

Can They and Will They? Exploring Proxy Response of Sexual Orientation and Gender Identity in the Current Population Survey

Jessica Holzberg¹, Renee Ellis¹, Robin Kaplan², Matt Virgile¹, and Jennifer Edgar²

Within the United States Federal Statistical System, there has been interest in capturing sexual orientation (SO) and gender identity (GI), collectively known as SOGI, on surveys to allow researchers to estimate the size and distribution of sexual and gender minority populations. SOGI measurement in federal surveys may also help to identify disparities between people who identify as lesbian, gay, bisexual, and transgender (LGBT) and those who do not in domains such as health, crime, or employment. Although research has been conducted on best practices for SOGI measurement in surveys, it has largely been limited to examination of self-reports. Many federal surveys use proxy reports, when one person generally responds for all household members. This research used cognitive interviews and focus groups to explore proxy response to SOGI questions. We explored potential sources of measurement error in proxy responses to SOGI questions, including sensitivity, difficulty, as well as the willingness and ability of respondents to answer SOGI questions about other household members. We also conducted paired interviews with members of the same household to assess level of agreement for SOGI questions. Findings suggest that measuring SOGI by proxy may be feasible in federal large-scale, general population surveys.

Key words: SOGI measurement; proxy reports; federal surveys.

1. Background

Within the United States Federal Statistical System, there has been interest in capturing sexual orientation (SO) and gender identity (GI), collectively known as SOGI, on federal surveys. In this article, we refer to SO in terms of sexual identity, or the way in which people identify with a given sexual orientation, such as lesbian, gay, bisexual, or heterosexual/straight, and to GI as one's personally-identified sense of gender, such as male, female, or transgender (Sexual Minority Assessment Research Team, SMART 2009; Federal Interagency Working Group 2016).

SOGI measurement in federal surveys would allow researchers to estimate the size and distribution of sexual and gender minority populations and identify disparities between

¹ US. Census Bureau, Center for Behavioral Science Methods, 4600 Silver Hill Road, Washington, DC 20233, U.S.A. Emails: Jessica.Holzberg@census.gov, Renee.Ellis@census.gov, Matthew.Virgile@census.gov

² US. Bureau of Labor Statistics, Office of Survey Methods Research, Suite 5930, 2 Massachusetts Ave., NE, Washington, DC 20212, U.S.A. Emails: Kaplan.Robin@bls.gov and Edgar.Jennifer@bls.gov

Acknowledgments: The authors sincerely thank Joanne Pascale, Research Social Science Analyst at the Census Bureau, and Polly Phipps, Senior Survey Methodologist at the Bureau of Labor Statistics, for their significant contributions to the analysis, structure, and writing of this article. The article is released 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 or the Bureau of Labor Statistics.

people who identify as lesbian, gay, bisexual, and transgender (LGBT) and those who do not in domains such as health, crime, or employment. Currently, 11 federal surveys collect data on SO, and among these, seven also ask about GI ([Federal Interagency Working Group 2016](#)). These surveys vary on features such as SO and/or GI question wording, mode of survey response, primary topic of the survey, and population being surveyed. All of these surveys ask respondents to self-report on SO and/or GI and not to report for other household members. However, many federal surveys use proxy reports, when one person generally responds for all household members. Proxy reports of SOGI measures have received less research attention than self-reports.

1.1. Self-Reports of SOGI

A fair amount of research has been conducted on SOGI measurement in a variety of fields, such as psychology, sex research, and survey methods (e.g., [Galupo et al. 2014](#); [Gates 2011](#); [Lombardi and Banik 2016](#); [McCabe et al. 2012](#)). Research has also been conducted on best practices for self-reports of SOGI in federal surveys, primarily conducted in English, in the United States. This includes the Williams Institute at the University of California, Los Angeles's Sexual Minority Assessment Research Team report ([SMART Report 2009](#)), and the Gender Identity in U.S. Surveillance Group report ([GenIUSS Group 2014](#)). Other reports have been written by the Federal Office of Management and Budget's Interagency Working Group on Measuring Sexual Orientation and Gender Identity (e.g., [Federal Interagency Working Group 2016](#)). Generally, most respondents do not seem to have major difficulties answering questions about SOGI for themselves ([Cahill et al. 2014](#)). In addition, SOGI items do not have higher rates of nonresponse than other sensitive questions, such as questions about earnings or income, disability, or health ([Dahlhamer et al. 2014](#); [Joloza et al. 2010](#)), nor do they lead to higher attrition rates in panel surveys ([Joloza et al. 2009](#)). However, for some respondents, SOGI questions may be perceived as sensitive or personal, subject to social desirability bias, or elicit confidentiality concerns ([Tourangeau and Yan 2007](#)). SOGI questions may also be difficult to answer to the extent that respondents exhibit comprehension problems with the terms used in the questions, experience fluidity in SOGI over time, or are not willing to share the information ([Dahlhamer et al. 2014](#)). Overall, the literature suggests that, in both U.S. federal surveys and in other contexts, respondents generally understand SOGI questions, and most are willing and able to answer them.

1.2. Proxy Reports for Non-SOGI Questions

Many federal surveys use proxy response, in which one person responds for all eligible household members. Proxy response is used primarily to reduce costs, time, and nonresponse ([Pierce et al. 1993](#); [Park 2015](#)). Although small differences in agreement between proxy and self-responses are more common than large differences ([Mellow and Sider 1983](#); [Boehm 1989](#); [Tamborini and Kim 2013](#)), the quality of proxy response is difficult to assess due to a lack of systematic studies. For instance, most surveys do not randomly assign who serves as a proxy for the household, but instead survey whoever is available, and most prior research does not examine the validity of responses ([Moore 1988](#); [Cobb 2018](#)). Thus, the literature remains mixed on the quality of proxy responses,

with most studies showing at least moderate agreement between self and proxy reports (Krosnick et al. 2015; Garbarski 2014; Schwarz and Wellens 1997). Like self-response, proxy reporting is also prone to measurement error during the survey response process, where factors such as question sensitivity and difficulty reporting an answer may affect data quality (Lee et al. 2004; Tourangeau, 1984; Bickart et al. 1990).

1.2.1. Sensitivity in Proxy Reporting

One reason for measurement error between self and proxy reports is question sensitivity, or how personal, invasive, threatening, or uneasy a question makes respondents feel, which can lead to increased item nonresponse, refusals, and other threats to data quality (Tourangeau and Yan 2007). Sensitivity in proxy reporting may occur when people feel uncomfortable or reluctant to report personal information about another household member. For example, King et al. (2012) found that proxies did not always feel comfortable revealing personal information about others, such as where they live or how long they lived there, due to privacy concerns. Proxies may not have explicit permission to reveal the information, they may feel they are unable to report accurately on subjective or personal questions about others, or they may fear disclosure to third parties (Sudman et al. 1994; King et al. 2012; Mingay et al. 1994). Thus, sensitivity in proxy reporting can lead to higher levels of unit and item nonresponse, privacy concerns, and other threats to data quality (Todorov and Kirchner 2000).

1.2.2. Difficulty in Proxy Reporting

Proxy respondents may also encounter difficulties when formulating a response and reporting an answer on behalf of other household members. This may occur when proxy respondents have insufficient knowledge to answer a question about another household member (Cobb 2018). For example, in a survey about living situations, proxy respondents provided more “don’t know” responses and less complete information about the household than those who self-reported (King et al. 2012). Level of knowledge about other household members can also be tied to social distance, or how often household members interact, have discussions, and share experiences together (Bickart et al. 1990; Bickart et al. 2006; Pascale 2016). The greater the social distance, the more likely it is that proxy respondents will have insufficient knowledge to report accurately on other household members.

Conducting paired interviews with members of the same household is a common method to understand the accuracy or level of agreement between self and proxy responses, and this research has found that proxies may also find questions difficult to answer because the survey topic is subjective or not observable. For instance, Boehm (1989) conducted paired interviews with household members that contained questions about employment and demographics. On average, all items had an agreement rate of 70% between self and proxy reports. Items that caused the most disagreement tended to be activities that were not directly observable, such as how long the target person had been looking for work, when they last worked, their income, education level, and whether they worked overtime or belonged to a union. Similarly, Davis et al. (2017) examined proxy response across pairs of respondents from the same household on topics not readily observable, such as civic engagement and volunteerism. The researchers created a measure of accuracy based on exact matches (identical responses between the pair) or near

matches (responses separated by one response category) and found an average match rate of 50%. Thus, more communication amongst household members and the observability of the target behaviors on a survey are associated with greater agreement between proxy and self-reports (Kojetin and Tanur 1996).

Finally, difficulties in proxy reporting may occur due to differences in how people process information about themselves versus others (Cobb 2018). The actor/observer bias suggests that people encode information more deeply about themselves than they do about others, leading people to rely more on general information when reporting on others (Jones and Nisbett 1971). This may lessen the motivation to search memory for relevant information about other household members to answer survey questions (Sudman et al. 1994). Unless household members have joint experiences or discuss the question topic frequently, proxies are less likely to have salient and accessible memories to provide information about other household members (Schwarz and Wellens 1997). Proxies may instead use judgment and estimation methods, relying more on heuristics and general knowledge about what they know about others' dispositions to arrive at their answers (Blair et al. 1991; Bickart et al. 1990). This is especially true when proxies have insufficient knowledge to answer for the target person (Todorov 2003).

1.3. Proxy Reports of SOGI Questions

Although there is substantial literature on proxy reporting in general, little is known about proxy reports of SOGI questions, including how sensitive or difficult respondents would find these questions, and whether proxies have the knowledge and willingness to report the information. Qualitative research conducted in the United Kingdom, including interviews and focus groups, has examined the feasibility of including proxy reports of SOGI questions on social surveys. The findings showed that some respondents had concerns about reporting accurately and the confidentiality of responses (Joloza et al. 2009; Joloza et al. 2010; Park 2015).

One of the only quantitative studies conducted on proxy reports of SOGI questions employed an online nonprobability panel to test SOGI questions (Ortman et al. 2017). Respondents were asked to self and proxy report on other eligible household members (age 16 or older). They found overall low rates of nonresponse for the SOGI questions. Item nonresponse for these questions was lower than for income, which is also considered to be a sensitive question (Moore and Welniak 2000). However, nonresponse to the SOGI questions was significantly higher for proxy reports compared to self-reports.

Because SOGI questions may be considered sensitive (SMART Report 2009; GenIUSS Group 2014), this may raise concerns about privacy and confidentiality in proxy response. Proxies may not want to disclose private information about another household member to the federal government, the interviewer, or to other household members who may be present during the interview. Proxies may not have permission from other household members to reveal the information, feel the question invades their privacy, or fear discrimination or stigma.

Factors that contribute to difficulty reporting SOGI for other household members may also affect the data quality of proxy reports of SOGI questions. The quality of SOGI proxy reports may depend in part on the extent to which household members discuss SOGI and

how open household members are with one other. If the information is not shared, observed, or discussed amongst household members, it is more likely to be underreported (Magaziner et al. 1996). In addition, SOGI may be dynamic and subject to changes and fluidity over time (Dahlhamer et al. 2014; Federal Interagency Working Group, 2016). However, a gap in the literature remains regarding the issues surrounding proxy reporting of SOGI questions.

2. Motivation

This research was designed in part to address the gap in the literature on proxy reports of SOGI questions, as well as the feasibility of including SOGI measures in a US federal survey that does not currently include SOGI questions, the Current Population Survey (CPS). The CPS is sponsored jointly by the Bureau of Labor Statistics (BLS) and the U.S. Census Bureau, and serves as the primary source of labor force statistics for the U.S. population. The CPS differs from federal surveys currently collecting SOGI information in its use of proxy response, where one household respondent reports for themselves as well as all other eligible household members age 16 or older. In addition to the CPS, many other large-scale federal household surveys in the United States also employ proxy response. Because little is known about proxy reporting of SOGI information, this remains a major factor for federal surveys considering SOGI questions.

This study sought to explore the processes surrounding proxy response to SOGI questions. Because LGBT individuals have experienced social stigma in the United States, SOGI questions may elicit more privacy and confidentiality concerns compared to other questions typically included on federal surveys (Fisher et al. 2017; SMART Report 2009; GenIUSS Group 2014). These privacy and confidentiality concerns may cause respondent reluctance to self or proxy report SOGI on a federal survey, or to not disclose their status to other household members, creating difficulties in proxy response. This research is a critical step before attempting to collect SOGI data by proxy. By conducting in-depth qualitative research to understand whether respondents are willing and able to report this information, we will be able to determine whether such collection is feasible, where measurement error may arise, and how to mitigate measurement error in the future. Because the size of the LGBT population is estimated to be small relative to the rest of the population, even slight measurement error issues (e.g., under- or overreporting) can have a large impact on estimates (DeMaio et al. 2013). To address this, cognitive interviews and focus groups were conducted to collect information about the sensitivity and difficulty (i.e., willingness and ability to report) associated with reporting SOGI information by proxy, for both LGBT and non-LGBT respondents.

3. Methods

3.1. Sample

We conducted 132 cognitive interviews in and near Washington, D.C., Portland, Oregon; Nashville, Tennessee; and Fargo, North Dakota. These cities were selected to represent different geographic regions of the United States, with the assumption that these regions would also vary on attitudes, political experiences, and other factors that would influence respondent experiences and opinions. Half of the cognitive interviews were conducted with respondents in LGBT households, which we defined as a household with at least one

LGBT person over the age of 15, and half were conducted with respondents in non-LGBT households. Interviews were conducted by staff from the U.S. Census Bureau, the BLS, and a contractor.

Of the 132 cognitive interviews, 52 were conducted with individuals from 26 unique households for “paired interviews.” Each respondent in the cognitive interview pairs was interviewed separately, providing information about themselves, each other, and (in households with three or more members) all the other members of their household. This allowed us to directly compare responses between household members to create a measure of accuracy based on match rates (Davis et al. 2017).

We also conducted four focus groups (one in each of the four cognitive interview cities) with 29 transgender respondents, since very little research has been conducted with transgender respondents in the survey methods literature. Using a focus group methodology was the preferred approach for answering our research questions about gender minorities due to this lack of prior research. Focus groups are typically homogenous to enhance self-disclosure and increase comfort level among respondents. Our transgender-only focus groups were designed to foster a rich group discussion of issues specific to measurement of gender minority status, to allow respondents to build on each other’s comments, to gain a deeper understanding of how gender minorities would react to questions about gender identity, and to identify any sensitivity or difficulty associated with proxy reporting this information. We did not discuss SO in the focus groups because it was examined in depth in the 132 cognitive interviews and because we were constrained by time in the groups.

Recruiting was done by the U.S. Census Bureau or the contractor, using a combination of flyers, [Craigslist.com](https://www.craigslist.com) advertisements, a broadcast message sent to all U.S. Census Bureau employees who work in the Suitland, Maryland headquarters, and posts on Facebook pages for LGBT groups. All respondents were screened prior to being scheduled. Screening included questions on respondents’ age, race, ethnicity, employment status, household size and composition, geographic area (urban versus rural), and LGBT status. For the majority of respondents, geographic area classification was based on whether respondents’ zip code fell within the bounds of Census Bureau defined urbanized areas (50,000 or more people) or urban clusters (2,500–49,999 people). If not, respondents were classified as rural. Respondents’ self-description of their community was used to aid classification in a few instances.

To identify LGBT individuals, we used screener questions that differed from the SOGI questions being tested. For the testing in Washington, D.C., respondents were asked for their gender (male, female, or transgender) and whether anyone in their household age 15 and over, including themselves, identified as LGBT. For the testing in other cities, respondents were asked for their gender (male, female, or transgender), an open-ended SO question, and whether anyone in their household age 15 and over, including themselves, identified as LGBT.

For the cognitive interviews, recruiting equal numbers of respondents in LGBT and non-LGBT households and recruiting paired interview respondents was the most important, but we also set goals for diversity on characteristics such as race, household size and composition, urbanicity, and education, as these factors were hypothesized to have an effect on respondents’ reactions to answering SOGI questions for themselves and other household members. We were successful in meeting these goals (see [Table 1](#) for

Table 1. Respondent characteristics, by interview type.

	Cognitive interviews		Focus groups
	Individual	Paired	
n	80	52	29
LGBT/non-LGBT			
LGBT	32 (40.0%)	33 (63.5%)	29 (100%)
Non-LGBT	48 (60.0%)	19 (36.5%)	0
Age			
18–25	14 (17.5%)	14 (26.9%)	8 (28.6%)
26–35	21 (26.3%)	12 (23.1%)	12 (42.9%)
36–50	25 (31.3%)	14 (26.9%)	5 (17.9%)
Over 50	20 (25.0%)	12 (23.1%)	3 (10.7%)
Race			
White, non-Hispanic	44 (55.0%)	35 (67.3%)	21 (75.0%)
Black, non-Hispanic	13 (16.3%)	7 (13.5%)	4 (14.3%)
Other/multi-race, non-Hispanic	17 (21.3%)	4 (7.7%)	0
Hispanic	6 (7.5%)	6 (11.5%)	3 (10.7%)
Education			
Less than bachelor’s degree	43 (53.8%)	34 (65.4%)	10 (35.7%)
Bachelor’s degree	24 (30.0%)	9 (17.3%)	14 (50.0%)
Higher than a bachelor’s degree	13 (16.3%)	9 (17.3%)	4 (14.3%)
Household size			
Lives alone	0	0	2 (10.0%)
Lives with one other person	28 (35.0%)	26 (50.0%)	9 (45.0%)
Lives with at least two other people	52 (65.0%)	26 (50.0%)	9 (45.0%)
Household composition			
Lives alone	0	0	2 (10.0%)
Lives only with immediate family member(s)	57 (71.3%)	32 (61.5%)	12 (60.0%)
Lives with any extended family or non-family member(s)	23 (28.8%)	20 (38.5%)	6 (30.0%)
Age of household members			
Lives in household with any member age 15–25	39 (48.8%)	28 (53.8%)	15 (53.6%)
Lives in household with no members age 15–25	41 (51.3%)	24 (46.2%)	13 (46.4%)
Region			
Lives in urban area	42 (52.5%)	32 (61.5%)	25 (89.3%)
Lives in rural area	38 (47.5%)	20 (38.5%)	3 (10.7%)

respondent characteristics). Note that of the eight cognitive interview respondents who identified as transgender, seven also identified as LGB; the one transgender respondent who did not identify as LGB was counted as LGBT. For the focus groups, we recruited only transgender respondents to gain deeper insight into this small, hard-to-reach population. As with the cognitive interviews, we tried to recruit respondents with a range of demographic characteristics (see [Table 1](#)). Due to revisions to the recruitment protocol, we had incomplete screener information for 10 focus group respondents.

3.2. *Protocols*

All respondents were provided with a consent form before beginning the interview session. They were also told that information they provided would be confidential and audio-recorded, and were notified if there were any observers.

3.2.1. Cognitive Interview Protocol

In the cognitive interviews, the SOGI questions were embedded into a subset of 46 questions selected from the CPS core instrument. Selected non-SOGI items represented a range of topic areas and varied in terms of their difficulty and sensitivity. We did this in order to mimic the CPS interview, thus providing a realistic context for the SOGI items if they were to be included on the survey, and also to serve as a point of comparison for analysis of SOGI versus non-SOGI items. The topic areas of non-SOGI items included basic demographics (age, marital status, race, etc.), employment, disability, and household income. The SOGI questions were asked relatively early in the questionnaire, after a household roster was collected and questions on date of birth and age were asked.

At the beginning of each cognitive interview, interviewers explained that the purpose of the study was to test new questions developed for the CPS – the primary source of labor force statistics, like employment and unemployment – in the nation. No mention was made at the start about testing of SOGI questions. Then, the standardized questionnaire was administered to respondents by cognitive interviewers via computer-assisted personal interview (CAPI). After completing the standardized questionnaire, cognitive interview respondents completed several debriefing tasks designed to collect information about their response process and reactions. The tasks followed a semi-structured protocol and included a general debriefing, question-specific debriefing, and card sort exercise. In the card sort, respondents were asked to sort index cards listing 15 of the 46 administered CPS questions (including the SOGI questions) into two piles depending on whether they thought the questions were “sensitive” or “not sensitive.” They then ranked the cards in the sensitive pile from most to least sensitive. A second card sort task was then administered, with respondents sorting cards into “difficult” and “not difficult” piles and ranking the difficult cards by most to least difficult. Additionally, respondents in the paired cognitive interviews answered a set of debriefing questions about their reactions to reporting for the other person and having the other person report for them. At the end of the interview, cognitive interview respondents were asked about their reactions and how other people in their household might react to SOGI questions being asked in a federal survey. Detailed information about the protocols can be found in [Ellis et al. \(2017\)](#).

3.2.2. Focus Group Protocol

For each of the four focus groups, a semi-structured protocol was followed that concentrated on questions about GI. The moderator followed the protocol for most of the discussion, but spontaneously added or eliminated probes as needed to encourage discussion and elicit feedback from focus group respondents. There were some modifications between each group, but the main focus of the sections was generally consistent for all four. Each focus group followed the same order, starting with an explanation that the research was one of the first steps being taken as a part of research to understand people's reactions to potentially including questions about GI on one of our national surveys. Focus group respondents introduced themselves and then answered questions about their reactions to adding GI to a government survey about employment. The proxy reporting concept was introduced, and the groups discussed whether they thought other household members would be able and willing to answer GI on their behalf. Detailed information about the protocol can be found in [Holzberg et al. \(2017\)](#).

3.2.3. SOGI Question Wordings

After reviewing wording used in a variety of other federal surveys, we used a common version of the SO question and a two-question approach for GI in the cognitive interviews (see [Figure 1](#)). In the focus groups, we were interested in gauging respondents' reactions to question wording, so we presented respondents with the two-step GI question used in the cognitive interviews, as well as a one-step question and three other versions of the two-step question (see [Figure 2](#), Supplemental material).

3.3. Analysis

3.3.1. Cognitive Interviews

Researchers conducting the cognitive interviews wrote summaries for each individual interview. Summary data were analyzed for evidence of recurring themes and patterns ([Willis 2015](#)), both across cognitive interview respondents and within subgroups (e.g., respondents in LGBT versus non-LGBT households). We also developed a coding scheme to flag the data as sensitive or difficult to answer. For each, we also capture whether the sensitivity or difficulty was related to a self-report or a proxy report.

In general, questions were coded as difficult if there was any evidence that cognitive interview respondents thought that they or others (within or outside the household) would be unable to answer the question, because they either lacked the relevant knowledge or they did not understand the question and/or certain terms in the question. Questions were coded as sensitive if there was any evidence that cognitive interview respondents thought they or others would be unwilling to answer them, or in a few cases if the cognitive interview respondent had a negative emotional reaction to the question (e.g., becoming visibly upset or angry), which was rare. The evidence used for the coding was generally verbal – that is, cognitive interview respondents' responses to a debriefing probe or a spontaneous comment made during the standardized questionnaire. In some instances, the difficulty/sensitivity was directly reported (e.g., "I don't know" or "I would not feel

<p><u>Sexual orientation</u></p> <p>[Self-response]: Which of the following best represents how you think of yourself ?</p> <p>[Proxy response]: To the best of your knowledge, which of the following best represents how [NAME] thinks of themselves?</p> <ul style="list-style-type: none"> • Gay or Lesbian • Straight, that is not gay, lesbian, or bisexual • Bisexual • Something else <p><u>Gender identity</u></p> <p>Question 1: Sex at birth</p> <p>[Self-response]: Was your sex recorded as male or female at birth?</p> <p>[Proxy response]: To the best of your knowledge, was [NAME's] sex recorded as male or female at birth?</p> <ul style="list-style-type: none"> • Male • Female <p>Question 2: Current gender identity</p> <p>[Self-response]: Do you describe yourself as male, female, or transgender?</p> <p>[Proxy response]: To the best of your knowledge, does [NAME] describe themselves as male, female, or transgender?</p> <ul style="list-style-type: none"> • Male • Female • Transgender

Fig. 1. SOGI question wording – cognitive interviews.

comfortable answering”), and other times the difficulty/sensitivity was indicated by the nature of cognitive interview respondents’ comments during debriefing.

For the disability and income questions in the standardized questionnaire, it was somewhat ambiguous whether the questions were asking for self or proxy response because they were asked at the household level (i.e., “Because of a physical, mental, or emotional condition, does anyone have difficulty [insert task]?”). For these questions, it was not always clear whether the cognitive interview respondent found the question difficult/sensitive for themselves, or for other household members. Unless the indication was clearly related to self-response, we coded difficulty as being related to proxy reporting, as the problems identified were related to arriving at an answer about their household as a whole. For sensitivity, on the other hand, we coded comments as pertaining to self-response by default unless cognitive interview respondents were explicit that their reaction was based on the income or disability status of someone else, and not themselves.

Two staff members, working independently, coded the cognitive interview data. Once the independent coding was completed, final consensus codes were assigned through adjudication, either by a third staff member or by a discussion among the coders (e.g., [Kvale and Brinkmann 2015](#); [Saldaña 2015](#)). The coders then counted the number of instances in which cognitive interview respondents indicated that they found a given question sensitive or difficult throughout the interview.

We also evaluated the degree to which answers from paired interview respondents matched and interpreted this as a measure of accuracy and ability to proxy report for both SOGI and non-SOGI questions (see Table 6, Supplemental material, for more details). Responses were categorized as an exact match, near match, or mismatch using a methodology similar to [Davis et al. \(2017\)](#). Criteria for this depended on the complexity of the question. For questions with yes/no response options, a mismatch occurred when the

pair chose different options; there are no possible near matches. For more complex questions such as education and income, response options that were next to each other in the list were considered a near match, and responses that were not next to each other in the list of response options were considered a mismatch. Age was handled similarly. For the SOGI questions and the remaining non-SOGI questions, responses that did not match but would still lead to the same final disposition or classification were considered a near match (e.g., a household member identified as either gay or as bisexual would still lead to this person being classified as LGBT). Discrepant responses that would lead to a different disposition or classification were considered a mismatch (e.g., a household member identified as bisexual versus as straight would lead to this person being classified as either LGBT or as non-LGBT, depending on whose response was used).

Finally, to evaluate respondents' willingness to report SOGI and non-SOGI information, we analyzed item response rates in the standardized questionnaire. We also examined spontaneous comments and responses to probes during the debriefing.

3.3.2. Focus Groups

Following data collection, each focus group was transcribed verbatim. Personally identifiable information (PII) – such as names and places of employment. – were not included in the transcription, and focus group respondents were referred to using ID numbers. The introduction, off-topic comments, and moderator probes were summarized rather than transcribed. We then created a summary document organized by the research questions and sections of the focus group moderator's guide. Each section of this document corresponded to probes or groups of probes. Although a full analysis of the summary document was done to answer several research questions, only the results related to proxy reporting of SOGI information are presented here (full results can be found in [Holzberg et al. 2017](#)).

4. Results

Overall, most cognitive interview respondents understood the SOGI questions as intended, and most did not indicate sensitivity or difficulty when proxy reporting SOGI (see [Table 2](#)). In addition, all cognitive interview respondents were willing to provide answers to the SOGI questions about other household members, and all but one had the information to proxy report. Most of the cognitive interview respondents in paired interviews gave the same responses to the SOGI questions, indicating that they had the knowledge to report this information.

Of the 132 cognitive interview respondents, 105 did not indicate sensitivity proxy reporting SO, and 115 did not indicate sensitivity proxy reporting GI. Most cognitive interview respondents said they were comfortable answering about other household members, and believed SOGI questions were acceptable on a survey like the CPS:

“We are pretty open about everything.”

“This sort of question is becoming more prevalent in society. I don't believe it's a very intrusive question.”

Table 2. Indications of sensitivity and difficulty for proxy and self-reporting, by question – cognitive interviews.

Question	n	Sensitivity		Difficulty	
		Proxy reports	Self reports	Proxy reports	Self reports
SOGI					
Sexual orientation	132	27 (20.5%)	42 (31.8%)	20 (15.2%)	14 (10.6%)
Gender identity (includes sex at birth and current gender identity)	132	17 (12.9%)	27 (20.5%)	6 (4.5%)	11 (8.3%)
Non-SOGI					
Date of birth	132	2 (1.5%)	5 (3.8%)	28 (21.2%)	0
Relationship to reference person	132	3 (2.3%)	0	1 (0.8%)	2 (1.5%)
Marital status	132	6 (4.5%)	9 (6.8%)	2 (1.5%)	10 (7.6%)
Education	132	6 (4.5%)	7 (5.3%)	13 (9.8%)	4 (3.0%)
Hispanic origin	132	6 (4.5%)	3 (2.3%)	2 (1.5%)	3 (2.3%)
Race	132	6 (4.5%)	7 (5.3%)	7 (5.3%)	12 (9.1%)
Boyfriend, girlfriend, partner living in household	132	5 (3.8%)	11 (8.3%)	1 (0.8%)	0
Military service	132	1 (0.8%)	3 (2.3%)	2 (1.5%)	1 (0.8%)
Name of employer	101	5 (5.0%)	10 (9.9%)	5 (5.0%)	4 (4.0%)
Worked for pay	132	2 (1.5%)	6 (4.5%)	9 (6.8%)	10 (7.6%)
Job type	101	0	0	2 (2.0%)	3 (3.0%)
Hours worked	101	0	0	3 (3.0%)	3 (3.0%)
Second job	101	0	0	3 (3.0%)	2 (2.0%)
Could start job	31	0	0	0	1 (3.2%)
Disability (concentrating)	132	58 (43.9%)	35 (26.5%)	23 (17.4%)	21 (15.9%)
Disability (doing errands)	132	42 (31.8%)	27 (20.5%)	16 (12.1%)	9 (6.8%)
Income	132	18 (13.6%)	26 (19.7%)	85 (64.4%)	31 (23.5%)

Not many cognitive interview respondents indicated difficulty answering the SOGI questions. Of the 132 cognitive interview respondents, 112 did not indicate difficulty proxy reporting SO, and 126 did not indicate difficulty proxy reporting GI:

“Certainly not [difficult] for myself, not for the others.”

“[We’re] married [so] it’s pretty clear-cut.”

While indications of sensitivity and difficulty were generally low for SOGI questions, there were more indications of sensitivity and difficulty for cognitive interview respondents in LGBT households (see Table 3). Respondents could have reported difficulty/sensitivity both when self-reporting and when proxy reporting. Interestingly, cognitive interview respondents indicated more sensitivity when self-reporting SO and more difficulty self-reporting GI than when they were proxy reporting. There was no clear pattern in the data to explain this finding, though we hypothesize that this may be

Table 3. Indications of sensitivity and difficulty for proxy and self-reporting to SOGI questions, by household LGBTstatus – cognitive interviews.

	All (n = 132)		LGBT (n = 65)		Non-LGBT (n = 67)	
	Proxy	Self	Proxy	Self	Proxy	Self
Sensitivity						
Sexual orientation	27 (20.5%)	42 (31.8%)	16 (24.6%)	23 (35.4%)	11 (16.4%)	19 (28.4%)
Gender identity (includes sex at birth and current gender identity)	17 (12.9%)	27 (20.5%)	10 (15.4%)	12 (18.5%)	7 (10.4%)	15 (22.4%)
Difficulty						
Sexual orientation	20 (15.2%)	14 (10.6%)	15 (23.1%)	13 (20.0%)	5 (7.5%)	1 (1.5%)
Gender identity (includes sex at birth and current gender identity)	6 (4.5%)	11 (8.3%)	4 (6.2%)	10 (15.4%)	2 (3.0%)	1 (1.5%)

explained in part by respondents not wanting to repeat their comments if they found SOGI sensitive and/or difficult for themselves.

In the following sections, we describe the themes that emerged from the comments of cognitive interview and focus group respondents who indicated sensitivity and/or difficulty proxy reporting. This includes a discussion of differences between cognitive interview respondents in LGBT versus non-LGBT households when observed. We also include a comparison between the sensitivity and difficulty themes that emerged for the SOGI questions versus non-SOGI questions on income, disability, education, and so on. We discuss the level of agreement between the answers of paired cognitive interview respondents, and conclude with an examination of the willingness of respondents to provide proxy responses to SOGI and non-SOGI questions.

4.1. Sensitivity

Although most cognitive interview respondents did not indicate any sensitivity when reporting SOGI for other household members, it was one of the more sensitive questions for respondents, and thus it is important to explore the patterns amongst those who found it sensitive. This may inform future research to reduce sensitivity in surveys considering using proxy to collect SOGI information.

4.1.1. Sexual Orientation

When sensitivity occurred for cognitive interview respondents in non-LGBT households, it was because SO was viewed as a generally sensitive topic that was uncomfortable or inappropriate for people to discuss:

“[My husband] would find the gay and lesbian, the transgender, and the [disability questions] sensitive. . . . He was raised in Alabama as a Baptist.”

“[My elder mother] would feel frustrated by this question. It is not something they talk about. She knows he is gay. . . .but it is not talked about.”

For cognitive interview respondents in LGBT households indicating sensitivity, SO was viewed as more of a personally private matter. Eight cognitive interview respondents, most of whom were from LGBT households, indicated that they were uncomfortable responding about other household members in particular. A few cognitive interview respondents also said that they were uncomfortable choosing a response option on behalf of other household members:

“Feels uncomfortable answering about anyone else, whether they are in the room or not, because it’s a little bit of a personal statement.”

“Would not want to answer for others. [I] would prefer they answer for themselves.”

“Answering for relatives, not knowing exactly how they identify or their own history, was sensitive.”

4.1.2. Gender Identity

The reasons given by the few cognitive interview respondents who found GI sensitive were similar to those given for SO. Cognitive interview respondents in non-LGBT

households found GI to be private because they felt it was a topic that is generally sensitive for many people in the United States, and thus felt uncomfortable discussing it.

“We live in a strange time... [gender] is in everyone’s face.”

“What was your gender at birth. . .when I was born, this would not be asked.”

On the other hand, those in LGBT households viewed GI questions as a personally private matter, and were concerned about the ramifications of answering for household members. Ten cognitive interview respondents living in LGBT households said that their household members would find GI questions sensitive because of their own personal GI, or expressed a preference for household members to answer for themselves instead:

“Because they don’t get a say, don’t know what I’m saying about them, [it] makes it more sensitive for [household member].”

Nearly all transgender focus group respondents found proxy reporting very sensitive, due to general confidentiality concerns, belief that it is inappropriate to answer on someone else’s behalf, the potential to accidentally “out” someone (i.e., disclose a person’s GI without their knowledge or consent), and risks to transgender household members’ safety:

“One hundred percent I would not want anybody to answer this for me at all.”

“[My wife]’s scared of putting out too much because I might get hurt.”

“I would want to know how identifiable this is. Is this just how you’re referring to someone or [do you have] their actual name? I am much more comfortable fitting myself into the boxes that forms have than I am doing that for others. . .I wouldn’t feel comfortable making that choice for someone else.”

Some focus group respondents even commented that they thought people should *not* feel comfortable answering on the behalf of someone who is transgender:

“I would be worried if they were comfortable to answer for me. I wouldn’t really know if I could trust them if they were comfortable to answer this for me.”

4.1.3. Non-SOGI Questions

Most of the non-SOGI questions had similarly low levels of sensitivity, with few cognitive interview respondents indicating any sensitivity when proxy reporting (see [Table 2](#)). Overall, the questions about disability had more indications of sensitivity than the SOGI questions. Income had a similar rate to GI.

“It’s very personal to me because a person in my household has a disability.”

Like SOGI, cognitive interview respondents who found non-SOGI questions sensitive often did not want to answer for other household members because they preferred household members respond for themselves. In some cases, cognitive interview respondents felt uncomfortable because they were concerned they might answer incorrectly for the other person; this reason was most frequently given for disability. In

other cases, cognitive interview respondents thought a household member would find questions sensitive to answer generally, for both themselves and other household members. This reason was most frequently cited for disability and income:

“She does not want people to know her income.”

“My mother does not like to admit her disability.”

4.2. Difficulty

Fewer cognitive interview respondents indicated difficulty than indicated sensitivity when reporting SOGI for other household members. However, it is important to understand the reasons why cognitive interview respondents indicated difficulty, as respondents who find it difficult to proxy report SOGI may be unlikely to be able to provide accurate answers.

4.2.1. Sexual Orientation

Only one of our cognitive interview respondents provided a “Don’t Know” response when proxy reporting SO. This respondent was an LGBT teenager who said they had not talked to their parents about their SO, and thus could not be certain of the correct answer. While all other respondents were able to provide an answer, some cognitive interview respondents from both LGBT and non-LGBT households indicated difficulty proxy reporting because household members’ identities may be fluid. Cognitive interview respondents in LGBT households said that people’s identities may be fluid generally:

“They see sexuality [as] more fluid. They might answer it ‘lesbian,’ might answer ‘bisexual.’”

Cognitive interview respondents in non-LGBT households said this was more because they recognized teenagers or younger members of the household might not have fully developed their identities, and they did not want to assume that they were straight:

“. . . except my son. I don’t want to label him if I don’t know. Until then I assume he’s straight.”

“My daughter is in college; she could be experimenting.”

“My son is still young and society is still not 100 percent accepting, so it is still possible that my son may be bisexual or something rather than straight and not told me.”

Most of the cognitive interview respondents in LGBT households indicating difficulty proxy reporting knew how the household member identified, but did not see a suitable response option in the question or preferred to select more than one option:

“They would want to answer the sexual orientation question as ‘queer.’”

“He is asexual. Straight, and something else.”

4.2.2. Gender Identity

None of our cognitive interview respondents provided “Don’t Know” responses when proxy reporting GI. However, a few cognitive interview respondents in non-LGBT

households indicated difficulty proxy reporting GI because they were unsure of the correct answer. For example, two cognitive interview respondents said they lacked knowledge about their roommates' GI:

“Think I know the answer but I don’t know the roommates very well. They could be transgender, but I don’t think so.”

Three cognitive interview respondents in non-LGBT households indicating difficulty for GI said that older household members would have difficulty answering because of the language used in the questions:

“She would not relate to the language used. . . concepts would be [unfamiliar] to her.”

In the focus groups, transgender respondents said they thought household members would have difficulty responding on their behalf and would likely be inaccurate in their responses. Focus group respondents believed this would be due primarily to a lack of knowledge of household members' correct GI, either generally or within the “transgender” umbrella term, or a refusal to accept and acknowledge household members' GI:

“They’re going to mark you as whatever they see you as, and you’re not getting the authentic [answer] because you haven’t told them what your situation is.”

“It’s very problematic because if you were to ask my dad, his answer would be totally different from mine, and anybody who isn’t out to their head of household, they’re not going to know any of that information or might refuse to answer because they don’t want you to know.”

“I live with 4 roommates, and 3 out of those 4 people, I don’t know if they know or not about my [GI], and I don’t know if they think they know, or if they don’t think they know. I don’t know what they would say. I have literally no idea.”

“If I was living with my parents, absolutely not, but I live with another [transgender] woman as a roommate so I feel like we would both be able to.”

A few cognitive interview respondents gave evidence of similar types of difficulty. One respondent said that other household members would have difficulty because the response options were not inclusive enough and did not allow for the selection of more than one option. Two transgender cognitive interview respondents also expressed that they were uncertain about how their household members would identify them or thought they would be identified incorrectly:

“I don’t know that they would answer that I am transgender or male.”

“My mother would answer as male [though I am gender-fluid]. She would feel certain of that answer. She does not understand all this.”

4.2.3. Non-SOGI Questions

Looking at the difficulty indicated for the non-SOGI questions, income stood out as the most difficult, with more than half of the cognitive interview respondents indicating difficulty proxy reporting (see [Table 2](#)). Date of birth and the disability question on ability

to concentrate had the next most indications of difficulty for the non-SOGI questions. All three had slightly more indications of difficulty than SO, and far more than GI.

Across the non-SOGI questions, the most frequent reason for difficulty was an inability to select from the response categories that were provided. This is similar to what we observed for the SOGI questions:

“There isn’t quite a [race] option because she’s completely Hispanic.”

“[Income options are] overwhelming . . . It’s just very visually challenging in that format.”

Difficulty for non-SOGI questions was also more about a lack of understanding of the questions’ intended meaning than about knowing the answer. We did not observe comprehension difficulties for the SOGI questions:

“Nowadays so many people have mental and emotional difficulties. . .Give examples, or define [disability] a little bit more.”

“OK, so our household is not a family. So how do you want me to break [income] out? . . . OK, so family. That’s me.”

The frequency of “Don’t Know” responses is commonly used to assess the difficulty respondents have reporting. Cognitive interview respondents were able to provide a response when proxy reporting for almost all of the non-SOGI and SOGI questions (see Table 4). Date of birth had the most cognitive interview respondents saying they did not know the answer. The SOGI questions had similarly low “Don’t Know” rates to those found for the non-SOGI questions.

4.3. Match Rates

The majority of answers for paired cognitive interview respondents matched, on both SOGI and non-SOGI questions (see Table 5). Almost all of the paired respondents selected the same response options for both SOGI and non-SOGI questions. The exceptions were income and education, where half or more than half of paired cognitive interview respondents had mismatches. For income, respondents were divided evenly into near

Table 4. Questions with any “don’t know” proxy responses – cognitive interviews (n = 132).

Question	“Don’t know” proxy responses
SOGI	
Sexual orientation	1 (0.8%)
Gender identity (includes sex at birth and current gender identity)	0
Non-SOGI (only questions with any “don’t know” responses shown)	
Date of birth	16 (12.1%)
Education	5 (3.8%)
Hispanic origin	2 (1.5%)
Race	2 (1.5%)
Income	2 (1.5%)

Table 5. Household mismatches and near matches, by question – paired cognitive interviews.

Question	n	Mismatch	Near match	Total with any mismatch
SOGI				
Sexual Orientation	26	3 (11.5%)	2 (7.7%)	5 (19.2%)
Gender Identity (includes sex at birth and current gender identity)	26	1 (3.8%)	2 (7.7%)	2 (7.7%)
Non-SOGI				
Date of birth	26	5 (19.2%)	1 (3.8%)	5 (19.2%)
Age	26	1 (3.8%)	4 (15.4%)	5 (19.2%)
Relationship to reference person	26	0	1 (3.8%)	1 (3.8%)
Marital status	26	0	3 (11.5%)	3 (11.5%)
Education	26	6 (23.1%)	9 (34.6%)	13 (50.0%)
Hispanic origin	26	1 (3.8%)	–	1 (3.8%)
Race	26	4 (15.4%)	–	4 (15.4%)
Number of people in household	26	1 (3.8%)	–	1 (3.8%)
Worked for pay/did not work for pay	26	3 (11.5%)	–	3 (11.5%)
Could/could not start job (only asked for those not currently employed)	12	4 (33.3%)	–	4 (33.3%)
Disability/no disability in HH (concentrating)	26	6 (23.1%)	–	6 (23.1%)
Disability/no disability in HH (dressing/bathing)	26	1 (3.8%)	–	1 (3.8%)
Disability/no disability in HH (doing errands)	26	2 (7.7%)	–	2 (7.7%)
Income	26	8 (30.8%)	8 (30.8%)	16 (61.5%)

matches and mismatches; for education, there were more near matches than mismatches. Most of the mismatches for the non-SOGI questions were due to a lack of knowledge or issues with comprehension of the question, in contrast to the SOGI questions, where comprehension was not a problem and lack of knowledge was less common:

“I don’t know [income]. [We] don’t talk about that.”

“[My partner] may be still technically married.”

“Not sure if I should include [my spouse’s] inheritance [for income].”

There were more mismatches for SO, but near matches were the same between SO and GI. Near matches for SOGI occurred when respondents did not see a preferred response option when reporting for a household member who was a sexual minority, and thus selected different alternatives from the options available (for example, selecting “bisexual” and “something else” for SO, or “male/female” and “male/transgender” for GI).

Reasons for SOGI mismatches were more varied than for the near matches. One case appeared to be a knowledge issue, where a parent identified one of their children as bisexual but their other child identified their sibling as straight. The other two mismatches on SO and the only GI mismatch appeared to be more about willingness to identify a household member as a sexual minority. For the mismatches on SO, in both households the female respondents identified themselves as “bisexual” or “something else” and their male partners identified them as straight; both members of both pairs commented that they were aware that they would not match. For the mismatch on GI, one respondent answered “female/female” for themselves but commented later that they are transgender and just do not like to identify that way; their partner identified them as transgender.

4.4. Willingness to Report

Overall, the cognitive interviews suggest that respondents are willing to proxy report both SOGI and non-SOGI information. No cognitive interview respondent refused to report SOGI for other household members. However, while we did not have a direct measure of willingness in the focus groups, some transgender focus group respondents made comments indicating that they thought members of their household would refuse to proxy report GI on their behalf due to concerns about confidentiality:

“I think my partner would be like, ‘you can [expletive] right off.’”

“I think there’s a strong sense of you don’t disclose other people’s sex or gender; you don’t speak for them, whether it’s a case of accidentally outing them or giving more information than they’re comfortable with.”

Like SOGI, the majority of the non-SOGI questions also received no refusals in the cognitive interviews, with two exceptions. Both refusals occurred due to sensitivity concerns about proxy reporting on behalf of a household member. One refusal was to name the employer, out of concerns about identity theft, and the other refusal occurred because the cognitive interview respondent felt the household member should answer for themselves about their disability.

5. Discussion

Overall, we found that most cognitive interview respondents did not indicate difficulty or sensitivity when proxy reporting SOGI information. Rates of difficulty and sensitivity were similar or slightly higher than observed for other questions. Respondents were generally willing to report SOGI for themselves and other household members, with no respondents refusing and only one respondent saying they did not know the answer. In the paired interviews, there was a high level of agreement in responses. Of the few paired responses that did not match, about half were near matches (a similar rate to other research on paired interviews, e.g., [Davis et al. 2017](#)), where paired respondents still agreed that a person should be classified as a sexual or gender minority. Thus, cognitive interview results suggest that asking SOGI by proxy may be feasible in large-scale, general population surveys.

Of the cognitive interview respondents who did have difficulty, respondents in LGBT households found proxy response more difficult than respondents in non-LGBT households, largely due to issues of question wording. There were also more indications of sensitivity from cognitive interview respondents in LGBT households, who viewed SOGI as a personally private matter. Generally, most difficulty and sensitivity issues for cognitive interview respondents from LGBT households seemed to arise because these respondents were *very much* aware of their household members' LGBT status. Therefore, cognitive interview respondents were concerned about specifying LGBT status accurately and had difficulty when the response options did not meet their needs (e.g., wanting to select more than one option or report a preferred term not provided in the question). They were also concerned about respecting their household members' privacy. We note that transgender respondents in the focus groups raised serious concerns about the difficulty and sensitivity of proxy response for GI, for similar reasons. The results suggest that future research should pay particular attention to how the SOGI questions function differently for subpopulations. Different issues may arise in terms of proxy response measurement error or data quality depending on household composition (e.g., having an LGBT household member or not) and/or social distance.

5.1. Limitations

The present study explored the feasibility of proxy response for a large-scale household survey, the CPS. While this research included some quantitative components, the main data collected were qualitative, and respondents are not meant to be representative of any given population. Cognitive interviews were conducted in select locations (in or near a major US city), excluded single-person households, and all research respondents were paid volunteers willing to respond to a recruitment ad. Respondents may have been more cooperative and comfortable with the collection of SOGI than actual survey respondents outside of the lab setting. Only eight of the 132 cognitive interview respondents were transgender, and only two additional households had transgender household members. However, compared to most qualitative research studies, our sample was large and diverse in terms of demographic, household, and geographic characteristics. Additionally, this study tested the SOGI questions in a single context, within the CPS demographic section, in English, and only as interviewer-administered.

5.2. Future Research

A number of questions remain about the feasibility of SOGI proxy response in large-scale household surveys. There is a clear need to field test how SOGI proxy response would function in the production setting of a real, large-scale household survey; this would provide more information on item nonresponse, response distributions, impact on response rate, attrition, and any systematic bias involved in SOGI measurement that may be magnified for small populations such as sexual and gender minorities.

From our qualitative research, it does not seem as though SOGI proxy response is problematic for most people. There is some difficulty and sensitivity for LGBT respondents, especially for transgender respondents in the focus groups regarding GI. Some of these issues stemmed from question wording; respondents were still generally able to answer whether a person should be considered a sexual or gender minority. As

attitudes towards LGBT individuals have changed rapidly over recent years (Westgate et al. 2015), ongoing research on SOGI measurement is needed, since people's attitudes and perceived sensitivity of these items is unlikely to be static.

The comments from transgender focus group respondents suggest a risk of an undercount of that population, due to both inability and lack of willingness to proxy report the GI of transgender individuals due to sensitivity concerns. However, we note that their comments were made following only a brief description of the CPS and its methodology. Lack of exposure to the CPS interview itself meant we were unable to see how transgender focus group respondents would react to being asked to answer GI by proxy in context. The focus groups were more hypothetical in nature, and opinions may or may not be predictive of behaviors (Fazio 1986; Horwitz and Finamore 2017).

We recommend conducting additional cognitive interviews with respondents in LGBT households, with an emphasis on those with a transgender household member. These interviews would shed more light into the validity of the focus group concerns by giving a more direct measure of proxy response behavior during the survey interview, and would also provide an opportunity to address the question wording concerns from both LGBT cognitive interview and transgender focus group respondents, particularly for GI. Future testing of GI question wording might include the ability to mark all that apply and the inclusion of a "something else" or "other" response option. In addition, conducting this research in other modes, such as self-administered, online questionnaires, may reveal additional findings that could not be observed in traditional, in-person cognitive interviews or focus groups (e.g., Edgar et al. 2016). Respondents may feel more anonymous when answering SOGI questions in an online, self-administered mode, and may provide more honest responses and varying perspectives (e.g., Kreuter et al., 2008; Robertson et al. 2018).

Future research on proxy reporting of SOGI should also explore the role of social distance between household members. For other types of survey questions, the social distance between household members, or how often they interact, have discussions, and share experiences affects proxy reports (Bickart et al. 1990; Bickart et al. 2006; Pascale 2016). Respondents most familiar with other household members, such as spouses, tend to be better proxies than other relatives or household members (Boyle et al. 1992; Kojetin and Mullin 1995; Grieco and Armstrong 2014; Jäger 2005). Other relationships, such as a parent and child, might be subject to other systematic biases. For instance, children may not fully disclose information to their parents, and parents may misreport on their children in a socially desirable direction (Reynolds and Wenger 2012), in particular for socially stigmatizing conditions (Lavtar et al. 2016). These issues could be compounded in households consisting of multiple generations or age groups.

We were unable to explore social distance between household members in the present study because most of our respondents lived in households with immediate family members. For spouses, answering SOGI questions may be relatively simple for the majority of respondents (Ortman et al. 2017). However, for more distant relationships, proxy reports may become more difficult due to insufficient knowledge about other household members (Kojetin and Miller 1993; Kojetin and Tanur 1996). Furthermore, because SOGI questions are about personal identification, which can be fluid over time, and not necessarily observable behavior, the quality of proxy reports may depend in part on the extent to which household members discuss SOGI (Magaziner et al. 1996).

In order to explore the role of social distance between household members in proxy reporting of SOGI, we recommend conducting cognitive interviews with respondents in more complex households (e.g., those with distantly related or unrelated household members). This line of research will be particularly important as households grow more complex, and will benefit all U.S. federal surveys using proxy response that are considering implementing SOGI questions.

6. References

- Bickart, B.A., J. Blair, G. Menon, and S. Sudman. 1990. "Cognitive Aspects of Proxy Reporting of Behavior." In *Advances in Consumer Research 17*, edited by M. Goldberg, G. Gorn, and R. Pollay, 198–206. Provo, UT: Association for Consumer Research.
- Bickart, B.A., J.M. Phillips, and J. Blair. 2006. "The Effects of Discussion and Question Wording on Self and Proxy Reports of Behavioral Frequencies." *Marketing Letters 17*: 167–180. DOI: <https://doi.org/10.1007/s11002-006-5232-1>.
- Blair, J., G. Menon, and B. Bickart. 1991. "Measurement Effects in Self vs. Proxy Responses: An Information-Processing Perspective." In *Measurement Errors in Surveys*, edited by P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman, 145–166. New York: Wiley.
- Boehm, L.M. 1989. "Reliability of Proxy Response in the Current Population Survey." In Proceedings of the Survey Research Methods Section: American Statistical Association, 486–489. Alexandria, VA: American Statistical Association. Available at: https://ww2.amstat.org/sections/srms/Proceedings/papers/1989_086.pdf (accessed February 2018).
- Boyle, C.A., E.A. Brann, and Selected Cancers Cooperative Study Group. 1992. "Proxy Respondents and the Validity of Occupational and Other Exposure Data." *American Journal of Epidemiology 136*: 712–721. DOI: <https://doi.org/10.1093/oxfordjournals.aje.a116550>.
- Cahill, S., R. Singal, C. Grasso, D. King, K. Mayer, K. Baker, and H. Makadon. 2014. "Do Ask, Do Tell: High Levels of Acceptability by Patients of Routine Collection of Sexual Orientation and Gender Identity in Four Diverse American Community Health Centers." *PLoS One 9*: 1–8. DOI: <https://doi.org/10.1371/journal.pone.0107104>.
- Cobb, C. 2018. "Answering for Someone Else: Proxy Reports in Survey Research." In *The Palgrave Handbook of Survey Research*, edited by D.L. Vannette and J.A. Krosnick, 87–93. Cham: Palgrave Macmillan. DOI: <https://doi.org/10.1007/978-3-319-54395-6>.
- Dahlhamer, J.M., A.M. Galinsky, S.S. Joestl, and B.W. Ward. 2014. "Sexual Orientation in the 2013 National Health Interview Survey: A Quality Assessment." *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research 169*: 1–32.
- Davis, M.C., J. Fulton, and A. Henderson. 2017. "Is a Proxy Response Good Enough? Using Paired Cognitive Interviews to Assess the Accuracy of Proxy Responses." Presentation at the DC-AAPOR/WSS Summer Preview/Review Conference. Available at: <http://washstat.org/presentations/20170724/Davis.pdf> (accessed February 2018).
- DeMaio, T.J., N. Bates, and M. O'Connell. 2013. "Exploring Measurement Error Issues in Reporting of Same-Sex Couples." *Public Opinion Quarterly 77*: 145–158. DOI: <https://doi.org/10.1093/poq/nfs066>.

- Edgar, J., J. Murphy, and M. Keating. 2016. "Comparing Traditional and Crowdsourcing Methods for Pretesting Survey Questions." *SAGE Open* 6: 1–14. DOI: <https://doi.org/10.1177/2158244016671770>.
- Ellis, R., M. Virgile, J. Holzberg, D. Nelson, J. Edgar, P. Phipps, and R. Kaplan. 2017. "Assessing the Feasibility of Asking about Sexual Orientation and Gender Identity in the Current Population Survey: Results from Cognitive Interviews." Technical Report, Center for Survey Measurement, U.S. Census Bureau; Office of Survey Methods Research, Bureau of Labor Statistics. Available at: https://www.bls.gov/osmr/pdf/cps_sogi_cognitive_interview_report.pdf (accessed February 2018).
- Fazio, R.H. 1986. "How Do Attitudes Guide Behavior?" In *Handbook of Motivation and Cognition: Foundations of Social Behavior*, edited by R.M. Sorrentino and E.T. Higgins, 204–243. New York: Guilford Press.
- Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys. 2016. "Current Measures of Sexual Orientation and Gender Identity in Federal Surveys." Available at: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/WorkingGroupPaper1_CurrentMeasures_08-16.pdf (accessed February 2018).
- Fisher, A.D., G. Castellini, J. Ristori, H. Casale, G. Giovanardi, N. Carone, E. Fanni, M. Mosconi, G. Ciocca, E.A. Jannini, and V. Ricca. 2017. "Who Has the Worst Attitudes Toward Sexual Minorities? Comparison of Transphobia and Homophobia Levels in Gender Dysphoric Individuals, the General Population and Health Care Providers." *Journal of Endocrinological Investigation* 40: 263–273. DOI: <https://doi.org/10.1007/s40618-016-0552-3>.
- Galupo, M.P., K.S. Davis, A.L. Gryniewicz, and R.C. Mitchell. 2014. "Conceptualization of Sexual Orientation Identity among Sexual Minorities: Patterns across Sexual and Gender Identity." *Journal of Bisexuality* 14: 433–456. DOI: <https://doi.org/10.1080/15299716.2014.933466>.
- Garbarski, D. 2014. "Comparing Self and Maternal Reports of Adolescents' General Health Status: Do Self and Proxy Reports Differ in Their Relationships with Covariates?" *Quality of Life Research* 23: 1953–1965. DOI: <https://doi.org/10.1007/s11136-014-0649-0>.
- Gates, G. 2011. "How Many People are Lesbian, Gay, Bisexual, and Transgender?" Los Angeles, CA: The Williams Institute. Available at: <https://escholarship.org/uc/item/09h684x2> (accessed September 2018).
- GenIUSS Group. 2014. "Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys." Los Angeles, CA: The Williams Institute.
- Grieco, E.M. and D.M. Armstrong. 2014. "Assessing the 'Year of Naturalization' Data in the American Community Survey: Characteristics of Naturalized Foreign Born Who Report – and Don't Report – the Year They Obtained Citizenship." Presentation at the Applied Demography Conference, San Antonio, TX.
- Holzberg, J., R. Ellis, M. Virgile, D. Nelson, J. Edgar, P. Phipps, and R. Kaplan. 2017. "Assessing the Feasibility of Asking about Gender Identity in the Current Population Survey: Results from Focus Groups with Members of the Transgender Population." Technical Report, Center for Survey Measurement, U.S. Census Bureau; Office of

- Survey Methods Research, Bureau of Labor Statistics. Available at: https://www.bls.gov/osmr/pdf/cps_sogi_focus_group_report.pdf (accessed February 2018).
- Horwitz, R. and J. Finamore. 2017. "Do Focus Group Participants Mean What They Say? A Field test of Mailing Materials Updated Based on Qualitative Feedback." Presentation at the Conference of the European Survey Research Association: Lisbon, Portugal.
- Jäger, A. 2005. "Explaining the 'Accuracy' of Proxy-Reports on Attitudes Towards Immigrants in Germany: Two Approaches Compared." *Advances in Methodology and Statistics 2*: 27–57.
- Joloza, T., J. Traynor, and L. Haselden. 2009. "Developing Survey Questions on Sexual Identity: Report on the General Lifestyle (GLF) Split-Sample Pilot." Newport: for National Statistics (Household, Labour Market and Social Wellbeing Division).
- Joloza, T., J. Evans, and R. O'Brien. 2010. "Measuring Sexual Identity: An Evaluation Report." Office for National Statistics, United Kingdom. Available at: <https://www.ons.gov.uk/ons/rel/ethnicity/measuring-sexual-identity--evaluation-report/2010/sexual-identity.pdf> (accessed February 2018).
- Jones, E.E. and R.E. Nisbett. 1971. "The Actor and the Observer: Divergent Perceptions of Behavior." In *Attribution: Perceiving the Causes of Behavior*, edited by E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, and B. Weiner, 79–94. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- King, T., S. Cook, and J.H. Childs. 2012. "Interviewing Proxy versus Self-Reporting Respondents to Obtain Information Regarding Living Situations." In Proceedings for the Joint Statistical Meetings, Survey Research Methods Section, 5667–5677. Alexandria, VA: American Statistical Association. Available at: https://www.amstat.org/sections/srms/proceedings/y2012/files/400243_500698.pdf (accessed February 2018).
- Kojetin, B.A. and L. Miller. 1993. "The Intrahousehold Communications Study: Estimating the Accuracy of Proxy Responses at the Dyadic Level." In Proceedings of the Section on Survey Research Methods, 1095–1100. Alexandria, VA: American Statistical Association. Available at: http://www.asasrms.org/Proceedings/papers/1993_188.pdf (accessed October 2019).
- Kojetin, B.A. and P. Mullin. 1995. "The Quality of Proxy Reports on the Current Population Survey (CPS)." In Proceedings of the Section on Survey Research Methods, 1110–1115. Alexandria, VA: American Statistical Association. Available at: http://www.asasrms.org/Proceedings/papers/1995_193.pdf (accessed October 2019).
- Kojetin, B.A. and J.M. Tanur. 1996. "Proxies for Youths and Adults: Communication and Reports of Job Search." In Proceedings of the Section on Survey Research Methods, 254–259. Alexandria, VA: American Statistical Association. Available at: http://www2.amstat.org/sections/srms/Proceedings/papers/1996_039.pdf (accessed February 2018).
- Kreuter, F., S. Presser, and R. Tourangeau. 2008. "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity." *Public Opinion Quarterly* 72: 847–865. DOI: <https://doi.org/10.1093/poq/nfn063>.
- Krosnick, J.A., S. Presser, K. Husbands, S. Fealing, S. Ruggles, and D.L. Vannette. 2015. "The Future of Survey Research: Challenges and Opportunities." The National Science Foundation Advisory Committee for the Social, Behavioral and Economic Sciences

- Subcommittee on Advancing SBE Survey Research. Available at: http://www.nsf.gov/sbe/AC_Materials/The_Future_of_Survey_Research.pdf (accessed February 2018).
- Kvale, S. and S. Brinkmann. 2015. *InterViews: Learning the Craft of Qualitative Research Interviewing (3rd edition)*. Thousand Oaks, CA: Sage.
- Lavtar, D., M. Zaletel, and A. Rogelj. 2016. "Proxy Responses to Subjective Questions: The Influence on the Results of the Health Expectancy Indicator." European Conference on Quality in Official Statistics, Madrid. Available at: <http://www.ine.es/q2016/docs/q2016Final00147.pdf> (accessed September 2018).
- Lee, S., N.A. Mathiowetz, and R. Tourangeau. 2004. "Perceptions of Disability: The Effect of Self-and Proxy Response." *Journal of Official Statistics* 20: 671–686.
- Lombardi, E. and S. Banik. 2016. "The Utility of the Two-Step Gender Measure within Trans and Cis Populations." *Sexuality Research and Social Policy* 13: 288–296. DOI: <https://doi.org/10.1007/s13178-016-0220-6>.
- Magaziner, J., S.S. Bassett, J.R. Hebel, and A. Gruber-Baldini. 1996. "Use of Proxies to Measure Health and Functional Status in Epidemiologic Studies of Community-Dwelling Women Aged 65 years and Older." *American Journal of Epidemiology* 143: 283–292.
- McCabe, S.E., T.L. Hughes, W. Bostwick, M. Morales, and C.J. Boyd. 2012. "Measurement of Sexual Identity in Surveys: Implications for Substance Abuse Research." *Archives of Sexual Behavior* 41: 649–657. DOI: <https://doi.org/10.1007/s10508-011-9768-7>.
- Mellow, W. and H. Sider. 1983. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics* 1: 331–344.
- Mingay, D.J., S.K. Shevell, N.M. Bradburn, and C. Ramirez. 1994. "Self and Proxy Reports of Everyday Events." In *Autobiographical Memory and the Validity of Retrospective Reports*, edited by N. Schwarz and S. Sudman, 235–250. New York: Springer.
- Moore, J.C. 1988. "Miscellanea, Self/Proxy Response Status and Survey Response Quality: A Review of the Literature." *Journal of Official Statistics* 4: 155–172.
- Moore, J.C. and E.J. Welniak. 2000. "Income Measurement Error in Surveys: A Review." *Journal of Official Statistics* 16: 331–361.
- Ortman, J.M., N. Bates, A. Brown, and R.C. Sawyer. 2017. "Optimizing Self and Proxy Response to Survey Questions on Sexual Orientation and Gender Identity." Presentation at the Annual Meeting of the Population Association of America, Chicago, IL, April, 2017. Available at <https://paa.confex.com/paa/2017/mediafile/ExtendedAbstract/Paper13962/Ortman%20Bates%20Brown%20Sawyer%20PAA%202017.pdf> (accessed February 2018).
- Pascale, J. 2016. "Modernizing a Major Federal Government Survey: A Review of the Redesign of the Current Population Survey Health Insurance Questions." *Journal of Official Statistics* 32: 461–486. DOI: <http://dx.doi.org/10.1515/JOS-2016-0024>.
- Park, J.E. 2015. "Proxy Reporting of Sexual Orientation Information in Household Surveys." Internal Presentation, Office of Statistical and Science Policy, Office of Management and Budget.
- Pierce, J.P., A. Farkas, N. Evans, C. Berry, W. Choi, B. Rosbrook, M. Johnson, and D.G. Bal. 1993. "Tobacco Use in California: A Focus on Preventing Uptake in Adolescents." Sacramento, CA: California Department of Health Services.

- Reynolds, J. and J.B. Wenger. 2012. "He Said, She Said: The Gender Wage Gap According to Self and Proxy Reports in the Current Population Survey." *Social Science Research* 41: 392–411. DOI: <https://doi.org/10.1016/j.ssresearch.2011.10.005>.
- Robertson, R.E., F.W. Tran, L.N. Lewark, and R. Epstein. 2018. "Estimates of Non-Heterosexual Prevalence: The Roles of Anonymity and Privacy in Survey Methodology." *Archives of Sexual Behavior* 47: 1069–1084. DOI: <https://doi.org/10.1007/s10508-017-1044-z>.
- Saldaña, J. 2015. *The Coding Manual for Qualitative Researchers*. Thousand Oaks, CA: Sage.
- Schwarz, N. and T. Wellens. 1997. "Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers." *Journal of Official Statistics* 13: 159–179.
- SMART (Sexual Minority Assessment Research Team). 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys*. Los Angeles, CA: The Williams Institute, UCLA School of Law. Available at: <https://escholarship.org/uc/item/706057d5> (accessed November 2019).
- Sudman, S., B. Bickart, J. Blair, and G. Menon. 1994. "The Effect of Participation Level on Reports of Behavior and Attitudes by Proxy Reporters." In *Autobiographical Memory and the Validity of Retrospective Reports*, edited by N. Schwarz and S. Sudman, 251–265. New York: Springer. DOI: https://doi.org/10.1007/978-1-4612-2624-6_17.
- Tamborini, C.R. and C. Kim. 2013. "Are Proxy Interviews Associated with Biased Earnings Reports? Marital Status and Gender Effects of Proxy." *Social Science Research* 42: 499–512. DOI: <https://doi.org/10.1016/j.ssresearch.2012.11.004>.
- Todorov, A. 2003. "Cognitive Procedures for Correcting Proxy-Response Biases in Surveys." *Applied Cognitive Psychology* 17: 215–224.
- Todorov, A. and C. Kirchner. 2000. "Bias in Proxies' Reports of Disability: Data from the National Health Interview Survey on Disability." *American Journal of Public Health* 90: 1248–1253. DOI: <http://dx.doi.org/10.1002/acp.850>.
- Tourangeau, R. 1984. "Cognitive Sciences and Survey Methods." In *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*, edited by T. Jabine, M.L. Straf, J.M. Tanur, and R. Tourangeau, 73–100. Washington, D.C. National Academy Press.
- Tourangeau, R. and T. Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133: 859–883. DOI: <http://dx.doi.org/10.1037/0033-2909.133.5.859>.
- Westgate, E., R. Riskind, and B. Nosek. 2015. "Implicit Preferences for Straight People over Lesbian Women and Gay Men Weakened from 2006 to 2013." *Collabra: Psychology* 1: 1–10. DOI: <http://doi.org/10.1525/collabra>.
- Willis, G.B. 2015. *Analysis of the Cognitive Interview in Questionnaire Design*. New York: Oxford University Press.

Received March 2018

Revised September 2018

Accepted February 2019