

Who Has What Information About Others: Proxy Reporting, Knowledge and Willingness¹

Katherine Kenward¹, Eleanor Gerber¹, Alisu Schoua-Glusberg¹, Patricia Goerman², Elizabeth Nichols², Murrey Olmstead³

¹Research Support Services, Inc, 906 Ridge Avenue, Evanston, IL 60202

²U.S. Census, 4600 Silver Hill Rd, Suitland, MD 20746

³RTI International, 3040 East Cornwallis Road, Research Triangle Park, NC 27709-2194

Abstract

The decennial and other surveys conducted by the U.S. Census Bureau typically collect data from households by asking a single household respondent to provide information about others who live in the dwelling. This method of enumeration assumes that the household respondent can act as an accurate proxy for all other household members and that he or she is willing to share information about all household members. This paper explores the cognitive strategies that people use when they are unaware or uncertain of the information they are being asked to provide as proxies and the extent to which it is possible to determine the quality of proxy responses in an actual enumeration. We also explore the reported willingness and/or barriers that exist when reporting for others in the household, especially those unrelated to the proxy. To explore these issues, we use data from cognitive interviews conducted with census questions asking respondents about alternate addresses where household members may live or stay, such as former addresses, seasonal homes, or relatives' homes. We report what respondents think about responding for themselves, their family members, and those living at the same address who are unrelated or only tenuously attached to the household. We also describe strategies that can be used to determine the likelihood that the data are accurate and complete; also we identify alternative data collection strategies that may be warranted for households that include roommates, boarders, or tenuously attached household occupants. Finally, the implications of the findings for the decennial census and other household surveys will be discussed.

Key Words: Decennial Census, Cognitive Testing, Duplication, Census Coverage

1. Introduction

Many surveys in this country collect data from households by asking a single household respondent to provide information about others who live in the dwelling. This method of enumeration assumes that the household respondent can act as an accurate proxy for all other household members and that he or she is willing to share information about household members. Although this may be possible in traditional and stable families, it is not always the case when people are living in situations where they are less familiar with the other household members, or do not feel entitled to divulge information. In this paper we explore the cognitive strategies people use when they are unaware or uncertain

¹*Disclaimer:* This paper was written to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau, Research Support Services, Inc. or RTI International.

of the information they are being asked to provide as proxies, such as the names of the people living in their household and other places where those people live or stay. We also explore the reported willingness and/or barriers that exist when reporting for oneself and others in the household, especially for those not related to the proxy.

To explore these issues, we use data from cognitive interviews conducted using experimental U.S. Census Bureau questionnaire instruments collecting short form decennial census information. Decennial census questionnaires are typically completed by one household member. The instruments tested include questions asking about other addresses where the respondent and other household members may live or stay. These questions are intended to collect data about places where individuals could be enumerated, or may be duplicated, such as former addresses, seasonal homes, or relatives' homes. In this analysis, we summarize what respondents reported for themselves, their family members, and those living at the same address who are unrelated or only tenuously attached to the household. The implications of the findings for the decennial census and other household surveys are also discussed.

Methodology

In 2011, the Census Bureau undertook a project to develop innovative enumeration techniques and to better understand issues of the under- and over-reporting of individuals on census forms, together known as "coverage." Typically, census respondents are given instructions, or asked questions, to identify who should and should not be included in the household, according to census residence rules. Since these rules and the instructions based on them are subject to respondent interpretation, it is occasionally the case that the household is not enumerated accurately according to the rules.² "Undercoverage" occurs when individuals who should be listed are omitted, such as persons staying temporarily who have no other place to live. "Overcoverage" occurs when absent household members are reported, such as college students or members of the military living elsewhere.

The study, still in progress at the time of this analysis, consisted of three iterative rounds of cognitive interviewing with 80 interviews conducted in each round. Only data from the second round of testing were analyzed in this paper.

In order to test the census form with as many different types of households as possible, 20 respondent categories were defined and recruited that represented different types of living arrangements. Some examples are parents with college students who live away at college, multiple families living in one residence, or extended family living together. Of particular interest for this paper were the categories representing living arrangements where the household members are not intimate with each other, such as roommates who share housing for financial reasons rather than because of a preexisting friendship, or a homeowner who rents out a bedroom or basement to boarders. Also of interest were extended family relationships where household members may know some but not all information about their extended kin either because of the distance of the relationship or the tenuousness of the living arrangement.

² "Task Order 001: Qualitative Interviewing with Suspected Duplicates and Cognitive Testing of the Targeted Coverage Follow-up (TCFU) Interview Final Report", Prepared for the U.S. Census Bureau, September 27, 2011 by RTI International and RSS, Inc. RTI Project Number 0212349.001

For the cognitive tests, the Census Bureau used three different form types to collect data from different types of places in which people reside:

1. A Household form in 2010 was mailed to every known U.S. residence. For 2020, Household questionnaires are being tested for both paper (mailed) administration and electronic devices.³ These forms can be completed by the householder or by an enumerator in Non-Response Follow-up (NRFU) operations.⁴
2. A Be Counted form designed to collect data from individuals who do not believe they or their households will be counted elsewhere. For example, a tenant who lives in a basement apartment that does not have a separate address may believe the upstairs landlord would not include him or his family on the census form sent to that address. The respondent can then obtain and complete a Be Counted form to include his family in the census. In rounds 1 and 2 both electronic and paper Be Counted forms were tested.
3. Finally, we tested a Group Quarters (GQ) form, which can be used to capture data about people who reside in facilities such as prisons, college dorms, nursing homes and homeless shelters. This project included only an electronic version of the GQ form.⁵

As is the case in typical census operations, the GQ and Household forms in our study started with an address printed (or prefilled). The form then asked respondents to provide additional addresses for any other places household members had stayed. The Be Counted form required an individual to provide the primary address along with any additional addresses where the person may have stayed around Census Day. The goal of collecting the additional addresses is to obtain other addresses where a household member might also be enumerated so that “unduplication” can later be carried out. In Round 2 of this study, respondents were selected to test a version of the decennial census enumeration forms based on their pre-screened living situations. Respondents were selected if they had living situations involving potentially interesting enumeration problems. These included situations in which household members were highly mobile or tenuously attached, in which core household members lived elsewhere part-time, or in which the household consisted of unrelated persons.

Assessment of Respondent Knowledge

In this paper, we examine the respondents’ ability and willingness to provide particular kinds of information about the people in their households. These include: the full names of household members and any additional addresses where household members may be enumerated. Ability to provide this information was judged in terms of the accuracy and completeness of the responses. Because the data were collected in cognitive interviews, we were able to assess accuracy, based on three sources:

- Standard probes. Respondents were asked about the completeness of names they provided, and were probed about any missing data that was evident in address fields. Standard probes were also included about the sensitivity of questions.
- General understanding of the living situation. Since the respondents had been pre-

³ For this project we tested tablet and smart phone instruments.

⁴ In this project, the paper and device formats differed slightly to accommodate the different formatting needed on the electronic devices however within the device format wording was identical between NRFU and self-report instruments.

⁵ Although the GQ form requires only self-report, we include it in our examination of proxy reporting because respondents are asked for address data of places they have stayed, which often requires them to provide data about other people or households.

screened, enumerators had a small amount of background knowledge about the potential complexity of the person's living situation, and were encouraged to ask questions to elicit a full understanding.

- Spontaneous remarks that respondents made while filling out the forms. This proved particularly fruitful for understanding respondents' habits and attitudes relating to requests for information. We examined respondent reactions to the requests for this information, including sensitivity of the questions; their willingness to report for themselves and as proxy for their immediate family members and others living in the household.

Findings: Name information

The most basic item of information in census enumeration is the name of each person in the household. Complete and accurate name information is critical to the census enumeration for any subsequent matching and unduplication analysis. Particularly in machine editing, the exact form and spelling of the name may influence the likelihood of recognizing a match.

Name information was collected using several different formats in this research. For mobile device enumeration, the respondent was asked for the full first name, full middle name and full last name of each household member. For the paper enumeration, the respondent was asked for full first name, middle initial and full last name. Cognitive probes were used to determine why any name fields were left blank, whether any of the provided names were nicknames or if the respondent had any other issues or sensitivities with providing names for each household member.

In this analysis, we examine the completeness of the names provided according to the closeness of the respondent to the individual about whom, they are reporting. The categories⁶ we have used, and the number of persons in each category are reported below.

1. The respondent (80 people);
2. What we define as "immediate family," including spouses, parents, children including adopted and step children, and unmarried partners (82 people);
3. What we have termed "close non-relatives" such as boyfriends and girlfriends of the respondent and their children, who live with the respondent full or part time (9 people)
4. "Extended family," which we define as cousins, aunts, nephews or nieces, in-laws and other relatives in the extended family (33 people); and
5. "Non-relatives," such as roommates (55 people).

Our hypothesis was that the closer the relationship, the more likely respondents would be to know and provide complete and accurate information. This was largely borne out. As one would expect, respondents were better able to provide complete and accurate full names for themselves and their immediate family members than for others.

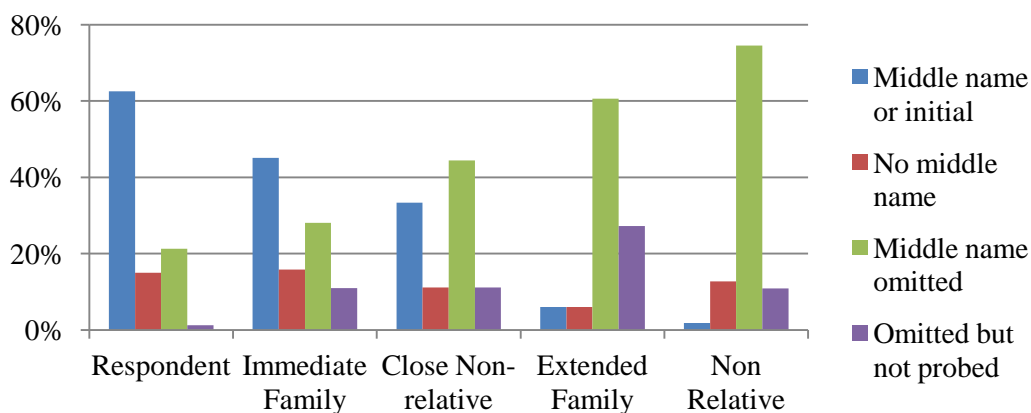
The vast majority of respondents provided full and complete first names for everyone in the household. However, in five of the 259 person pages filled out by respondents, non-standard or inaccurate forms of first names were provided by respondents. For example, two respondents entered nicknames, one respondent provided his own nickname, and another respondent used a nickname for her husband. One respondent, who had two international student boarders, was unable to provide accurate first names for her

⁶ The categories were defined and ordered for this paper on the basis of the likelihood of long term co-residence with the respondent, since this may influence mutual knowledge.

boarders, instead providing their Americanized first names and explaining in probing that they were foreign born and she only knew the names they used to refer to themselves in the U.S. A final respondent chose to write in 'three kids' in a first name field without providing additional information rather than complete an enumeration for three of the people in her household. Only in this last instance would a researcher be able to identify an error with the first name without using follow-up cognitive probes.

Middle names proved the most problematic for respondents. Middle names were omitted most frequently for more distant household members, particularly extended relatives and non-relatives. As one would expect, respondents included the middle name for themselves, their immediate family members and those close non-relatives with whom they lived in family-like relationships more often than with more distant kin and non-related household members. (See Figure 1)

Figure 1: Presence of Middle Name by Relationship to Respondent



Respondents did not generally withhold information about their own middle names if they were willing to provide them for their other household members (regardless of degree of closeness.) In all but one case when respondents provided complete middle name information for household members, they also had included middle name information for themselves. This may indicate that providing middle names is a matter of preference or habit for some people and that if they divulge information about others they will provide it for themselves as well. In fact, when probed, respondents generally expressed a preference for using or not using middle names in "official documents" such as the census. In the only case where a respondent did not report his own information while reporting for others, he explained to the interviewer that he does not like his middle name and thus never uses it on any forms, again a matter of habit. The inverse, however was not the case. Respondents sometimes provided middle names for themselves but not others and gave several reasons for the omission.

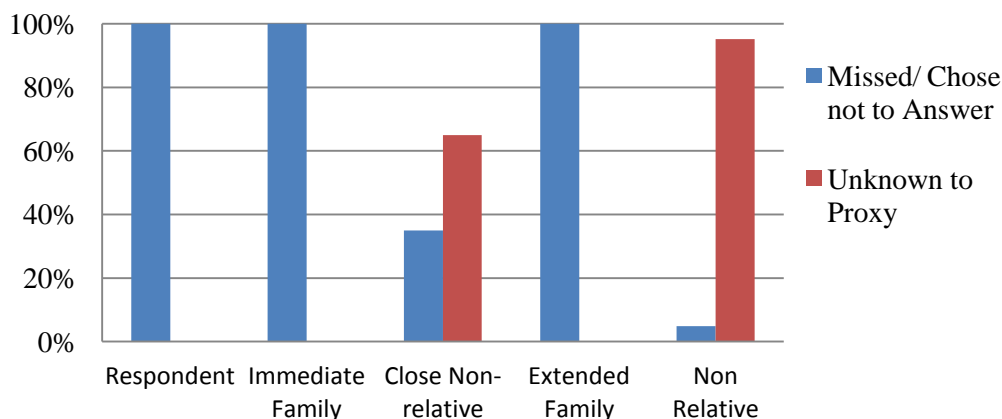
Those who provided middle names seemed to believe it was required or "more accurate," while those who did not provide them often expressed the idea that there was "no need" to supply that much information. Other reasons for not reporting middle name were that they did not like their middle name so avoided reporting it. Not knowing someone else's middle name, and not feeling comfortable giving a middle name for themselves or someone else were also reasons reported.

The reasons respondents gave for omitting the middle name of more distant household members were distinctly different than their explanations for omitting middle names of closer relatives. Figure 2 compares the reasons elicited for omitting middle name across different types of household members. For immediate family, respondents reported that they did not include the middle name because they “missed it,” chose not to answer it or that there was no middle name, the same reasons they gave for themselves. But when reporting on “extended family” and “non relatives” an additional reason emerged; respondents reported that they did not have information about the middle name, that is, they did not know the middle name or were unsure if there was a middle name. For extended family, in cases where a middle name was omitted, respondents reported more than half the time that they were unsure or didn’t know the middle name. Most importantly, for ‘non-relatives,’ in only two cases of omission did the respondent indicate that they actually knew the middle name. In both of those cases the respondent did not feel authorized to divulge the information and noted that they intentionally chose not to put it down. After providing only the first and last name of his roommate, one respondent said that he would not provide the middle name because it was private and should not be provided without permission. A second respondent chose not to provide the middle name for the three people for whom she acted as proxy, although she acknowledged that she knew the names. The interviewer reported that the respondent was more concerned about privacy when providing information about others than about herself.

In one additional case the sensitivity of providing a middle name arose. In this case a respondent explained that he chose not to enter his own middle initial or name. When asked why he hadn’t written a middle initial, the respondent stated:

No, that’s what I always write. If there’s room for the whole middle name, I put it in. But I’d rather not put it in. [PROBE: Why is that?] I don’t know. I mean, why do you need it? I don’t use it.” ...[PROBE: Did you see the box marked MI?] Yes, I noticed it, but I just ignored it.

In cases where respondents claimed not to have seen the middle name/initial field, it was difficult to evaluate whether this was the real reason for their omission. When probed about a missing response, respondents may have found it more convenient or polite to claim they had just “missed” it rather than reporting that they had chosen to ignore it or leave it blank, particularly for close relatives for whom an accidental omission may seem more acceptable than a deliberate one. In addition, we speculate that respondents may find the request to report middle names for household members other than themselves burdensome or unnecessary. In the nine largest households, with over five members each, respondents sometimes reported middle names for themselves or close family members (immediate family and close-non relatives) but never reported middle names for more distant relationships. In addition to the issues already discussed, it is possible that for these large households, their attention to the form became diminished over time as they filled out information for numerous people.

Figure 2: Cited Reason for Omitting Middle Name

Most last name omissions were due to a proxy not knowing the last name of the nonrelatives about whom they were reporting. However in one case, a different issue arose. In this case a Hispanic woman entered only her first of two last names, although in probing she noted that her legal documents include two last names and no middle name as is typical in her country of origin.⁷ In two cases respondents said that they “forgot” to include their own last names. Two people did not provide last names of extended family, with one stating that he was unsure of the last names of his nephews and nieces since they have different names than his sister. In the other case, the respondent refused to provide names for the three children residing only part time in the house because she did not think it was necessary as they were the children of the cousin she had already enumerated and they did not “live” in the household but rather visited for one week per month. Four respondents were unable to report last names for non-relative household members. These household members were not close to the respondent but rather associated with someone else in the house, such as a roommate’s girlfriend, a sister’s friend, a cousin’s girlfriend, and so on. One respondent was unable to report the last names of 10 of the 27 non-related people who lived in her household.

When last names are omitted, except in the case of the Hispanic respondent, the omission is clear to researchers or automated systems. This is not the case with middle names, nicknames, or ‘Americanized’ names used in place of a first name. In these situations the suspect quality of data may be more easily overlooked while appearing accurate and complete.

Although we might wish to assume that people know the names of those who live with them, this may in fact be more common knowledge for people in a family-like living arrangement. It may also point out that Census procedures are somewhat at variance with the expectations of respondents, who do not expect to have to provide information for certain kinds of persons.

⁷ On paper and one of the device versions, first, middle and last names had separate fields for entry. A second version of the device form had a single box where all name information was collected. This Hispanic respondent received administration of the form type where three boxes were labeled: First, Middle and Last.

Findings – Address information

Good quality address data are particularly critical in the census for assessing the proper enumeration of household members. The aim is to determine the proper place to count each individual, and to find other places where they may be duplicated in the census. However, this determination cannot be made unless good address data are available for both the usual residences and for other places where these people may have been listed.

The address formats used in these census instruments provide the respondent several options for providing address information. The traditional postal address format includes a request for county, in addition to city and state. The apartment number is separated from the building number and appears after street name. A separate field asks for “facility name” if that is applicable. In addition, space is provided for descriptive information, such as cross streets in case the respondent does not know certain requested items. This information is designed to provide some information, in the absence of other more accurate data that will allow at least a census block designation to be assigned (See Figure 3).

- 13. If you marked yes above indicating that Person 1 stays somewhere else or has moved, please provide the full address of the other place mentioned in Question 10, 11 or 12. Provide as much address information as possible. If you do not know the full address provide whatever you can such as neighborhood, cross streets or facility name. If there is more than one place, provide the other address where Person 1 was most of the time.**

Street Address (House Number and Street Name)

Apartment Number

City State ZIP Code

County/Township/Parish

Facility name (if necessary):

Other location information (if necessary):

- 14. Where does Person 1 live or stay most of the time?**

- The address printed on the back of this questionnaire
 The address or location listed in Question 13
 Both places equally

Figure 3: Round 2 Paper form address collection fields

When reviewing address data, there are two ways to think about how easy or difficult it might be to provide an address. First, we can consider how important the respondent perceives this information to be and thus how much effort he or she puts into accurately reporting an address. For example, in one case, a respondent took the time to look up an address on her phone. And second, we should not forget that a respondent may wish to complete the information but may not know either all or part of the address.

For both self- and interviewer-administered interviews, most addresses offered were relatively complete⁸. They included at least house number, street name, city and state. ZIP code was usually included whereas, on paper forms where it was requested, county was the most commonly omitted information. However the omission of either zip code or county is generally not problematic as these are easily added.

Overall, good quality addresses were provided by respondents who had actually stayed at the address provided. These included movers, people with second homes (including vacation homes and apartments kept to be closer to work), and with Be Counted respondents, places the respondent had stayed on a temporary basis such as with friends and relatives.

Table 1 shows the quality of the address data provided by respondents, comparing places that the respondent him/herself stayed with places where he/she had not. Respondents were accurately able to report addresses where they had lived or stayed, such as the previously shared homes of ex-spouses or where they still sometimes stayed such as vacation homes.

Table 1: Address quality for places where R has stayed/lived compared to places R does not stay

	<u>R stays</u> ¹	<u>R does not stay</u>
Complete ²	38	13
Missing House Number	1	2
Cross streets or Major neighborhood road	1	2
City and State		4
Unknown address /description 'house'; 'overseas'		4
Unknown/skipped field or wrote 'unknown'		9
Total	40	34

¹ Includes addresses where R has lived in the past such as moves and the previously shared home of ex-spouse

² Complete addresses includes addresses with only the zip code missing

In all but two cases when a respondent stayed at an address he or she provided a complete address for that location. These reports occurred both for themselves, such as when they reported moves or vacation homes, and proxy reports such as reporting for their children who stay with the other parent in the previous family home. Of the two instances in which the respondent did not provide a street address, one was a case where the

⁸ There were minor differences between what was collected among the various form types being tested.

respondent was unsure of an address where she had stayed with friends and so provided only the cross streets. The other case was a respondent who declined to share the street address, citing privacy reasons.

In comparison, in 17 of the 34 times a respondent reported that a household member had stayed somewhere where the respondent had never stayed, s/he was unable to provide any information beyond city. In the thirteen proxy reports where the respondent was not able to provide any address information, they either attempted some form of descriptor such as “mother’s house,” left the field blank, or wrote ‘unknown’ in the address field.⁹

The case of a respondent who answered for himself and as a proxy for his roommate illustrates both the respondent’s ability to report an address because he had stayed at an address and the issue of reporting on more distant household relationships. The roommates are both students and visit their parents on holidays. Although the respondent was able to report the complete address of his parents’ house, he could only report the city and state in which his roommate’s parents lived. In this and so many other households, the concept of a head of household who knows details about all of his household or family members does not fit; instead, this respondent is in a shared home with only limited additional information about the person with whom he lives. One might argue that by the respondent’s definitions of “household,” in which the institution is construed as an interacting economic and social unit, these situations would be described as more than one household living in the same dwelling.

One particular case illustrates how relationships and geographical proximity between addresses can affect accuracy. In this case, a man lived with his mother, his cousin, and his nephew. The respondent’s mother frequently stayed at her boyfriend’s home, the cousin had recently moved in after breaking up with her live-in boyfriend and the nephew had lived both with the respondent and with his paternal grandparents. The respondent was able to report the complete address for his mother’s boyfriend, in the same city where they lived. But he provided only a city and state for his nephew’s grandparents, who lived in another part of the country, where he had never visited.

When respondents have only partial information, their responses can be difficult to interpret. The same respondent was unable to provide a fully accurate alternate address for his cousin. He initially reported what appeared to be a complete address. However, in probing the interviewer confirmed that the street number, which ended in 00, referred to the block, and not the specific address on the heavily populated urban street.

The general location information provided for the cousin’s address, to the block level, may be the most problematic of the three alternative address types provided by this respondent, because it carries the appearance of accuracy while being --in fact-- more generalized address information. Researchers may find it difficult to assess how accurate their data set is when such errors are introduced. Address data at this level may be sufficient for some research needs but will not allow address matching to confirm possible duplications or give enough information for locating purposes since neither the boyfriend’s name, nor other addressee were collected with the address data.

⁹ There were two additional cases which create an editing issue but are not an address quality issues. In these cases, the respondent was able to report a complete address for a relative but for her two children she wrote “same as above.” Presumably these could be dealt with in data editing and so are not addressed here.

As with name data, respondents were generally able to report higher quality address data for people closer to them than for the more distant household relationships, such as the roommate discussed above. (See Figure 4)

Where the respondent was reporting for immediate family, errors tended to be minor omissions. In the two cases where parents were reporting for their children, they were able to provide nearly complete addresses. A mother was able to report all address information except the zip code for her daughter's college dorm and a father was able to report the address of his son's military barracks address.¹⁰

Figure 4: Quality of Address by Relationship to Participant

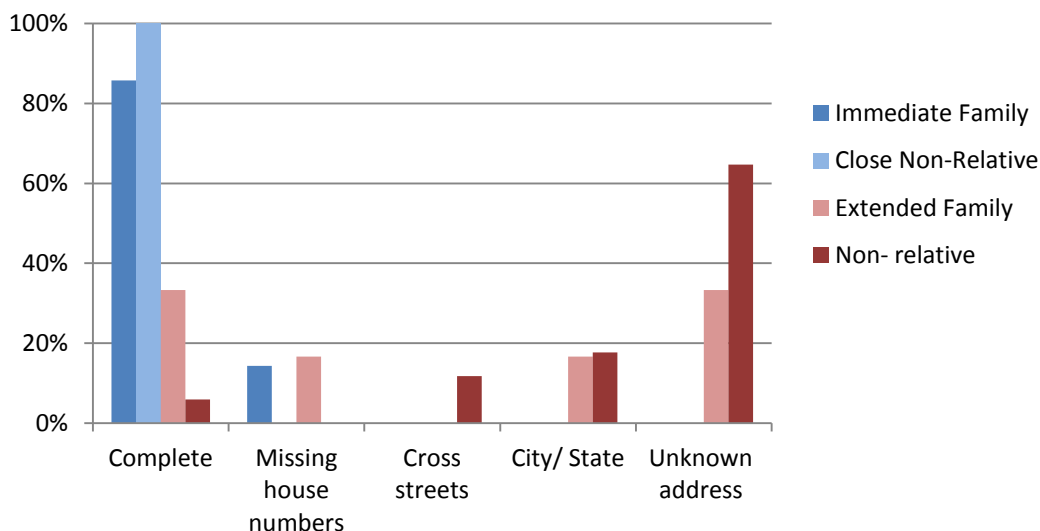


Figure 4 shows that respondents are able to provide complete addresses for some types of household members but not others. Not obvious however, are the reasons why this might be the case. Two of our cases are useful in distinguishing the reasons. In one case a woman shared a home with her boyfriend and their children. Her children from a previous marriage stayed at both her home and her ex-husband's, and her boyfriend's children stayed both at their shared home and their mother's home. In this case, the respondent is involved in the lives of all the children. It was described as a fluid arrangement based on school, work, vacation and sports activities. In a case such as this where driving children to and from their various homes and where all homes are situated in relative proximity, the respondent was able to provide complete alternate addresses for all household members. In comparison, the respondent mentioned above who had two international boarders could provide no additional detail about either their full names or their alternate addresses, although she knew they stayed elsewhere. In both cases, the

¹⁰ There were issues with the format of the address fields that may be problematic with large data sets but are not the focus of this paper. These include, the father of the military son who was not able to make the military barracks address fit in the traditional format of street/city/state/zip code provided, so wrote the address without regard to the instructions. Likewise, a homeless individual wrote explanations about the business address, including the nickname of the owner, for the business outside of which he sometimes slept.

respondent is proxy reporting for non-relatives, but the relationships the respondents have with these household members are completely different. It appears that personal experience with the alternate location is likely to be more important than relationship per se in determining address quality, but both are important.

Address information for people who completed Be Counted and GQ forms had similar outcomes when respondents reported alternate addresses. GQ respondents, who live in group quarters such as dorms or shelters, were asked to report only for themselves and, not surprisingly, they were able to provide complete and full addresses for the other places they had stayed, such as their parent's homes. College students in dorms reported their parents' complete addresses and in one case a friend's address. Only one person living in another type of facility provided a second address: that respondent provided a complete address of the friend's home in which he had stayed.

Be Counted respondents, who have tenuous living situations such as those experiencing homelessness, had the option to report for others, such as their children. Be Counted respondents, generally reported just for themselves and, of the 16 respondents who responded via an automated device, all but three were able to provide a complete initial address where they were currently staying. Of the three, two, who were homeless, provided no street information and one misspelled the street name. Be Counted respondents also have the opportunity to provide an additional address where they stay. Seven respondents provided at least one more address. These addresses included homes of close and extended family members and friends. One respondent did not know his uncle's complete address and noted that he would normally make up the street numbers; he provided street, city and state and zip code. Another respondent was unsure how to indicate that she slept in her car. She entered the word "car" on the space allocated for "If there is no street address, or if this is a facility, please type a description in the box below." The rest were able to report a complete address successfully. This is consistent with the findings in the household cases where respondents were able to provide addresses where they had stayed better than where other people had stayed.

Discussion and Future Directions

When data are missing from a form or questionnaire, especially data provided by proxy, we must ask ourselves, was the data not included because it was not known or because the question was not applicable to the respondent? Was the proxy willing but unable to answer or unwilling to answer? Are we expecting too much of respondents when we ask them to report about others? Have we provided them an opportunity to let us know why they did not complete the form or questionnaire in the way we had hoped? How can we encourage them to provide more information? And we must ask ourselves, do we have a preconceived notion of "household" that leads us to assume that household members have intimate knowledge of others in the household and is that notion reasonable? Further research might help to elucidate the extent to which improving information given to respondents or changes in survey procedures improve the quality of household enumeration.

The cognitive interviews discussed above indicate that we need to keep in mind, as we construct forms and questionnaires, that not all households are "families" and that not all household members have intimate knowledge of each other, travel and commune together, but rather that there are households where the space is shared while the lives are not. Proxy respondents are not always privy to information about others, especially those who are more distantly related or unrelated. Places that the respondent has not

himself visited may be hard for him or her to report on accurately and sometimes those inaccuracies may not be apparent. We must also consider that not all proxy respondents feel comfortable providing information about others, especially those with whom they are not particularly close. Habit and sensitivity also factor into people's decision-making while completing forms, especially when the respondent considers answering for him or herself or close family members. However, lack of knowledge is also important when a respondent is a proxy for persons who are less well known. We cannot tell if the information was correctly omitted or was too difficult or too sensitive to provide. Different kinds of data intersect differently as well when proxy reporting. Simple facts are more likely to be known. Facts are more likely to be known about closely related people. And for complex information, such as addresses, the respondent's direct experience appears to play a greater role in their knowledge.

Future research should explore methods that encourage proxy respondents to report when they do have the information and how to most effectively alert the researcher to the lack of either knowledge or willingness when questions are left unanswered. In addition, the effects of these knowledge patterns on the ability to unduplicate individuals in the census, and to assign them to the proper place, should be examined. If the interaction between this process and the respondent's ability and willingness to provide information were better understood, efforts to elicit information could be targeted to the the most effective and useful data that respondents are, in practice, able to offer.

Acknowledgements

The authors would like to thank Jennifer Hunter Childs, at the Census Bureau and Mandy Sha and Emily McFarlane Geisen of RTI International. We would also like to acknowledge interviewers on this project: Pat Marnell, Dánae Corado and Betsy Strick of Research Support Services, Inc. and Emily McFarlane Geisen, Tim Flanigan, Georgina McAvinchey, and Tiffany King from RTI International. We would also like to acknowledge the recruiters Lisa Andrews, Dánae Corado, and Daniela Glusberg from Research Support Services and Jon O'Connor Headway/RTI. Finally, we would like to thank Thatiana Salgado of Research Support Services for her help in coordinating materials.

References

- Bates, Nancy and Eleanor R. Gerber (1998.) Temporary Mobility and the Reporting of Usual Residence. *Survey Methodology*, June, 1998, Vol 24, No. 1, pp. 89-98
- Fay, R. E. (2004) "An Analysis of Person Duplication in Census 2000". 2004 Proceedings of the Joint Statistical Meetings on CD_ROM, American Statistical Association, Alexandria, VA, pp.3478 -3485.
- Gerber, Eleanor R. (1994.) The Language of Residence: Respondent Understandings and Census Rules. Final Report of the Cognitive Study of Living Situations. Report for the Center for Survey Methods Research, US Bureau of the Census.
- Sha, Mandy., Kenward, Katherine., Feldman, Julie., & Heime Sarah. "Managing quality on a large qualitative research study with complex respondent recruitment criteria." In *JSM Proceedings, 67th Annual Conference of the American Association for Public Opinion Research, American Statistical Association* May, 2012. pp. 5477-5491.