 years/high school graduate/GED, 1 or 2 years of college or junior college degree, 3 or more years of college but no degree, college graduate, or an advanced degree?" (Prisons)

"What is the last grade or class you completed in school?" [None, or grades 1-8; high school incomplete (grades 9-11); high school graduate (grade 12 or GED certificate); technical, trade or vocational school after high school; some college, no 4-year degree (includes associate degree); college graduate (B.S., B.A., or other 4-year degree), post-graduate training/professional school after college (toward a Masters/Ph.D., Law or Medical school)] (Social)

RACE/ETHNICITY

“Are you of Hispanic origin or Background?” [If yes] “Are you White Hispanic or Black Hispanic?” [If no] “Are you White, Black, or some other race?” (Abcnuke)

“Are you of Hispanic origin or Background?” [If yes] “Are you White Hispanic or Black Hispanic?” [If no] “Are you White, Black, or some other race?” (Abcpoll)

“Are you of Hispanic origin or descent, or not? Are you White, Black, Asian, or some other race?” (Cbcpoll)

“Are you of Hispanic of Latin origin or descent? What race do you consider yourself to be, White, Black, Asian, Native American, or some other race?” (Century)

“Please choose from this list the number which best describes your race.” [Showcard with responses: 1. Asian; 2. Black; 3. Hispanic; 4. White] (Cid)

“What is your racial or ethnic identity – White, not Hispanic, African American, White Hispanic/Latinos/Latinas, Asian or Pacific Islander, American Indian or Alaskan Native, other (specify)?” (Courts)

“Would you describe yourself as White, Black or African-American, Hispanic, Native American, Asian, or something else?” (Firearms)

“Are you Hispanic or Latino? Which one or more of the following would you say is your race? Are you American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, or White?” (Hints)

“What race do you consider yourself to be – White, Black and/or African American, Native American or Aleutian Island/Eskimo, Asian or Pacific Islander, other (please specify), multiracial (please specify)?” (Midus)

“Are you of Hispanic, Latino, or Spanish descent? Which of these racial/ethnic groups best describes you – White, Black or African-American, American Indian/Alaskan Native, Asian or Pacific Islander, other (specify)?” (Prisons)

“Are you, yourself, of Hispanic or Latino origin or descent, such as Mexican, Puerto Rican, Cuban, or some other Latin American background? What is your race? Are you White, Black, Asian, or some other race?” (Social)

GENDER

[Interviewer recorded respondent gender.] (Abcnuke)

[Interviewer recorded respondent gender.] (Abcpoll)

[Gender coded by interview, but asked if necessary] “For survey purposes, I need to ask are you male or female?” (Century)

[Respondent gender recorded from interviewer observation.] (CIDsurvey)

[Interviewer recorded respondent gender.] (Courts)

[Interviewer recorded respondent gender.] (Firearms)

“Are you male or female?” (Hints)

[Interviewer recorded respondent gender.] (Midus)

[Interviewer recorded respondent gender. If unsure, asked] “I am required to ask, are you male or female?” (Prisons)

[Interviewer record respondent gender.] (Social)

Appendix B

Question Wording from 12 Surveys

ACTIVITIES & BEHAVIORS

“Do you happen to have a face mask that covers the nose and mouth for each person in your household in your home or not?” (Abcnuke)

“Do you happen to have duct tape in your home or not?” (Abcnuke)

“How about a plan for how members of your household will communicate or where you'll meet up in case of a serious emergency - do you happen to have that kind of plan or not?” (Abcnuke)

“If you drive, what kind of vehicle do you usually drive - a car, an SUV, a pickup truck, a van or mini-van, or what?” (Abcnuke)

“Have you ever been shot with a gun?” (Firearms)

“Have you smoked a pack or more of cigarettes in the past week?” (Firearms)

“In the last month, how many times have you had 5 or more alcoholic drinks on one occasion?” (Firearms)

“In the past 12 months, have you personally had a fire in your home in which the fire department came?” (Firearms)

“Before being contacted for this study, have you ever heard of the 1-800-4-Cancer information number?” (Hints)

“During a typical weekday, about how many hours do you listen to the radio?” (Hints)

“During a typical weekend, including both Saturday and Sunday, about how many hours do you listen to the radio?” (Hints)

“During a typical weekend, including both Saturday and Sunday, about how many hours do you watch television?” (Hints)

“During the past 30 days, how often did you drink 100% fruit juice such as orange, apple, and grape juice? Do not include fruit drinks like Kool-Aid or Hi-C.” (Hints)

“During the past 30 days, how often did you eat fruit? Include fresh, canned, or frozen fruit.” (Hints)

“During the past 30 days, how often did you eat potatoes? Do not include things like fried potatoes, french fries, or rice.” (Hints)

“During the past 30 days, how often did you eat vegetables other than potatoes? Include things like salad, cooked dried beans, corn, and broccoli.” (Hints)

“Have you smoked at least 100 cigarettes in your entire life?” (Hints)

“In a typical week, how many days do you do any moderate-intensity physical activity or exercise comparable to walking as if you were in a hurry?” (Hints)

“By one 'drink', we mean either a bottle of beer, a wine cooler, a glass of wine, a shot of liquor, or a mixed drink. With these definitions in mind, during that year you drank most, about how many drinks would you usually have on the days that you drank?” (Midus)

“How old were you when you had your first drink, not counting a sip of someone else's drink?” (Midus)

“How old you were the last time htat you regularly drank this often?” (Midus)

“At what age did you have your very first cigarette?” (Midus)

“Think about the one year in your life when you drank the most, during that year, how often did you typically have at least one drink: every day, five or six days a week, three or four days a week, one or two days a week, or less than one day a week?” (Midus)

“Was there ever a time in your life when you regularly had at least one drink three or more days a week?” (Midus)

“Has someone ever broken into or somehow illegally gotten into your home or apartment?” (Prisons)

“Have you ever served on a jury in a criminal trial?”(Prisons)

“Are you currently active in any consumer groups such as AAA or coupon sharing programs?” (Social)

“Are you currently active in any literary, discussion or study groups, such as a book club or reading group?” (Social)

“Are you currently active in any social or fraternal clubs, sororities, or fraternities?” (Social)

“About how often do you use the internet or email from home - several times a day, about once a day, 3-5 days a week, 1-2 days a week, every few weeks, less often or never?” (Social)

“Are you currently active in any sports fantasy leagues?” (Social)

“Are you currently active in any travel clubs?” (Social)

“Have you, personally, ever been part of a group that was created online but lasted less than 12 months?” (Social)

“Thinking just about yourself now, do you use the internet, at least occasionally?” (Social)

DEMOGRAPHIC AND EMPLOYMENT

“Are you now married, widowed, divorced, separated, or have you never been married?” (Cbspoll)

“How concerned are you that in the next 12 months you or someone else in your household might be out of work and looking for a job -- very concerned, somewhat concerned, or not concerned at all?” (Cbspoll)

“If you were asked to use one of these five names for your social class, which would you say you belong in -- upper class, upper-middle class, middle class, working class, or lower class?” (Cbspoll)

“What is your current status of employment - full-time, part-time, unemployed, homemaker, student, retired, disabled?” (Century)

“Could I ask about your current marital status? Which of the descriptions on this card apply to you: married, separated (legally married), divorced, widowed, never married?” (Cid)

“Out of every 100 people living in the U.S., how many do you think were born outside the country?” (Cid)

“How many people live in your household?” (Courts)

“In the past 12 months, have you personally gotten married?” (Firearms)

“Which best describes the community that you live in - is it a city, suburb, or rural area?” (Firearms)

“Are you married, divorced, widowed, separated, never been married, or living with a partner?” (Hints)

“Are you currently employed for wages, self-employed, out of work for more than one year, out of work for less than one year, a homemaker, a student, retired, or unable to work?” (Hints)

“(When you are working, about/About) how many hours do you work for pay in an average week on your main job?” (Midus)

“The next questions are about your employment history. How old were you when you first worked for pay for six months or more, whether part-time or full-time?” (Midus)

“Think about your main job (from which you are currently [on leave/laid off]). Do you supervise anyone on this job?” (Midus)

“Do you live in a city, suburb, small town, or rural area?” (Prisons)

“Are you currently married, living with a partner, divorced, separated, widowed, or have you never been married?” (Social)

“Which of the following best describes the place where you now live - a large city, a suburb near a large city, as small city or town, or a rural area?” (Social)

FAMILY AND FRIENDS

“Do you have any children? If yes: Are any of your children under 18?” (Cbspoll)

“In the last month, about how often did you communicate with your close friends by email or Internet: Every day, a few times a week, a few times, once, not in the last month?” (Cid)

“Still thinking about your close friends, are any of them on welfare?” (Cid)

“Still thinking about your close friends, are any of them recent immigrants to the United States? By recent immigrants I mean those who have been in the country for less than five years.” (Cid)

“Think of all the close friends you had contact with in the past month, whether the contact was in person, by telephone, or by email. Of all these people, about how many of them have different religious views from yours - one, almost none, a few, some, about half, many, most, almost all, all?” (Cid)

“Think of all the close friends you had contact with in the past month, whether the contact was in person, by telephone, or by email. Of all these people, about how many of them have roughly the same level of education as you - none, almost none, a few, some, about half, many, most, almost all, all?” (Cid)

“How many children under the age of 18 live in this household?” (Firearms)

“Have any of your family members ever had cancer? If so, what type of cancer was it?” (Hints)

“How frequently do you talk to these friends or family members about your health - very frequently, somewhat frequently, or not very frequently?” (Hints)

“Has anyone you know well -- such as your spouse, a family member, or a close friend -- ever had a major heart procedure?” (Midus)

“How many (other) people in your immediate biological family -- that is, your biological parents, brothers or sisters -- have ever had a heart attack?” (Midus)

“How many (other) children do you have, including step children, adopted children, and any others you helped to raise for at least five years?” (Midus)

“In what month and year were you married (for the first time)?” (Midus)

“What was the highest grade of school or year of college your mother completed?” (Midus)

“How many of your friends are not religious at all?” (Religion)

“How many of your friends go to a different church?” (Religion)

FINANCES & ECONOMICS

“Which of the following categories best describe your total annual household income before taxes, from all sources - under 20 thousand dollars, 20 to under 35 thousand, 35 to under 50 thousand, 50 to under 75 thousand, 75 to under 100 thousand, 100 thousand or more?” (Abcnuke)

“Changing topics, many economists say that using the standards they apply, the recession is probably over. Thinking about your own experience of economic conditions, would you say that from your point of view the recession is over, or not over?” (Abcpoll)

“Do you think the United States is in a long-term economic decline, or do you think the country's economic system is basically pretty solid?” (Abcpoll)

“How concerned are you about being able to maintain your current standard of living -- are you very worried, somewhat worried, not so worried, or not worried at all?” (Abcpoll)

“Is the current economic situation a cause of stress in your life or not? If yes: Is it a cause of serious stress, or stress but not serious?” (Abcpoll)

“Thinking now about job opportunities where you live, would you say there are plenty of jobs available in your community, or are jobs difficult to find?” (Abcpoll)

“Which of the following categories best describe your total annual household income before taxes, from all sources?” (Abcpoll)

“Based on what you've seen, do you think the economy in your local community is getting better, getting worse, or staying about the same?” (Cbcpoll)

“Compared to a year ago, do you owe more money now on your credit cards, less money on your credit cards, or about the same amount of money on your credit cards as you did a year ago?” (Cbcpoll)

“Compared to a year ago, have medical expenses become a larger part of your total credit card debt, a smaller part, or have they stayed about the same?” (Cbcpoll)

“Do you have any major credit cards?” (Cbcpoll)

“Do you think the economy is getting better, getting worse, or staying about the same?” (Cbcpoll)

“Do you think your household's financial situation is getting better, getting worse, or staying about the same?” (Cbcpoll)

“How much, if any, of your credit card debt is due to charging medical expenses such as doctor visits, prescription drugs, or hospital stays -- almost all of it, most of it, just some of it, or almost none of it?” (Cbcpoll)

“How would you describe the affordability of basic medical care for you and your family - would you say basic medical care is easy to afford, not easy to afford but manageable, hard to afford, or is basic medical care beyond your financial means right now?” (Cbcpoll)

“How would you rate the condition of the national economy these days? Is it very good, fairly good, fairly bad, or very bad?” (Cbcpoll)

“In this recession, which of the following has been the most difficult thing for you to cut back spending on -- vacations, or new clothes, or dining out, or alcohol?” (Cbcpoll)

“Was your total family income in the year 2008 under or over 50,000? Was it under 15,000, between 15,000 and 30,000, or between 30,000 and 50,000? Was it between 50,000 and 75,000, or between 75,000 and 100,000, or was it over 100,000?” (Cbcpoll)

“Please tell me if the total amount of income, before taxes received, by all members in your household during 2003 was above or below \$45,000?” (Century)

“Would you say that in the past year the national economy has gotten better, stayed the same, or gotten worse?” (Century)

“Which of the following best describes the total annual household income for all persons living in your household, under \$15,000, \$15,000-\$35,000, \$35,000-\$50,000, \$50,000-\$75,000, or over \$75,000?” (Firearms)

“At the end of the month, how much money are you able to put aside? Your best estimate is fine. Nothing, \$100 or less, \$101 to \$250, \$251 to \$500, \$501 to \$1,000, more than \$1,000?” (Hints)

“Thinking about members of your family living in this household, what is your combined annual income, meaning the total pre-tax income from all sources earned in the past year?” (Hints)

“During your childhood adolescence, was there ever a period of six months or more when your family was on welfare or ADC?” (Midus)

“Thinking of the income that your household earned or received from all sources in 1999, was the total amount before taxes and other deductions \$35,000 or more?” (Prisons)

HEALTH AND HEALTHCARE

“Do you have some form of health insurance or health care coverage, or not?” (Abcpoll)

“Are you covered by health insurance through an employer, a union, or through a plan you got on your own, or are you covered by Medicare or Medicaid?” (Cbcpoll)

“California and New York City now require calorie-counts to be printed on fast food menus. Which of the following other things would you most like to see happen to reduce obesity -- a fast food tax, or a soft drink tax, or a ban on using food stamps to buy high-fat food, or putting scales at restaurant entrances, or a tax credit for liposuction?” (Cbcpoll)

“Do you have some form of health insurance or health care coverage, or not?” (Cbcpoll)

“How serious a problem would you say the H1N1 or swine flu is? Would you say it is a very serious problem, a somewhat serious problem, not too serious of a problem, or not a serious problem at all?”(Cbcpoll)

“If you were to become seriously injured or seriously ill, how confident are you that your own insurance would pay for everything you needed to get better - are you very confident, somewhat confident, not too confident, or not at all confident that your insurance company would pay for everything you needed?” (Cbcpoll)

“Is your opinion of the health insurance industry very favorable, somewhat favorable, somewhat unfavorable, or very unfavorable?” (Cbcpoll)

“Thinking about the medical treatments and medicines that you currently need, would you say your own health insurance company currently pays for all of those costs, most of those costs, some of those costs, or almost none of those costs?” (Cbcpoll)

“Wal-Mart has announced that it is considering the addition of swine flu or H1N1 vaccination station to their 4,000 mega-stores across the country. If Wal-Mart offered healthcare services -- such as flu shots, dental, eye, or basic check-ups -- would you consider becoming a Wal-Mart patient?” (Cbspoll)

“In the past 12 months, have you personally been hospitalized for a fracture?” (Firearms)

“As far as you know, does physical activity or exercise increase the chances of getting some types of cancer, decrease the chances of getting some types of cancer, or does it not make much difference?” (Hints)

“Do you have any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare?” (Hints)

“Do you think that a low carbohydrate, high protein diet is a healthy way to lose weight?” (Hints)

“How many days a week of physical activity or exercise are recommended for the average adult to stay healthy? On those days, how long should the average adult be physically active to stay healthy?” (Hints)

“How many servings of fruits and vegetables do you think a person should eat each day for good health?” (Hints)

“In general, would you say your health is excellent, very good, good, fair, or poor?” (Hints)

“Assume for a moment that you had a heart condition and were told by your doctor that you had two choices -- either to have coronary bypass surgery or to take medication every day for the rest of your life. Which one do you think you would choose?” (Midus)

“At that time, was your blood pressure low, about normal, slightly raised, or high?” (Midus)

“Do you think your risk of a heart attack is higher, lower, or about the same as other (men/women) your age?” (Midus)

“Do you think your risk of getting cancer is higher, lower, or about the same as other (men/women) your age?” (Midus)

“Have you ever had a major heart procedure, such as catheterization, bypass surgery, or angioplasty?” (Midus)

“Have you ever had cancer?” (Midus)

“Have you ever had severe pain across the front of your chest lasting half an hour or more?” (Midus)

“How about your mental health at that time? Was it poor, fair, good, very good, or excellent?” (Midus)

“How long has it been since your last blood pressure test?” (Midus)

“How much do you worry about your heart, a lot, some, a little, or not at all?” (Midus)

“How sure are you that this [accept or get second opinion] is what you would do, very, somewhat, or not very sure?” (Midus)

“How sure are you that this [bypass surgery or medication] is what you would choose, very, somewhat, or not very sure?” (Midus)

“If your doctor recommended a major heart operation, would you accept this opinion or get a second opinion before deciding?” (Midus)

“In general, compared to most (men/women) your age, would you say your health is much better, somewhat better, about the same, somewhat worse, or much worse?” (Midus)

“In the past 30 days, how many days were you totally unable to go to work or carry out your normal household work activities because of your physical health or mental health?” (Midus)

“Now, think about when you were 16 years old. Was your physical health at that time poor, fair, good, very good, or excellent?” (Midus)

“The first questions are about your health. In general, would you say your physical health is poor, fair, good, very good, or excellent?” (Midus)

“The next set of questions is about your physical health. First, have you ever had heart trouble suspected or confirmed by a doctor?” (Midus)

“What about your mental or emotional health? Would you say it is poor, fair, good, very good, or excellent?” (Midus)

“What was the exact reading [blood pressure]? (Diastolic)” (Midus)

“What was the exact reading [blood pressure]? (Systolic)” (Midus)

PERSONALITY AND FEELINGS

“On a final subject, are you a fan of professional golf or not?” (Abcpoll)

“Overall do you have a favorable or unfavorable opinion of the professional golfer Tiger Woods? Do you feel that way strongly or somewhat?” (Abcpoll)

“Woods has been accused of having extramarital affairs, and has apologized for what he called “personal sins.” Given this situation, do you think companies should or should not continue to use Woods to endorse their products? Do you feel that way strongly or somewhat?” (Abcpoll)

“The American banking and car industries have suffered one of their worst years ever. Given that, which of the following best symbolizes America today -- Microsoft, Wal-Mart, Google, the National Football league, or Goldman Sachs?” (Cbcpoll)

“Which of the following celebrities' likenesses do you think is most likely to be used to endorse products 100 years from now - Michael Jackson, Miley Cyrus, James Dean, Albert Einstein, or Madonna?” (Cbcpoll)

“Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?” (Century)

“Giving money to charities - is this an obligation you owe to other Americans?” (Century)

“I would feel good if I were described as a typical American. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“In many respects, I am different from most Americans. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“When I think of the American people, I think of people who are a lot like me. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“For each of the following statements, please indicate whether you agree strongly, agree, are uncertain, disagree, or strongly disagree: Of all the different philosophies that exist in the world, there is probably only one that is correct.” (Cid)

“Here is a list of words that can be used to describe various political groups. Taking them one at a time, please tell me how you feel about [GROUP X]. [To what degree are you feeling angry? Showcard ranging from 1-7 with 1 indicating angry and 7 indicating indifference.]” (Cid)

“Please tell me on a scale of 0-10 how much you personally trust each of these institutions I read. 0 means you do not trust an institution at all, and 10 means you have complete trust. Unions?” (Cid)

“I am going to read you a list of statements about opportunity and discrimination. Please tell me how strongly you agree or disagree with each. Would you say you strongly agree, somewhat agree, somewhat disagree, or strongly disagree? America has provided me with an equal opportunity to get ahead in life.” (Courts)

“During the past 30 days, how often did you feel hopeless? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“During the past 30 days, how often did you feel nervous? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“During the past 30 days, how often did you feel restless or fidgety? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“During the past 30 days, how often did you feel so sad that nothing could cheer you up? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“During the past 30 days, how often did you feel that everything was an effort? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“During the past 30 days, how often did you feel worthless? Would you say all of the time, most of the time, some of the time, a little of the time, or none of the time?” (Hints)

“[For the final set of questions, please tell me how much each of the following describes you.] Curious? Does this describe you a lot, somewhat, a little, or not at all.” (Midus)

“And would you agree or disagree with the following statement: “In many ways, I feel disappointed about my achievements in life?”” (Midus)

“At present, how much control do you have over your life in general? Would you say a lot, somewhat, a little, or none at all?” (Midus)

“At present, how satisfied are you with your life? Would you say a lot, somewhat, a little, or not at all?” (Midus)

“Do you agree [with the statement “In many ways, I feel disappointed about my achievements in life?”] strongly, somewhat, or only a little?” (Midus)

“During the past 12 months, did you ever have a spell or an attack when all of a sudden you felt frightened, anxious, or very uneasy, in a situation when most people would not be afraid or anxious?” (Midus)

“During the past 12 months, did you ever have a spell or attack when for no reason your heart suddenly began to race, you felt faint, or you couldn't catch your breath? When we say 'for no reason,' we mean that it was not due to any physical cause, like a heart problem.” (Midus)

“During the past 12 months, was there ever a time lasting two weeks or more when you lost interest in most things like hobbies, work, or activities that usually give you pleasure?” (Midus)

“For the final set of questions, please tell me how much each of the following describes you. Outgoing? Does this describe you a lot, somewhat, a little, or not at all.” (Midus)

“Overall, how satisfied are you with your self? Would you say a lot, somewhat, a little, or not at all?” (Midus)

“People differ a lot in how much they worry. Considering how things have been going in your life over the past 12 months, do you worry more than most people in the same situation, less than most people, or about the same as most people in the same situation?” (Midus)

“The next questions are about your mood. During the past 12 months, was there ever a time when you felt sad, blue, or depressed for two weeks or more in a row?” (Midus)

“Thinking about the past 12 months, did you worry every day, just about every day, most days, about half the days, or less than half the days?” (Midus)

“When you think about your life as a whole up to the present, how would you rate your contribution to the welfare and well-being of other people: excellent, very good, good, fair, or poor?” (Midus)

“[For the final set of questions, please tell me how much each of the following describes you.] Optimistic? Does this describe you a lot, somewhat, a little, or not at all.” (Midus)

“[For the final set of questions, please tell me how much each of the following describes you.] Worrying? Does this describe you a lot, somewhat, a little, or not at all.” (Midus)

“Overall, how much impact do you think people like you can have in making your community a better place to live -- a big impact, a moderate impact, a small impact, or no impact at all?” (Social)

“Overall, how would you rate your community as a place to live? Would you say it is excellent, good, only fair, poor.” (Social)

“Thinking more generally, not just about your local community, generally speaking would you say that most people can be trusted or that you can't be too careful in dealing with people?” (Social)

POLITICS & GOVERNMENT

“Generally speaking, do you usually think of yourself as: a democrat, a republican, an independent, or what?” (Abcnuke)

“Would you say your views on most political matters are liberal, moderate, or conservative?” (Abcnuke)

“Do you approve or disapprove of the way Barack Obama is handling his job as president? If approve: Do you approve strongly or somewhat? If disapprove: Do you disapprove strongly or somewhat?” (Abcpoll)

“Generally speaking, do you usually think of yourself as: a democrat, a republican, an independent, or what?” (Abcpoll)

“Would you say your views on most political matters are liberal, moderate, or conservative?” (Abcpoll)

“Did you vote in the 2008 presidential election, did something prevent you from voting, or did you choose not to vote? If voted: Did you vote for Barack Obama, John McCain, or someone else?” (Cbcpoll)

“Do you approve or disapprove of the way Congress is handling its job?” (Cbcpoll)

“Do you think Barack Obama has strong qualities of leadership or not?” (Cbcpoll)

“Generally speaking, do you usually think of yourself as: a democrat, a republican, an independent, or what?” (Cbcpoll)

“How much do you think Barack Obama cares about the needs and problems of people like yourself - a lot, some, not much, or not at all?” (Cbcpoll)

“How much of the time do you think you can trust the government in Washington do what is right - just about always, most of the time, or only some of the time?” (Cbcpoll)

“How would you describe your views on most political matters? Generally do you think of yourself as liberal, moderate, or conservative?” (Cbcpoll)

“If the 2010 election for the U.S. House of Representatives were being held today, would you vote for the Republican candidate or the Democratic candidate in your district?” (Cbcpoll)

“These last few questions are for background only. Some people are registered to vote and others are not. Are you registered to vote in the precinct or election district where you now live, or aren't you?” (Cbcpoll)

“Generally speaking, do you think of yourself as conservative, moderate, or liberal?” (Century)

“How many years is the term of a United States Senator?” (Century)

“What are the first 10 amendments to the Constitution called?” [Open-ended response coded as correct or incorrect.] (Century)

“What position is held by William Rehnquist?” [Open-ended response coded as correct or incorrect.] (Century)

“Which political party has the most members in the House of Representatives -- the Democrats or the Republicans?” (Century)

“Do you think corruption by public officials in the United States is more widespread than it was a few decades ago, less widespread, or about the same?” (Cid)

“For each of the following statements, could you please indicate whether you agree strongly, agree, are uncertain, disagree, or disagree strongly: The Supreme Court can usually be trusted to make decisions that are right for the country as a whole?” (Cid)

“Generally speaking, do you usually think of yourself as: a democrat, a republican, an independent, or what?” (Cid)

“If corruption means the abuse of public office for private gain, in your opinion, how widespread is corruption in the United States?” [Showcard with response options: hardly anyone is involved; a small number of people are involved; a moderate number of people are involved; a lot of people are involved; almost everyone is involved.] (Cid)

“On the whole, how satisfied are you with the way democracy works in the United States?” [Showcard from 0-10 where 0 indicates extremely dissatisfied and 10 indicates extremely satisfied.] (Cid)

“Please tell me on a scale of 0-10 how much you personally trust each of these institutions I read. 0 means you do not trust an institution at all, and 10 means you have complete trust. Political parties?” (Cid)

“Please tell me on a scale of 0-10 how much you personally trust each of these institutions I read. 0 means you do not trust an institution at all, and 10 means you have complete trust. The US Supreme Court?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: Citizens do not have enough influence in policy decisions?” (Cid)

“We hear a lot of talk these days about liberals and conservatives. Here is a 0-10 scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale?” (Cid)

“Judges are honest in their case decisions. Would you say you strongly agree, somewhat agree, somewhat disagree, or strongly disagree?” (Courts)

“As of today, do you lean more to the Republican Party or more to the Democratic Party?”
(Social)

“In general, would you describe your political views as very conservative, conservative, moderate, liberal, or very liberal?” (Social)

“In politics today, do you consider yourself a Republican, Democrat, or Independent?”
(Social)

RACE/ETHNIC RELATIONS

“Immigrants should really know what's going on in the United States if they want to stay here, but a lot of them just don't want to be bothered. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“Immigrants today come to think of themselves as Americans just as much as immigrants from earlier eras did. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“Immigrants today take advantage of jobs and opportunities without doing enough to give back to the community? Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“In general, do you think discrimination against (country of origin fill) is a major problem, a minor problem, or not a problem in schools?” (Century)

“In general, do you think discrimination against (race fill) is a major problem, a minor problem, or not a problem in schools?” (Century)

“It has been reported that some police officers stop motorists of certain racial or ethnic groups because the officers believe that these groups are more likely than others to commit certain types of crime. This practice is known as racial profiling. Do you approve or disapprove of the use of racial profiling by police?” (Century)

“Over the last few years, immigrants have gotten less than they deserve. Do you strongly agree, somewhat agree, somewhat disagree, strongly disagree?” (Century)

“People have different ideas about how to teach non-English speaking children when they enter public schools. I am going to read a few of them. Please tell me which one comes closest to your view. First, all classes should be conducted only in English. Second, children should be able to take some classes in native language, just for a year or two. Or, third, children should be able to take classes in their native language all throughout high school.”
(Century)

“People have different ideas about what the main goal of bilingual education should be. I am going to read a few of them. Please tell me which one comes closest to your view. First, teach English to non-native children as quickly as possible. Second, teach English to non-native children while making sure they do not fall behind in other subjects. Or, third, teach English to non-native children while making sure they remain fluent in their first language.” (Century)

“What about in preventing (race fill) in general from succeeding in America? Do you think that is a major problem, a minor problem, or not a problem?” (Century)

“What about in the workplace? (Do you think discrimination against (country of origin fill) is a major problem, a minor problem, or not a problem in the workplace?)” (Century)

“What about people who immigrated illegally? Do you think they should be allowed to benefit from government assistance programs like Medicaid and food stamps?” (Century)

“What about the workplace? (Do you think discrimination against (race fill) is a major problem, a minor problem, or not a problem?)” (Century)

“You may have noticed that in some neighborhoods, the sign on the front of a store or a business will be in a language other than English. Over the past several years, some cities and towns have adopted policies that require a certain percentage of the words on such signs be in English. Do you think you would favor or oppose such a policy?” (Century)

“And to what degree do you feel hatred versus indifference toward [GROUP X]? [Showcard ranging from 1-7 with 1 indicating hatred and 7 indicating indifference.]” (Cid)

“Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?” (Cid)

“For each of the following statements, please indicate whether you agree strongly, agree, are uncertain, disagree, or strongly disagree: In the long run, the best way to live is to pick friends and associates whose tastes and beliefs are the same as one's own.” (Cid)

“Here is a list of the groups we have been talking about. Of the groups on this list, which do you dislike the most?” [Conservatives; the U.S. Communist Party; Christian fundamentalists; the Ku Klux Klan; people who are against all churches and religions; American Nazis; those who prohibit abortion; The Society for a New America; Liberals; those who would allow all abortions; those who advocate doing away with elections and letting the military run the country; gay rights activists; radical Muslims.] (Cid)

“Next, we would like to ask your opinion about different racial groups in the United States. Would you say you agree strongly, agree, are uncertain, disagree, or disagree strongly with the following statements: More than most groups [RACE 2] find it difficult to understand the customs and ways of [RACE].” (Cid)

“Of the people you interact with in your neighborhood, how many of them are recent immigrants to the United States?” [Showcard with the categories none (0%); almost none (5%); a few (10%); some (25%); about half (50%); many (75%); almost all (95%); all (100%)] (Cid)

“Please tell me how important you think each of these things should be in deciding whether someone born and raised outside the United States should be able to come and live here: Be oppressed in their country.” [Show card with responses ranging from 1-10 where 1 indicates extremely unimportant and 10 indicates extremely important.](Cid)

“Suppose two people - one white and one black - each appeared in court charged with a crime they did not commit. What do you think their chances are of being found guilty: the white person is more likely to be found guilty, both have the same chance, the black person is more likely to be found guilty?” (Cid)

“There has been much talk recently about whether gays and lesbians should have the legal right to marry someone of the same sex. Which of the following comes closest to your position on this issue? Do you support full marriage rights for gay and lesbian couples; do you support gay civil unions or partnerships, but not gay marriage; or, do you oppose any legal recognition for gay and lesbian couples?” (Cid)

“Using a scale of 0 to 10, where 0 means 'Cannot be trusted at all' and 10 means 'Can be trusted a lot', how much do you trust each of the following groups of people: People in your neighborhood?” (Cid)

“Using a scale of 0 to 10, where 0 means 'Cannot be trusted at all' and 10 means 'Can be trusted a lot', how much do you trust each of the following groups of people: People of a different race than yours?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: Gay men and lesbians should be free to live their own lives as they wish?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: It is important for government officials to represent my race?” (Cid)

“Using this card, please tell me how much you agree or disagree with each of these statements: Immigrant groups today tend to treat women less fairly and less equally than most Americans do.” [Showcard with responses ranging from 1-5 where 1 indicates agree strongly and 5 indicates disagree strongly.] (Cid)

“Using this card, please tell me how much you agree or disagree with each of these statements: Immigrants today are just as tolerant and open minded as most Americans.”

[Showcard with responses ranging from 1-5 where 1 indicates agree strongly and 5 indicates disagree strongly.] (Cid)

“Would you say you agree strongly, agree, are uncertain, disagree, or disagree strongly with the following statements: More than most groups [RACE 1] are selfish, and only look after the interests of their own group.” (Cid)

“Would you say you agree strongly, agree, are uncertain, disagree, or disagree strongly with the following statements: More than most groups [RACE 1] groups are likely to engage in crime?” (Cid)

“How often are each of the following groups of people treated worse than others by the courts? Do they always, often, sometimes, rarely, or never treat a Latino or Hispanic worse than they treat others?”(Courts)

“How often are each of the following groups of people treated worse than others by the courts? Do they always, often, sometimes, rarely, or never treat a Non-English speaker worse than they treat others?” (Courts)

“How often are each of the following groups of people treated worse than others by the courts? Do they always, often, sometimes, rarely, or never treat an African-American worse than they treat others?” (Courts)

“How often are each of the following groups of people treated worse than others by the courts? Do they always, often, sometimes, rarely, or never treat someone with a low income worse than they treat others?” (Courts)

“If you are out alone in your neighborhood at night do you feel very safe, somewhat safe, somewhat unsafe, very unsafe.” (Firearms)

“Next, I'd like to ask you about the amount of influence or power different groups have. Let's start with minority groups in the United States and their Civil Rights. Would you say that they have too few Civil Rights, about the right amount, or too many Civil Rights?” (Prisons)

RELIGION AND SPIRITUALITY

“What, if anything, is your religion?” (Abcpoll)

“Would you consider yourself a born-again or evangelical Christian, or not?” (Abcpoll)

“What is your religious preference today?” (Cbcpoll)

“What, if any, is your religious preference?” (Century)

“Would you call yourself born-again?” (Cid)

“Not including funerals and weddings, how often do you attend religious services? Would you say every week, once or twice a month, a few times a year, or never?” (Hints)

“By your best guess, how would your place of worship feel about the following behaviors - raising hands during the service - encourages it, allows it, discourages it, forbids it?” (Religion)

“Which one statement comes closest to your personal beliefs about the Bible - it means exactly what it says and should be taken literally, it is perfectly true and should not be taken literally, the bible contains some human error, the bible is an ancient book of history and legends?” (Religion)

“Which one statement comes closest to your personal view of religious salvation - my religion is the one true faith that leads to salvation, many religions lead to salvation, I do not believe in religious salvation?” (Religion)

“With what religious family do you most closely identify?” (Religion)

“Aside from weddings and funerals, how often do you attend religious services... more than once a week, once a week, once or twice a month, a few times a year, seldom, or never?” (Social)

SPECIFIC POLICIES

“Do you approve or disapprove of the way Obama is handling health care? If approve: Do you approve strongly or somewhat? If disapprove: Do you disapprove strongly or somewhat?” (Abcpoll)

“Do you approve or disapprove of the way Obama is handling the situation in Afghanistan? If approve: Do you approve strongly or somewhat? If disapprove: Do you disapprove strongly or somewhat?” (Abcpoll)

“Do you think the government should or should not expand Medicare to cover people between the ages of 55 and 64 who do not have health insurance? Do you feel that way strongly or somewhat?” (Abcpoll)

“In the long run, do you think the country's health care system overall will cost more if the changes that have been proposed become law or if the current system is left as it is now?” (Abcpoll)

“In the long run, do you think the quality of health care you receive will be better if the changes that have been proposed become law or if the current system is left as it is now?” (Abcpoll)

“In the long run, do you think your own health care will cost you more if the changes that have been proposed become law or if the current system is left as it is now?” (Abcpoll)

“It's been proposed that the United States and other developed countries contribute 10 billion dollars a year to help developing countries pay for reducing the amount of greenhouse gases they release. Is this something you support or oppose?” (Abcpoll)

“Just your best guess, do you think health care reform would increase the federal budget deficit, decrease it, or have no effect? If increase: Do you think that would be worth it, or not?” (Abcpoll)

“Medicare is the government health insurance program for people 65 and over. Do you think health care reform would strengthen the Medicare program, weaken Medicare, or have no effect on it?” (Abcpoll)

“On another subject, do you think the federal government should or should not regulate the release of greenhouse gases from sources like power plants, cars, and factories in an effort to reduce global warming? Do you feel that way strongly or somewhat?” (Abcpoll)

“Overall, given what you know about them, would you say you support or oppose the proposed changes to the health care system being developed by Congress and the Obama administration?” (Abcpoll)

“Which comes closer to the way you feel: government reform of the nation's health care system is necessary to control costs and expand coverage, or government action on health care will do more harm than good? Do you feel that way strongly or somewhat?” (Abcpoll)

“Do you approve or disapprove of the way Barack Obama is handling his job as president?” (Cbbspoll)

“Do you think it's possible or not possible to expand health care coverage to more Americans without increasing the budget deficit?” (Cbbspoll)

“Do you think it's possible or not possible to expand health care coverage to more Americans without raising taxes on the middle class?” (Cbbspoll)

“Do you think the death of Senator Ted Kennedy makes the passage of health care reform this year more likely, less likely, or doesn't it have any impact on whether or not health care reform will pass this year?” (Cbbspoll)

“Do you think the federal government should put a limit on the amount of money that senior executives can earn at financial institutions, or do you think this is something the federal government should not be involved in?” (Cbbspoll)

“Do you think the government would do a better or worse job than private insurance companies in holding down health care costs?” (Cbbspoll)

“Do you think the government would do a better or worse job than private insurance companies in providing medical coverage?” (Cbspoll)

“Do you think you understand the health care reforms under consideration in Congress, or are they confusing to you?” (Cbspoll)

“From what you've heard or read about, what do you like least about the current plans to reform the U.S. health care system?” (Cbspoll)

“From what you've heard or read about, what do you like most about the current plans to reform the U.S. health care system?” (Cbspoll)

“From what you've heard or read, do you think the health care reforms under consideration in Congress will mostly help middle class Americans, will mostly hurt middle class Americans, or don't you think they will have much of an effect on middle class Americans?” (Cbspoll)

“From what you've heard or read, do you think the health care reforms under consideration in Congress will mostly help senior citizens, will mostly hurt senior citizens, or don't you think they will have much of an effect on senior citizens?” (Cbspoll)

“From what you've heard or read, do you think the health care reforms under consideration in Congress will mostly help small businesses, will mostly hurt small businesses, or don't you think they will have much of an effect on small businesses?” (Cbspoll)

“Regardless of how you usually vote, who do you think has better ideas about reforming the health care system - Barack Obama, or the Republicans in Congress?” (Cbspoll)

“Regardless of your own stance on health care, do you think the protesters at the town hall meetings generally reflect the views of most Americans, or not?” (Cbspoll)

“Some lawmakers have proposed requiring health insurance companies to cover anyone who applies for health insurance regardless of whether or not they have an existing medical condition or a prior illness. Do you approve or disapprove of requiring health insurance companies to cover anyone who applies?” (Cbspoll)

“What do you think will eventually happen to the nation's health care system if there are no government reforms right now - will the health care system get better on its own, get worse on its own, or will the health care system stay the same as it is now?” (Cbspoll)

“Which of the following three statements comes closest to expressing your overall view of Medicare: 1. On the whole, Medicare works pretty well and only minor changes are necessary to make it work better, 2. There are some good things about Medicare, but fundamental changes are needed, 3. Medicare has so much wrong with it that we need to completely rebuild it?” (Cbspoll)

“Which of the following three statements comes closest to expressing your overall view of the health care system in the United States: 1. On the whole, the health care system works pretty well and only minor changes are necessary to make it work better, 2. There are some good things in our health care system but fundamental changes are needed, 3. Our healthcare system has so much wrong with it that we need to completely rebuild it?” (Cbspoll)

“Would you approve or disapprove of the government providing subsidies to help low-income people buy their own health insurance from private health insurance companies?” (Cbspoll)

“Would you approve or disapprove of the government setting limits on the amount that health insurance companies can charge people for insurance premiums, co-pays, and out-of-pocket expenses?” (Cbspoll)

“Some judges in the U.S. serve for a set number of years; others serve a life term. Do you happen to know whether the justices of the U.S. Supreme Court serve for a set number of years or whether they serve a life term?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: The government should take measures to reduce differences in income levels?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: The U.S. government should increase foreign aid to poor countries?” (Cid)

“Trial courts handle different types of cases. As you know, there are civil cases and criminal cases but courts also handle other kinds of cases involving, for example, family problems and juvenile delinquency. On a scale from 1 to 5 with 1 being the very lowest rating and 5 being the very highest, please tell me how well you think the courts in your community handle each of the following types of cases: Criminal cases involving drug abusers or drunk drivers.” (Courts)

“Trial courts handle different types of cases. As you know, there are civil cases and criminal cases but courts also handle other kinds of cases involving, for example, family problems and juvenile delinquency. On a scale from 1 to 5 with 1 being the very lowest rating and 5 being the very highest, please tell me how well you think the courts in your community handle each of the following types of cases: Criminal cases involving violence such as robbery.” (Courts)

“Using the same scale, with 1 being least favorable and 5 being most favorable, how would you rate how you feel about your local police?” (Courts)

“Using the same scale, with 1 being least favorable and 5 being most favorable, how would you rate how you feel about your local schools?” (Courts)

“Do you think gun injuries in your community are a major problem, minor problem, not much of a problem?” (Firearms)

“Please consider the following hypothetical situation. Your community is thinking about adding a program to reduce the number of gunshot injuries by 30%. The program would have no other effect. If this program added \$100 to your taxes, would you favor or oppose this program?” (Firearms)

“Please consider the following hypothetical situation. Your community is thinking about adding a program to reduce the number of gunshot injuries by 30%. The program would have no other effect. If this program added \$50 to your taxes, would you favor or oppose this program?” (Firearms)

“Please consider the following hypothetical situation. Your community is thinking about adding a program to reduce the number of gunshot injuries by 30%. The program would have no other effect. If you had to pay extra taxes for this program to reduce the number of gunshot injuries, would you favor or oppose this program?” (Firearms)

“For the next few questions, I want you to put yourself in the shoes of your local mayor. The Federal government has given your city [\$100, \$1,000] per household. You may spend all or part of that money on crime control or crime prevention, or you may give all or part of it back to your local residents. Four different crime control strategies have been recommended to you: one - add more prisons, two - add more drug and alcohol treatment programs for offenders convicted of nonviolent crimes, three - add more police on the street, and four - add more prevention programs to help keep youth out of trouble. Once again, you have [\$100/\$1,000] per household to allocate these programs or to rebate to local residents. Would you spend any of this money to add more prisons?” (Prisons)

“We are faced with many problems in this country, none of which can be solved easily or inexpensively. For each of the following, please tell me whether you think we're sending too much money on it, too little money, or about the right amount. Police on the street?” (Prisons)

“We are faced with many problems in this country, none of which can be solved easily or inexpensively. For each of the following, please tell me whether you think we're sending too much money on it, too little money, or about the right amount. Programs designed to get people off drugs?” (Prisons)

“Would you spend any of this money to add drug and alcohol treatment programs for offenders convicted of nonviolent crimes?” (Prisons)

“Would you spend any of this money to add more police on the street?” (Prisons)

“To what extent do you agree or disagree that the federal government should abolish the death penalty - strongly agree, agree, disagree, strongly disagree?” (Religion)

“To what extent do you agree or disagree that the federal government should distribute the wealth more evenly - strongly agree, agree, disagree, strongly disagree?” (Religion)

WAR AND TERRORISM

[If needed: “In the event of a terrorist attack using a nuclear bomb or so called dirty bomb in the area where you live”] “Do you feel you'd know where to turn for information on what to do, or not?” (Abcnuke)

[If needed: How confident are you that (Item) are being adequately protected against terrorists?] “Lower-level radioactive materials in medical or industrial waste as it's transported to storage sites - very confident, somewhat confident, not so confident, not confident at all?” (Abcnuke)

“Compared to before September 11, 2001, do you think the country today is safer from terrorism or less safe from terrorism? [If safer] Would you say the country is much safer or somewhat safer?” (Abcnuke)

“Do you think the United States is doing all it reasonably can do to try to prevent further terrorist attacks, or do you think it should do more?” (Abcnuke)

“Do you think the United States is or is not doing enough to try to prevent Airline hijackings?” (Abcnuke)

“Do you think the United States is or is not doing enough to try to prevent car bombs or suicide bombers?” (Abcnuke)

“How confident are you in the ability of your local hospitals to respond effectively to a terrorist attack using nuclear or radioactive materials - very confident, somewhat confident, not so confident, or not confident at all?” (Abcnuke)

“If a terrorist attack using a so called “dirty bomb” combining conventional explosives with radioactive material happened in their area would most people you know be alarmed but not panicked, concerned but not alarmed, or not concerned?” (Abcnuke)

“Now I'll name a few specific types of nuclear or radiological materials. For each one, please tell me how confident you are that it's being adequately protected against terrorists - very confident, somewhat confident, not so confident, or not confident at all? Nuclear power plants in the United States.” (Abcnuke)

“Thinking specifically of a terrorist attack using a nuclear bomb or so-called dirty bomb, would you know what to do if that kind of attack happened in the area you live, or not?” (Abcnuke)

“Do you support or oppose Obama's decision to send approximately 30,000 additional U.S. Military forces to Afghanistan? Do you feel that way strongly or somewhat?” (Abcpoll)

“Do you think the United States must win the war in Afghanistan in order for the broader war on terrorism to be a success, or do you think the war on terrorism can be a success without the United States winning the war in Afghanistan?” (Abcpoll)

“How confident are you that Obama's new strategy for Afghanistan will succeed - very confident, somewhat confident, not so confident, or not confident at all?” (Abcpoll)

“Obama has said U.S. forces will start withdrawing soon from Afghanistan in summer 2011. Do you think they should start withdrawing sooner than that, later than that, or is this schedule about right?” (Abcpoll)

“On another subject, all in all, considering the costs to the United States versus the benefits to the United States, do you think the war in Afghanistan has been worth the fighting, or not? Do you feel that way strongly or somewhat?” (Abcpoll)

“Regardless of when you think the U.S. withdrawal should begin, do you support or oppose Obama having set a specific deadline for this to occur?” (Abcpoll)

“Barack Obama has set a timetable for a troop pullout in Iraq by 2011. When do you think Barack Obama will announce that he has set a timetable for removing U.S. troops from Afghanistan? 1. In a year or so, or 2. When the next presidential campaign begins, or 3. Barack Obama won't set a timetable for removing U.S. troops from Afghanistan?” (Cbcpoll)

“Do you approve of the way Barack Obama is handling the situation in Afghanistan?” (Cbcpoll)

“Do you think the policies of the Obama Administration have made the United States safer from terrorism, less safe from terrorism, or have the policies of the Obama Administration not affected the U.S.' safety from terrorism?” (Cbcpoll)

“From what you have seen or heard about the situation in Afghanistan, what should the United States do now -- should the U.S. increase the number of U.S. troops in Afghanistan, keep the number of U.S. troops in Afghanistan as there are now, or decrease the numbers of troops in Afghanistan?” (Cbcpoll)

“How likely do you think it is that there will be another terrorist attack in the United States within the next few months - very likely, somewhat likely, not very likely, or not likely at all?” (Cbcpoll)

“In general, do you think the United States is adequately prepared to deal with another terrorist attack, or not?” (Cbcpoll)

“The U.S. Justice Department has appointed a special prosecutor to investigate the tactics that were used by some people at the CIA when questioning suspected terrorists. Do you think appointing a special prosecutor is a good idea or a bad idea?” (Cbcpoll)

“What is your impression of how the war in Afghanistan is going for the United States right now -- very well, somewhat well, somewhat badly, or very badly?” (Cbspoll)

“I am now going to read you a list of some policy positions that some people in the U.S. are talking about. For each policy, please tell me if you strongly support it, support it, oppose it, or strongly oppose it: Allowing law enforcement officials to search a property without a warrant solely on the suspicion that a terrorist act is being planned there?” (Cid)

“I am now going to read you a list of some policy positions that some people in the U.S. are talking about. For each policy, please tell me if you strongly support it, support it, oppose it, or strongly oppose it: Making it legal for the government to be able to arrest and detain a non-citizen indefinitely if that person is suspected of belonging to a terrorist organization?” (Cid)

“To be a good citizen, how important would you say it is for men to serve in the military when the country is at war?” [Showcard from 0-10 where 0 indicates extremely unimportant and 10 indicates extremely important.] (Cid)

“Using a scale of 0 to 10, where 0 means 'Cannot be trusted at all' and 10 means 'Can be trusted a lot', how much do you trust each of the following groups of people: People who have different religious faith than you?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: If a dictatorship is very dangerous to the world, it is acceptable for democratic governments to use military force to overthrow it?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: If it will better protect Americans from terrorism, the government should reintroduce the draft?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: The measures that the government has taken since 9/11 have helped me feel safer?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: The U.S. was justified in going to war with Iraq?” (Cid)

“Using this card, please say to what extent you agree or disagree with each of the following statements. Do you agree strongly, agree, are uncertain, do you disagree, or disagree strongly with: When America is at war, people should not criticize the government?” (Cid)

“To what extent do you agree or disagree with the following statements about the war in Iraq and the Middle East - Saddam Hussein was somehow involved in the 9/11 attacks?”

(Religion)

“To what extent do you agree or disagree with the following statements about the war in Iraq and the Middle East - the United States was justified in entering Iraq - strongly agree, agree, disagree, strongly disagree?” (Religion)

Rebekah Young

Curriculum Vita

Collaborative Health Studies Coordinating Center
University of Washington
6200 NE 74th Street, Box 354922
Seattle, WA 98115

Email: rlyoung@uw.edu
Phone (CHSCC): 206-685-7123
Phone (Biostatistics): 206-543-1044
Fax: 206-616-4075

Current Position Postdoctoral Researcher, Department of Biostatistics, University of Washington

Education Ph.D., Sociology and Demography, Pennsylvania State University, 2012
M.A., Sociology and Demography, Pennsylvania State University, 2009
B.S., Summa Cum Laude, University of Oregon, 2007

Research Interests Survey methods, quantitative methods, missing data (item and unit nonresponse), family demography, health

Publications

Johnson, David R. and Rebekah Young. 2011. "Toward Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations." *Journal of Marriage and Family*, 73:926-945.

Johnson, David R., Veronica Roth and Rebekah Young. 2011. "Planned Missing Data Designs in Health Surveys." In *Proceedings of the Tenth Conference on Health Survey Research Methods*.

Young, Rebekah and David R. Johnson. 2010. "Imputing the Missing Y's: Implications for Survey Producers and Survey Users." In *Joint Statistical Meetings Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Johnson, David R. and Rebekah Young. 2009. "Improving the Utility of Imputed Values in Survey Datasets." In *Joint Statistical Meetings Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Work Experience

Research Assistant, Dr. David R. Johnson, Penn State University, 2007-2011

Statistical Consultant, Social Science Research Solutions (SSRS), Media, PA, 2011

Research Assistant, Dr. Jeremy Staff, Penn State University, 2011

Survey Statistician (Intern), U.S. Census Bureau, Washington, D.C., 2010

Contractor, Population Reference Bureau, Washington, D.C., 2009

Data Analyst, Social Science Research Institute, University Park, PA, 2008

Research Assistant (Intern), Dr. Marion S. Forgatch, Oregon Social Learning Center, 2007