Journal of Official Statistics
vol. 35, 4 (Dec 2019)

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 699–707, http://dx.doi.org/10.2478/JOS-2019-0029

Preface

1. Introduction to the Special Issue on Measuring LGBT Populations

In recent years, much progress has been made in the United States, Europe and beyond with regard to legislation that is supportive and protective of LGBT populations (lesbian, gay, bisexual, transgender persons). For example, in 2011 in the United States, the Institute of Medicine released the watershed monograph "The Health of Lesbian, Gay, Bisexual and Transgender People" (IOM 2011). In 2015, the US Supreme Court ruled that same-sex couples could legally marry in all 50 US states. Furthermore, US public opinion toward gay marriage changed dramatically in a relatively short amount of time. In Pew Research Center polling in 2004, Americans opposed same-sex marriage by a margin of 60% to 31%. By 2019, support for same-sex marriage had flipped with a majority of Americans (61%) supporting same-sex marriage, while 31% oppose it (Pew Research Center 2019). Most recently in the United States, the National Academies of Sciences, Engineering, and Medicine is undertaking a consensus study that will review the available data and future research needs on persons of diverse sexualities and genders.

Also across Europe the legal situation of LGBT persons has improved over the past decades and their visibility has increased in various areas of society (Fischer 2019). Some basic LGBT rights in the European Union (EU) are protected under EU treaties and law. Same-sex sexual activity is legal in all EU Member States and discrimination in employment has been banned since 2000. However, EU Member States have different laws when it comes to any greater protection for same-sex civil unions, same-sex marriage and adoption rights of same-sex couples. For instance, during the last 20 years, legal protection against hate speech towards sexual minorities has increased from five to 23 countries in 2019 (ILGA 2019). With respect to marriage and family laws, same-sex marriage is currently legal in 16 European countries (Pew Research Center 2019). While in 2001, the Netherlands was the first European country to open marriage to same-sex couples, Austria became the latest European nation to legalize the practice at the beginning of 2019. The change in Austria's marriage laws was prompted by its highest court, which in 2017 ruled that the country was discriminating against gay and lesbian couples by not granting them full marriage rights. Concerning the acceptance of same-sex marriage, however, European countries display a huge heterogeneity: with strong support in Sweden (88%), Denmark (86%), and the Netherlands (86%), and distinctly less support in post-socialist countries, such as Poland (32%) and Hungary (27%) (Pew Research Center 2019). Support of same-sex marriage is strongly correlated with acceptance of homosexuality (Hooghe and Meeusen 2013; Takács et al. 2016) in Europe. For instance, acceptance of homosexuality increased much faster in countries where same-sex marriage

is legal after those countries adopted same-sex relationship recognition policies (OECD 2019; Aksoy et al. 2018).

While these achievements are laudable, it is important to keep evaluating to what extent structural obstacles to equality remain. A recent report by the Organisation for Economic Co-operation and Development (OECD) concluded that worldwide, LGBT people report widespread discrimination, are at greater risk for mental health disorders, and have worse labor market outcomes than non-LGBT people (OECD 2019). Yet, compared to research on other minority groups, sexual and gender minorities have been studied quantitatively much less in the social sciences, which in part is related to the numerous challenges associated with collecting suitable survey data (OECD 2019; Umberson et al. 2015).

1.1. How Many People are Gay, Lesbian, Bisexual and Transgender?

A first fundamental challenge is due to the fact that neither sexual orientation nor gender identity are routinely collected in many of the nationally representative surveys sponsored by national statistical agencies. Only 15 OECD countries have included a question on sexual orientation in at least one of their nationally representative surveys sponsored by national statistics offices or other public organizations, and only three OECD countries collect information on gender identity (for a detailed overview, see OECD 2019). Additionally, no census has ever asked questions on sexual orientation and/or gender identity to identify LGB and transgender people (OECD 2019). However, at least some statistical offices are experimenting with it. For instance, in Europe, the United Kingdom is planning to include an item on sexual self-identification in the 2021 Census (ONS 2018). In a related pilot study, the overall conclusion was that including a question on sexual self-identification in the 2021 Census would not significantly impact the overall response and that responses to this question are of acceptable quality. However, such attempts at the national level remain an exception. Instead, the bulk of population-based surveys identify sexual minorities indirectly, using reported sex of a respondent's partner, which is only a sub-population of the total LGB population.

When it comes to cross-nationally comparable data at European level, so far only two surveys directly ask about sexual and gender identity, namely the non-probability based EU LGBT survey conducted in 2013 by the EU Fundamental Rights Agency (FRA 2014) and the periodic probability-based Eurobarometer surveys on discrimination carried out by the European Commission (European Commission 2012, 2015). This remains the largest body of comparative data in Europe and it provides a key source on LGBT experiences across many dimensions of social life. In the United States, the National Health Interview Survey (NHIS) began collecting data on sexual orientation in 2013. To date, the NHIS is the only continuing nationally representative survey of the entire adult population to do so.

1.2. The Lack of Standards How to Best Measure SOGI

Besides the question of how to reach LGBT populations, another challenge is the lack of standards for how best to measure the constructs of sexual orientation and gender identity (SOGI). The Williams Institute has produced two reports with suggested wordings and best practices when asking questions about SOGI in population-based surveys, but both

guidelines are now considered somewhat outdated (Williams Institute 2009, 2014). In the United States, the oversight body that regulates how demographics such as race and ethnicity are to be measured (the Office of Management and Budget, OMB), has no standards when it comes to operational definitions of SOGI. However, in 2015, the OMB did form an Interagency Working Group on Measuring SOGI whose mission was to "explore measurement of SOGI, considering multiple different dimensions of sex, gender and sexuality". This workgroup continues today under the auspices of the Federal Committee on Statistical Methodology and is co-chaired by one of the special issue editors (Nancy Bates). Among other things, the group endeavors to address two of the recommendations from the 2011 IOM report. Specifically, those relevant to survey practitioners and those who depend upon data from nationally representative surveys:

> "Recommendation 2. Data on sexual orientation and gender identity should be collected in federally funded surveys administered by the Health and Human Services and in other relevant federally funded surveys" (IOM 2011, 299)

> "Recommendation 4. National Institute of Health (NIH) should support the development and standardization of sexual orientation and gender identity measures" (IOM 2011, 303).

At the international level, the UN Expert Group on International Statistical Classifications engages in the discussion about concepts and standards about sex and gender. In addition, in Europe, the Bureau of the Conference of European Statisticians (CES) in February 2019 published a working paper on the measurement of gender identity. This is a first collective attempt at the European level to summarize and discuss the main needs for statistical measurement of gender identity, the challenges posed, and the current practices in different countries. Moreover, the report ends with a rather cautious recommendation that future development in this area should be closely monitored by the United Nations Economic Commission for Europe (UNECE 2019).

2. Overview of the Special Issue

The idea for a special issue had its beginning at the 2017 meetings of the European Survey Research Association conference in Lisbon, Portugal. At the conference, the co-editors (Stephanie Steinmetz, Mirjam Fischer and Nancy Bates) organized three sessions devoted to sexual and gender minority populations and subsequently discovered a common interest in publishing a journal issue devoted to the topic. The aim of the special issue is to showcase research around the challenges, successes, and best practices when collecting data on sexual minorities.

The special issue starts with a more general question whether sexual minorities can be considered a hard-to-survey population (with the expectation of lower participation rates); (Magnani et al. 2005; Meyer and Wilson 2009; Tourangeau 2014). While this has often been claimed, there is little empirical evidence regarding this issue. In their contribution *"Are Sexual Minorities Hard-to-Survey? Insights from the 2020 Census Barriers, Attitudes, and Motivators Study"* Bates, Garcia Trejo and Vines showcase for the United States that there is no evidence that sexual minorities required higher levels of effort to secure participation in a survey. On the contrary, it seems that in comparison to straight

respondents, LGB persons are more intent to respond, as measured by intent to participate in the upcoming 2020 Decennial Census. The rest of the contributions to the special issue are centered predominantly around the two above highlighted challenges of how to sample LGBT populations and how to measure SOGI.

With respect to the question of sampling and identifying LGBT populations, the three present contributions address the following two questions: What are common strategies for designing sampling frames intended for capturing LGBT populations, and what advantages and disadvantages in terms of data quality can be detected? The first contribution in relation to sampling *"Test of a Hybrid Method of Sampling the LGBT Population: Web Respondent Driven Sampling with Seeds from a Probability Sample"* by Michaels, Pineau, Reimer, Ganesh and Dennis describes and assesses the results of a pilot study of a new sampling approach that combines an implementation of web-based respondent-driven sampling (RDS) with seeds drawn from a probability-based panel of the US population. This aims to develop a less expensive alternative to full probability sampling that could be used to generate large enough samples of sexual and gender minority persons (including smaller sub-groups) to be able to address a wide range of research questions about these populations. Based on testing of two types of respondent-driven recruiting, the authors conclude that, in principle, both techniques can work to generate new LGBT cases.

The second contribution, by Steinmetz and Fischer *"Surveying Persons in Same-Sex Relationships in a Probabilistic Way – An Example from The Netherlands"* focuses on an approximation sampling strategy for persons in same-sex couples and examines whether this strategy has paid off in terms of reaching the target population, as well as in terms of the quality of the survey data. While the authors conclude that the sampling strategy has paid off by accurately identifying same-sex couple-households, the question of representativeness remains a challenging task in surveying any LGB populations, and couples in particular. The authors point out that especially, aspects related to the sampling strategy, the mode of the survey and the covered topic of the research are central to understanding observed selection patterns in the examined mixed- and same-sex couple samples.

Finally, in the contribution *"Comparing Self-Reported and Partnership-Inferred Sexual Orientation in Household Surveys"* Kühne, Kroh and Richter contribute to the evaluation of the two common strategies of identifying sexual minorities in surveys using the German Socio-Economic panel. The analysis shows, on the one hand, that self-reported and partnership-inferred sexual orientation are not mutual substitutes. Instead, they lead to substantively different conclusions about differences between heterosexuals and LGB persons, which seems to be particularly related to partnership characteristics. On the other hand, the authors are able to show that implementing self-reports of sexual orientation in surveys also comes with the potential of error, as it is sensitive to the data collection mode and interviewer characteristics.

Concerning the question of measuring SOGI in large-scale, general population surveys, the four remaining contributions center on which measure and method is the best and whether sexual orientation can be collected by proxy. All four articles center on production surveys in the United States that produce official statistics. In "*Asking about Sexual Identity on the National Health Interview Survey: Does Mode Matter?*" Dahlhamer, Galinsky and Joestl report findings from a robust split-panel field test designed to measure differences in

sexual minority reporting between interviewer-administered versus a self-interview method with both conditions using computer-assisted interviewing in a personal visit setting. The article explores the prevalence of sexual minority reporting between the two conditions, as well as a subgroup analysis of item nonresponse. Next, Truman, Morgan, Gilbert and Vaghela detail the process of adding both sexual orientation and gender identity to a nationally representative crime victimization survey. "*Measuring Sexual Orientation and Gender Identity in the National Crime Victimization Survey*" details the addition of SOGI in a longitudinal production survey, including qualitative pretesting to minimize measurement error, the exploration of monitoring metrics to gauge degree of missingness, and the examination of population estimates resulting from the new items.

The third article "*Intersections between Sexual Identity, Sexual Attraction, and Sexual Behavior among a Nationally Representative Sample of American Men and Women*" by Mishel focusses more narrowly on the three constructs that comprise the broader concept of 'sexual orientation'. The paper examines how the different self-reports both overlap and diverge, and how estimates of sexual minority prevalence can vary depending upon which measure is used. The findings serve as useful guidelines when planning new data collections of LGB persons. The final article on measurement addresses an important yet understudied topic when surveying sexual and gender minorities, that is, whether SOGI can be accurately collected via proxy. In many population surveys, a single household informant commonly reports demographics, such as age, race/ethnicity, and sex for all household members. How and whether this is possible in the case of SOGI is the topic of "*Can They and Will They? Exploring Proxy Response of Sexual Orientation and Gender Identify in the Current Population Survey*". Holzberg, Ellis, Kaplan, Virgile and Edgar share findings from a large-scale qualitative testing project that included both cognitive interviews and focus groups, the latter of which included transgender persons. The feasibility of collecting SOGI by proxy are examined in terms of sensitivity, difficulty, and willingness to report on behalf of others.

3. Future Outlook – Making LGBT Populations Visible in National and International Statistics

Based on the findings of the special issue, as well as on a follow-up session at the European Survey Research Association conference 2019 in Zagreb, Croatia, a fundamental prerequisite for improving lives of LGBT persons is making them visible in national statistics (OECD 2019). In times when the empirically documented inequality of sexual and gender minorities is frequently at risk of being treated as a matter of opinion, rather than as the product of rigorous scientific work (Perl et al. 2018), it is more important than ever to expand on the methodological repertoire in this field. This calls for fine-tuning of the scientific tools to document, empirically study, and ultimately improve the lives of LGBT persons everywhere. Collectively, the contributions to this special issue not only provide an extensive overview of the many pitfalls that need to be considered, but also show creative approaches to engaging with these challenges. By addressing some of the common concerns of survey providers, which may have thus far prevented them from including SOGI measures into existing and new data collections, we hope to help remove some apprehension around the topic. With this special issue, we invite scientists and

survey providers to engage in the discussion and embrace the many exciting avenues that the inclusion of SOGI offers for future research. We would like to conclude by highlighting just a few of the most pressing directions that would help address the aim of making LGBT populations visible in national and international statistics.

One major future direction needs to be the large-scale testing of translations of SOGI items into languages other than English. The research in the field is heavy on English and a few other European languages, as is the research in this special issue. Large international production surveys have the unique opportunity to include SOGI items into existing structures of quality testing and translation processes. This is an efficient and cost-effective step with immeasurable scientific impact. When thoroughly tested translations become available to the wider international scientific community, it enhances comparability of SOGI data across national contexts. This becomes increasingly relevant as language on the topic of SOGI evolves differently across countries. The availability of thoroughly tested multi-language SOGI measures would allow research communities that operate in various non-English languages to use these items despite a possible lack of financial resources to conduct such testing. Lastly, language minorities within English-speaking countries could be included better if non-English translations of SOGI questions were readily available.

Moreover, research is needed to explore the feasibility of proxy-reports for SOGI questions in large international surveys. Often times, such surveys collect information about the household members via one single respondent who reports this information on others in the household. Since this is a widely used and cost-effective surveying technique, its feasibility for SOGI questions needs to be understood better (see also the contribution of Holzberg, Ellis, Kaplan, Virgile and Edgar on this issue). There is an urgent relevance in knowing the sexual and gender identity of multiple household members to make visible the many forms of cohabiting and family-making that exist among LGBT persons. Considering that LGBT persons face legal obstacles to family formation and need to creatively circumvent these, it is important to examine sexual and gender identities as a family matter and in the context of living arrangements and intimate relationships, as opposed to an individual in isolation.

Another future avenue is to address the growing flux of fluidity in categories to measure SOGI. For example, besides lesbian/gay/bisexual, younger cohorts are rejecting these labels in favor of more, broader and more inclusive labels such as queer, pansexual, asexual, and omnisexual (Trevor Project 2019). Likewise, for gender identity, we see a rejection of the male/female binary in favor of labels like non-binary, genderqueer, and genderfluid,. However, introducing such labels in a general population survey is tricky, and can result in large measurement error among the cisgender population. Future research needs to test empirically how we can better accommodate the diversification of SOGI labels in standardized survey formats and how to weigh the risks of possibly increasing measurement error.

In addition, the temporal fluidity of both sexual and gender identities is a pressing matter for future research. SOGI labels are not as constant as coming out narratives would suggest. Rather than ‘discovering’ a time invariant identity, which is then adopted for all time, identity-making is an ongoing process that can result in the changing of labels over the life course. While this is perhaps more tangible for gender identities, since a non-cis

identity by definition implies a change from an assigned identity to another one, this is also relevant for sexuality labels (Diamond and Savin-Williams 2000). Accommodating such fluidity in standardized survey formats means testing the feasibility of retrospective histories and repeat-measurements over time.

Great potential also lies in examining how probability and non-probability approaches can be combined to survey LGBT persons. Rather than dismissing non-probability approaches as unfit for generating generalizable knowledge, more research needs to focus on the creative integration with probability data from probability surveys (e.g., see Michaels, Pineau, Reimer, Ganesh and Dennis on this issue, and Berzofsky et al. 2019). One major advantage of such an approach is that certain underrepresented groups within the LGBT community, who might be difficult to reach via classical surveys, could be targeted and included more easily. The combination of probability and non-probability approaches would also allow to combat the problem of small group size and insufficient power in statistical analyses, when LGBT persons make up a very small fraction of respondents in large probability-based surveys.

Lastly, government administrative registers could be explored as a source for measuring prevalence and characteristics of gender minorities. A handful of countries that derive population statistics from registers are exploring the expansion of gender options to include a non-binary third response option (UNECE 2019). For example, the Netherlands is exploring the feasibility of adding a third response option. Additionally, some government entities now allow individuals who have transitioned to officially change their sex of record on birth certificates, driver's licenses, pension benefit records, and the like. We recommend that practitioners consider these data as another source in need of attention and research as it relates to SOGI measurement.

The opportunities for future research in this field are ample. We are grateful to the contributors to this special issue for filling important research voids in the SOGI field and we hope it serves as a helpful resource to readers as they navigate this timely topic.

Nancy Bates
Guest Editor

Stephanie Steinmetz
Mirjam Fischer
Guest Associate Editors

4. References

Aksoy, C., Carpenter, C., De Haas, R., and Tran, K. 2019. Do Laws Shape Attitudes? Evidence from Same-Sex Relationship Recognition Policies in Europe. IZA Discussion Paper Series IZA DP No. 11743. IZA: Bonn. Available at: http://ftp.iza.org/dp11743.pdf (accessed October 2019).

Berzofsky, M.E., Scruggs, C.B., McKay, T., and Dever, J.A. 2019. Using Social Media to Increase Sexual and Gender Minority Youth While Maintaining Generalizability: Sampling and Weighting Issues and Solutions. Paper presentation at the 2019 European Survey Research Association conference, Zagreb, Croatia. Available at: https://www.europeansurveyresearch.org/conferences/programme?sess=33#315 (accessed November 2019).

Diamond, L.M. and Savin-Williams, R.C. 2000. Explaining diversity in the development of same-sex sexuality among young women. *Journal of Social Issues* 56(2): 297–313. DOI: https://doi.org/10.1111/0022-4537.00167.

EU FRA (European Union Agency for Fundamental Rights). 2014. *EU LGBT survey – European Union lesbian, gay, bisexual and transgender survey – Main results*. Available at: https://fra.europa.eu/sites/default/files/fra-eu-lgbt-survey-main-results_tk3113640enc_1.pdf (accessed November 2019).

European Commission. 2012. *Discrimination in the EU in 2012 – Special Eurobarometer 393*. Brussels: European Commission. Available at: http://data.europa.eu/88u/dataset/eurobarometer-2012-on-perceptions-of-discrimination-in-the-eu (accessed November 2019).

European Commission. 2015. *Discrimination in the EU in 2015 – Special Eurobarometer and 437*. Brussels: European Commission. Available at: http://data.europa.eu/euodp/en/data/dataset/S2077_83_4_437_ENG (accessed November 2019).

Fischer, M.M. 2019. Free to Live their Lives as they Wish? The Social Well-Being Gap between Persons in Same-Sex and Mixed-Sex Relationships in Europe [Doctoral dissertation]. University of Amsterdam: The Netherlands. ISBN 978-94-92679-72-7.

Hooghe, M. and Meeusen, C. 2013. Is same-sex marriage legislation related to attitudes toward homosexuality? Trends in tolerance of homosexuality in European countries between 2002 and 2010. *Sexuality Research and Social Policy* 10(4): 258–268. DOI: https://doi.org/10.1007/s13178-013-0125-6.

ILGA. 2019. Rainbow Europe 2019. Available at: https://rainbow-europe.org/#1/8693/0 (accessed October 2019).

IOM (Institute of Medicine). 2011. *The Health of Lesbians, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press.

Magnani, R., Sabin, K., Saidel, T. and Heckathorn, D. 2005. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 19: S67–72. ISSN 0269-9370.

Meyer, I. and Wilson, P. 2009. Sampling lesbian, gay, and bisexual populations. *Journal of Counseling Psychology* 56: 23–31. DOI: https://doi.org/10.1037/a0014587.

OECD. 2019. *Society at a Glance 2019: OECD Social Indicators, A Spotlight on LGBT People*. OECD Publishing, Paris. DOI: https://doi.org/10.1787/soc_glance-2019-en.

ONS, Office for National Statistics. 2018. *2017 Census Test Report*. Available at: https://www.ons.gov.uk/census/censustransformationprogramme/2017censustest/2017census-testreport#sexual-orientation-question (accessed October 2019).

Perl, A., Howlett, M. and Ramesh, M. 2018. Policy-making and truthiness: Can existing policy models cope with politicized evidence and willful ignorance in a "postfact" world? *Policy Sciences* 51(4): 581–600. DOI: https://doi.org/10.1007/s11077-018-9334-4.

Pew Research Center. 2019. Attitudes on Same Sex Marriage. Available at: https://www.pewforum.org/fact-sheet/changing-attitudes-on-gay-marriage/ (accessed October 2019).

Takács, J., Szalma, I. and Bartus, T. 2016. Social attitudes toward adoption by same-sex couples in Europe. *Archives of Sexual Behavior* 45(7): 1787–1798. DOI: https://doi.org/10.1007/s10508-016-0691-9.

Trevor Project. 2019. National Survey on LGBTQ Youth Mental Health. Available at: https://www.thetrevorproject.org/survey-2019 (accessed October 2019).

Tourangeau, R. 2014. *Defining hard-to-survey populations*. In *Hard-to-survey populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates, 3–20. Cambridge: Cambridge University Press.

Umberson, D., M.B. Thomeer, R.A. Kroeger, A.C. Lodge, and M. Xu. 2015. Challenges and Opportunities for Research on Same-Sex Relationships. *Journal of Marriage and Family* 77(1): 96–111. DOI: https://doi.org/10.1111/jomf.12155.

UNECE. 2019. Measuring Gender Identity. Working Paper Series on Statistics, Issue 5. Available at: http://www.unece.org/statistics/working-paper-series-on-statistics.html (accessed on October 2019).

Williams Institute. 2009. Best Practices for Asking Questions about Sexual Orientation on Surveys. Available at: https://williamsinstitute.law.ucla.edu/research/census-lgbt-demographics-studies/best-practices-for-asking-questions-about-sexual-orientation-on-surveys/ (accessed October 2019).

Williams Institute. 2014. Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys. Available at: https://williamsinstitute.law.ucla.edu/research/census-lgbt-demographics-studies/geniuss-report-sept-2014/ (accessed October 2019).

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 709–729, http://dx.doi.org/10.2478/JOS-2019-0030

Are Sexual Minorities Hard-to-Survey? Insights from the 2020 Census Barriers, Attitudes, and Motivators Study (CBAMS) Survey

Nancy Bates[1], *Yazmín A. García Trejo*[1], *and Monica Vines*[1]

As a stigmatized and vulnerable population, sexual minorities are often assumed to also be a hard-to-survey population. Despite this implicit assumption, there is little empirical evidence on the topic. Using a nationally representative survey that included sexual orientation (the Census Barriers, Attitudes, and Motivators Survey), we examine level of effort, the Census Bureau's Low Response Score (LRS), and stated intent to respond to the 2020 Census as proxy measures to explore this assumption. We found no evidence that sexual minorities required higher levels of effort to secure participation in the survey. Additionally, we found that compared to straight respondents, lesbians, gays, and bisexuals had a higher intent to respond to the 2020 Census. We surmise the current social climate in the United States may be a contributing factor to these findings.

Key words: SOGI; LGBT; hard-to-count; HTC.

1. Introduction

As survey and census nonresponse continues to increase (Atrostic et al. 2001; Williams and Brick 2018; De Leeuw et al. 2018; Beullens et al. 2018), more methodologists are focusing attention on which populations should be defined as "hard-to-survey". Studies suggest that immigrants, persons experiencing homelessness, renters, and cultural, ethnic, and linguistic minorities, are often classified as such (Rossi et al. 1987; Groves and Couper 1998; Stoop 2005; Massey 2014; Glasser et al. 2014; Harkness et al. 2014; Schwede et al. 2014). Other research (Berry and Gunn 2014; Dewaele et al. 2014) includes vulnerable and stigmatized populations in the hard-to-survey spectrum, which includes sexual minorities. Reports issued by the U.S. Census Bureau also point to sexual minorities among their list of hard-to-count groups requiring extra resources to ensure accurate counts (Moohn 2012; U.S. Census Bureau 2016). However, besides Lee et al. (2018), there is little published on the topic of sexual minorities as a hard-to-survey population – our article adds to this body of research.

When publicly identified as such, stigmatized and vulnerable populations can be subject to social discrimination, physical harm, and other negative outcomes (Berg and Lien 2006;

[1] U.S. Census Bureau, 4600 Silver Hill Rd. Washington, D.C., 20233, U.S.A. Emails: nancy.a.bates@census.gov, yazmin.argen.garcia.trejo@census.gov, and monica.j.vines@census.gov

Acknowledgments: We thank Mary Mulry, Jennifer Shopkorn, the JOS Associate Editor, and two anonymous reviewers for their comments and thoughtful feedback on a previous version. Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. Data for the analysis come from the Census Barriers, Attitudes, Motivators Survey (CBAMS) public use microdata CBDRB-FY18-422 (August 13, 2018).

2008). In the United States and worldwide, sexual minorities report widespread discrimination and greater risk of health disorders (Institute of Medicine 2011; OECD 2019). Additionally, hate crimes against gays, lesbians, bi-sexual, and transgender persons are on the rise in the United States (FBI 2017; Human Rights Campaign 2018; Zauzmer and McCoy 2019). Finally, in the United States, there is no federal law preventing employment discrimination on the basis of sexual orientation and only 26 states (and the District of Columbia) have equal employment laws that ban employment discrimination based on sexual orientation. In sum, being outwardly identified as a sexual minority carries risks regarding general well-being.

In the United States, the 2020 Census form does not include a direct measure of sexual orientation, but it will classify and make estimates of same-sex couples. Two separate questions are used to make these estimates: relationship to householder and sex of household members. Following the 2010 Census, revisions were made to the relationship question to reduce measurement error when counting same-sex couples and the result is a new relationship question with categories that clearly delineate same-sex from opposite sex couples (Bates and DeMaio 2013; DeMaio et al. 2013). As a result, for sexual minorities living with a spouse or unmarried partner, the census form clearly communicates that sexual minorities will be identified in the once-in-a-decade count. This fact, combined with the collection of names matched to addresses raises the question of whether sexual minorities may be hard-to-survey in the Census, (See Figure 1 for the 2020 Census relationship question).

More recent qualitative evidence (Ellis et al. 2017; Holzberg et al. 2017) indicates that many LGB respondents are willing to self-identify with a sexual minority group (even in government surveys) but for some, the stigma remains. According to Ellis et al. (2017,iii) ". . . a few respondents did raise issues about SOGI (Sexual Orientation and Gender Identification) questions, discussing concerns over confidentiality, or mentioning that their responses could be less protected and/or used for discrimination in the current political climate." To our knowledge, only one paper has quantitatively examined whether sexual minorities are less likely than sexual majorities to participate in surveys. Using measures of contactability and reluctance, Lee et al. (2018) found no evidence that sexual minorities

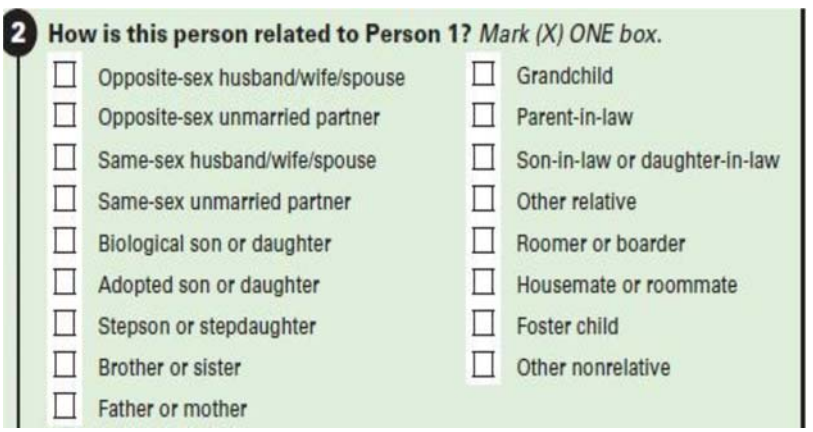


Fig. 1. 2020 U.S. Census relationship question.

had lower participation rates in the U.S. National Health Interview Survey. Our study adds to this literature with a focus on sexual orientation and potential participation in the 2020 Census.

2. Methods

To better understand whether sexual minorities are, in fact, hard-to-survey we explore three proxy measures of such: level of effort, geographic location, and stated intent to participate in the 2020 Census.

With increased availability to survey paradata, it is useful to calculate level-of-effort (LOE) metrics as a proxy way to classify hard-to-survey groups (Kreuter and Olson 2013; Lee et al. 2018). The continuum of resistance theory uses such paradata to study respondents from the tail end of data collection postulating they are similar to nonrespondents in that without the extra time, resources, and effort, they *would* be nonrespondents (Lin and Schaeffer 1995; Meiklejohn et al. 2012). While empirical evidence is mixed on this theory, such metrics do provide clues into late responders' response patterns and behavior.

Hypothesis 1: Sexual minorities will exhibit characteristics consistent with higher levels of effort (LOE).

Next, we will examine the distribution of sexual orientation by whether the sample unit was located in a hard-to-survey census tract. This designation is based on the Census Bureau's Low Response Score (LRS) (Erdman and Bates 2017) – a summary score predicting what percentage of households in a tract will fail to self-respond in the 2020 Census – the higher the LRS score, the harder the tract is to survey.

Hypothesis 2: Sexual minorities will over-represent in areas with a high LRS.

Finally, we will use stated intent to participate in the 2020 Census as a proxy indicator for actual behavior in the Census. Specifically, we examine the hypothesis of whether sexual orientation is a significant predictor of stated intent leading to our final hypothesis:

Hypothesis 3: Sexual minorities will be less likely to indicate a positive intention to participate in the 2020 Census.

2.1. Data: 2020 Census Barriers, Attitudes, and Motivator Study (CBAMS) Survey

Our article analyzes the 2020 CBAMS Survey. The U.S. Census Bureau administered this survey between February 20, 2018 and April 17, 2018 to 50,000 housing units in all 50 states and the District of Columbia. The survey contained questions designed to measure the public's attitudes, knowledge, and opinions regarding the 2020 Census. The results were primarily for the purposes of developing the creative platform and messaging for the 2020 Census Communications Campaign.

The sample design for the survey included stratifying the US population into eight strata based on a census tract's racial and ethnic makeup, as well as characteristics related to Internet response. Each household in the sample received a prepaid incentive and up to five mailings inviting them to participate by mail or Internet in either English or Spanish. We used characteristics related to Internet response to determine whether a

sampled address received an "Internet Choice" mailing strategy, where an invite to respond online was accompanied by a paper questionnaire in the first mailing or an "Internet First" that provided a paper questionnaire in the fourth mailing. (For more information on this methodology, see McGeeney et al. 2019.) There were 17,283 adults who responded to the survey, which was weighted to represent all householder adults in the United States. The final, weighted response rate was 39.4% and was calculated using a modified version of the American Association for Public Opinion Research (AAPOR) RR3 (AAPOR 2016). (See McGeeney et al. 2019 for more information on this survey.)

With the exception of reporting the overall response rate and item nonresponse rate to sexual orientation, all of our analysis and findings in this article use the public use microdata sample (PUMS) data set version of the 2020 CBAMS survey data set. The original data set was made differentially private to avoid disclosure of any personally identifiable information from 2020 CBAMS respondents (Dwork et al. 2006; Dajani et al. 2017). Each variable in the original data set was perturbed using either the multinomial generalization of randomized response mechanism for categorical variables (see McGeeney et al. 2019) or the Laplace mechanism for continuous variables with an epsilon parameter of seven. Findings in this paper include additional error from this disclosure avoidance process. To adjust for the complex sample design, we used SAS PROC SURVEYFREQ, SURVEYREG, and SURVEYMEANS for our analysis. All point and variance estimates take into account the sample design and final weights. Variance estimates were calculated using the Taylor series approximation. Any group differences noted in the text have undergone statistical testing in accordance with U.S. Census Bureau standards (U.S. Census Bureau 2013). A p-value of .05 or less was the threshold for our analysis.

3. Results

3.1. Operationalizing Sexual Orientation

The 2020 CBAMS survey asked sexual orientation as the last question in the survey (see Figure 2). The distribution for sexual orientation from the original data *before* editing, imputation, and noise infusion indicated that item nonresponse in the form of refusals and Don't Knows was low (1.95% and 2.01%, respectively). Table 1 shows the distribution from the CBAMS PUMS (*after* editing, imputation, and noise infusion). For purposes of analyses, we collapse gay or lesbian with the bisexual category to form a single category of sexual minorities. Combined, this group was 3.7% – a number somewhat higher than the percent of sexual minorities found in the 2015 National Health Interview Survey (NHIS), one of the few nationally representative surveys to ask sexual orientation (reported as 2.4%, NHIS 2015).

We also chose to exclude the Don't Know and Something Else cases for analyses. The latter decision was based on an examination of the raw write-ins to the Something Else category which revealed that, of the over 200 non-blank write-ins, only 16% represented other sexual minority labels such as "queer," "pansexual," or "asexual." Instead, the majority were write-ins such as "Christian male," "normal," and "not your business." For purposes of analyses, we used a dichotomous sexual orientation indicator of Lesbian/Gay/Bisexual versus Straight.

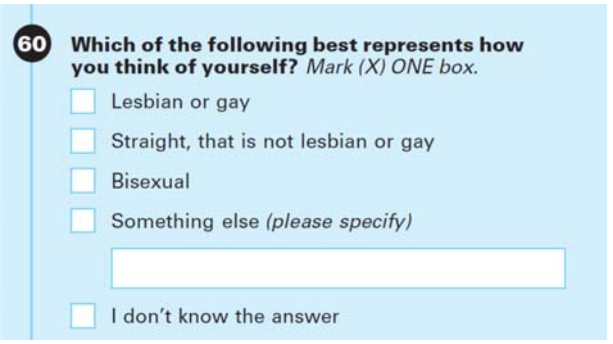


Fig. 2. 2020 CBAMS wording for the sexual orientation question.
Source: https://www2.census.gov/programs-surveys/decennial/2020/program-management/2020_cbams_questionnaire_final.pdf?#

To begin, we explored how sexual minorities compared to sexual majorities in terms of demographic characteristics. Overall, those who identified as a sexual minority skewed male, younger, with higher levels of education, lower income, and less likely to be married compared to those who reported being straight (see Section 6 Appendix, Table 7). Racially and ethnically, however, sexual minorities looked very similar to the rest of CBAMS respondents. In a later section when modeling intention to respond to the 2020 Census, we include many of these demographic variables as controls.

3.2. Level of Effort

Table 2 indicates that sexual minorities overwhelmingly responded by Internet compared to mail and at a higher rate than straight respondents (70.2% versus 61.9%). Online preference may be due, in part, to the fact that over one-third of sexual minorities were aged 18–34. 85% of sexual minority respondents also reported using the Internet "almost constantly/several times per day" compared to 68% of straight respondents (data not shown). Because online responses come in faster and at significantly lower cost than mail,

Table 1. Sexual orientation distribution.

Which of the following best represents how you think of yourself?	Weighted % (std. errors in parentheses)
Lesbian or gay	2.1 (0.134)
Straight	92.1 (0.252)
Bisexual	1.6 (0.119)
Something else	1.5 (0.117)
Don't know	2.6 (0.149)
N (Unweighted)	17283

Source: CBAMS 2018 PUMS file.

Table 2. Percent responding by mail versus Internet by sexual orientation (standard errors in parentheses).

	Gay/lesbian/bi	Straight	Total
Mail	29.8 (2.199)	38.2 (0.416)	37.8 (0.407)
Internet	70.2 (2.199)	61.9 (0.416)	62.2 (0.407)
N (Unweighted)	579	16016	16595
$X^2 = 18.4$, $df = 1$, $p < .001$			

Source: 2018 CBAMS PUMS.

the higher propensity for LGB to choose Internet mode of response is encouraging. Table 3 indicates that, compared to straight respondents, gay/lesbian/bisexual respondents took no more time to respond in either mode – the mean number of days for both Internet and mail did not differ significantly by sexual orientation. Taken together, these results do not indicate a higher LOE for sexual minorities.

3.3. Low Response Score Areas

Next, Table 4 contains the distribution of sexual orientation by whether the sample unit was designated as a "hard-to-survey" census tract. This designation is based on the U.S. Census Bureau's Low Response Score (LRS) (Erdman and Bates 2017) – a summary score predicting what percentage of households in a tract will fail to self-respond in the 2020 Census. The CBAMS used the LRS to stratify the sample by "hard-to-survey" tracts, defined as tracts with an LRS score of 30 or higher. Overall, around 7% of CBAMS respondents were located in such tracts compared to 12.4% who reported being sexual minorities ($X^2 = 34.1$, $df = 1$, $p < .0001$). Because the CBAMS PUMS file has removed the tract identifier, it is not possible to more closely examine these areas to determine if they are concentrated in a particular geography. However, a 2017 report by The Williams Institute (UCLA School of Law) found that LGBT households disproportionately face greater economic challenges and have fewer economic resources than their straight counterparts (Conron et al. 2018). This finding is the first CBAMS evidence to support the notion that sexual minorities tend to have a higher than average prevalence in census tracts classified as "hard-to-survey."

Table 3. Mean number days until response received by mode by sexual orientation.

	Mail		Internet	
	Gay/lesbian/bi	Straight	Gay/lesbian/bi	Straight
Mean days Response time	30.8	32.6	16.0	16.3
	t-value = −1.74, p = .08		t-value = −.46, p = .65	

Source: 2018 CBAMS PUMS.

Table 4. Percent of respondents located in hard-to-survey census tracts by sexual orientation (standard errors in parentheses).

Located in hard-to-survey tract?*	Gay/lesbian/bi	Straight	Total
Yes	12.4 (1.605)	6.5 (0.212)	6.7 (0.212)
No	87.6 (1.605)	93.5 (0.212)	93.3 (0.212)
N (unweighted)	579	16016	16595
$X^2 = 34.1$, $df = 1$, $p < .0001$			

Source: 2018 CBAMS PUMS.
*Hard-to-survey tracts defined as having a Low Response Score = >30 (see Erdman and Bates 2018).

3.4. Intent to Participate in the 2020 Census

Finally, we examine sexual orientation and stated intent to respond in the 2020 Census. Table 5 displays the simple crosstab of sexual orientation by the CBAMS question: *If the census were held today, how likely would you be to fill out the census form?*

Table 5 indicates that close to three-quarters of LGB respondents said they are "extremely" or "very" likely to complete the census form in 2020 (72.2%). Given the implicit assumption that survey response rates are lower among sexual minorities, this was an unexpected finding that we address in the Discussion section.

3.5. Modeling Intent to Respond to Census: Sexual Orientation as a Predictor

Table 6 contains results from two multiple regression models predicting the ordinal five-point scale indicating intent to respond to the 2020 Census. The first column reflects a

Table 5. Distribution of intent to participate in 2020 Census by sexual orientation (standard errors in parentheses).

	Gay/lesbian/bi %	Straight %	Total %
5- Extremely likely	44.2 (2.417)	29.5 (0.406)	30.1 (0.403)
4- Very likely	28.0 (2.170)	38.2 (0.444)	37.8 (0.435)
3- Somewhat likely	23.2 (2.115)	25.0 (0.412)	24.9 (0.404)
2- Not too likely	3.4 (1.033)	5.1 (0.226)	5.0 (0.221)
1- Not at all likely	1.2 (0.676) (0.150)	2.3 (0.147)	2.2
N (unweighted)	579	16016	16595
$X^2 = 68.6$, $df = 4$, $p < .0001$			

Source: 2018 CBAMS PUMS.

Table 6. Regression predicting INTENT (likelihood to participate in the Census along five-point Likert scale).

	Model 1		Model 2	
	Unstd. coeff.	Std. coeff.	Unstd. coeff.	Std. coeff.
Intercept	2.739*** (0.054)	0.000	2.423*** (0.052)	0.000
Predictor: sexual orient. (straight) Gay/lesbian/bisexual	0.314*** (0.048)	0.062	0.231*** (0.046)	0.046
Control variables:				
Age	0.110*** (0.007)	0.172	0.077*** (0.007)	0.120
Female	0.006 (0.018)	0.003	0.062** (0.017)	0.032
Race/Origin (NH, WHITE)				
Hispanic	0.147*** (0.032)	0.050	0.138*** (0.030)	0.047
NH, black	−0.001 (0.030)	0.000	0.083** (0.028)	0.028
NH, asian	−0.280*** (0.045)	−0.054	−0.126** (0.043)	−0.025
NH, other race	−0.052 (0.049)	−0.010	−0.002 (0.045)	0.000
Education	0.127*** (0.006)	0.220	0.048*** (0.006)	0.084
Income	0.040*** (0.005)	0.088	0.006 (0.005)	0.013
Rent (Homeowner)	−0.008 (0.022)	−0.004	0.014 (0.020)	0.007
Foreign Born	−0.062 (0.032)	−0.022	0.007 (0.030)	0.003
Non-English Spoken	−0.074* (0.033)	−0.030	−0.055 (0.030)	−0.022
Marital status (married)				
Div/separated/ widowed	−0.054* (0.022)	−0.025	−0.036 (0.020)	−0.017
Never married	−0.057* (0.028)	−0.024	−0.046 (0.026)	−0.019
Census knowledge			0.113*** (0.003)	0.331
Confidentiality concerns			−0.024** (0.008)	−0.027
Civic participation			0.052*** (0.003)	0.149

Note: ***p-value $<$ 0.001, **p $<$ 0.01, *p $<$ 0.05. Standard errors in parenthesis. NH: stands for Non-Hispanic. Logistic regressions were also performed with 1 = Extremely likely/Very Likely and 0 = Somewhat likely/Not too likely/Not at all likely. Results were very similar with no changes in the direction or significance of Sexual Orientation in either model.

Source: Authors' calculations from CBAMS PUMS 2020.

model containing the dichotomous sexual orientation as a predictor, along with a battery of demographic control variables. Demographic controls were selected based on one of two criteria; first, whether previous studies (as noted in the introduction) have identified the characteristic as being associated with survey nonresponse (e.g., race and ethnic minority, foreign born, and language minority) and/or second, whether the characteristic was found to skew for sexual minorities from the general population (i.e., age, sex, education, income and home ownership, as displayed in Section 6 Appendix Table 7).

In the first model, sexual orientation is a significant predictor with gay/lesbian/bisexual respondents indicating a *higher likelihood* of responding to the 2020 Census compared to straight respondents. While this finding comes with the caveats noted in the limitation section that follows, it is the only empirical evidence we are aware of to measure the relationship between sexual orientation and propensity to cooperate in a survey or census. The demographic only model had an R-square of 0.10.

The second column of Table 6 adds several important constructs found to predict census response behavior in the 2010 Census. In a survey that matched actual census behavior to Census knowledge, attitudes, and exposure to the Census advertising, Datta et al. (2012) reported that the higher the knowledge of Census uses and trust in Census data confidentiality, the higher the likelihood of having completed and mailed back the Census form. Additionally, Bates and Mulry (2012) reported that respondents having a civically-inclined "mindset" were also more likely to mail back a census form. Accordingly, in the second model we include three categorical indices based on three batteries of CBAMS questions. The first indicates the number of correct answers that respondents provided regarding applications/uses of Census data, a second reflects the level of civic participation, and a third measures the degree of concern about the confidentiality of the Census (see Section 6 Appendix for items comprising the indices).

In the second model, sexual orientation (gay/lesbian/bisexual) remains a *positive and significant* predictor of intent to respond. In addition, age, sex (female) and education also had significant and positive coefficients. Somewhat surprisingly, both Hispanics and blacks expressed a higher likelihood to participate compared to non-Hispanic whites, while Asians had a lower stated intent compared to non-Hispanic whites. Neither nativity, non-English language spoken at home, owner/renter status, nor marital status were significant predictors (at the .05 level).

Additionally, all three indices are significant and in line with prior studies. Specifically, higher knowledge and higher civic engagement is associated with higher intent to respond. Alternatively, the higher the privacy concerns, the lower the stated intent to respond. Addition of the knowledge, attitude, and confidentiality measures more than doubles the R-square from .10 to .23. We examined whether sexual minorities varied significantly from straights on all three indices (see Section 6 Appendix, Tables 8, 9, and 10). No significant differences were found for knowledge or confidentiality concerns. However, LGB respondents reported a significantly higher number of activities such as voting, participating in a protest or rally, or volunteering for an organization (see Section Appendix, Table 9).

4. Limitations

Our analyses come with a number of important limitations. First, our data come from a nationally representative sample of US households drawn from the U.S. Census Bureau's

Master Address File (MAF). While the sample frame contains some information (e.g., whether the unit is a single or multiunit, urban or rural) we have no frame information about the residents' sexual orientation. Furthermore, sexual orientation is not a demographic collected in the American Community Survey or Decennial Census. Consequently, we have no "gold standard" benchmarks to assess potential level of nonresponse bias. Second, our inferences are based on sexual minorities willing to identify as such in a government-sponsored survey. These respondents may or may not be similar to sexual minorities as a whole. Additionally, due to the small population size of sexual minorities, we cannot analyze lesbians, gay men, and bisexuals separately. Finally, previous research indicates that *stated intent to participate* in a census does not perfectly align with *actual behavior.* Research suggests that even individuals who report a commitment to participate may not follow through on their intention. Ajzen (1991) argues that there is often a gap between a person's intended and actual behavior, and unforeseen costs and circumstances ultimately prevent some people from carrying out their intended behavior. For example, in a 2010 Census match study, households who reported in a pre-census survey they would "definitely" mail back a census form were found, in fact, to have actually self-responded only 70.8% of the time (Datta et al. 2012).

5. Discussion

Marginalized and stigmatized groups are routinely classified as hard-to-survey populations and sexual minorities are commonly included in this category. Despite this implicit assumption, there is little empirical evidence on the topic. Our article adds to the literature on this assumption by examining whether sexual minorities (1) exhibit characteristics consistent with a population requiring higher levels of effort to secure participation, (2) tend to be located in hard-to-count areas, and/or (3) have lower intention to participate in the 2020 Census.

For the most part, our data did not support our hypotheses. The data did not support the notion that gay, lesbian, and bisexual respondents required higher levels of effort compared to sexual majorities; however, they were found to over-represent in census tracts classified as hard-to-survey as measured by the U.S. Census Bureau's Low Response Score.

We found that over 44.2% of sexual minorities said they were "extremely likely" to participate in the 2020 Census compared to less than 30% of straight respondents. Additionally, we found that sexual orientation (LGB) was a significant (positive) predictor of intent to participate, even after controlling for demographics such as age, nativity, race/ethnicity, education, level of civic participation, census knowledge, and privacy/confidentiality concerns.

To explain this somewhat surprising finding, we turn to the Groves and Couper (1998) conceptual model of survey cooperation. In this model, the social environment plays a critical role in the outcome of a survey request. Surveys are subject to societal change, for example, public opinion among the members of a society, level of trust in government, political alienation, and privacy and confidentiality concerns are all ways to measure the social climate of an environment. In addition, although the social environment is considered to be a fairly fixed attribute that the researcher is powerless to control, Groves and Couper (1998) warn that it should not be ignored. This is because it influences decision-making, its importance changes over time, and it exhibits variation among

subgroups of the population. We believe the social climate is particularly relevant for sexual minorities for a variety of reasons and that may help explain our findings.

First, public opinion in the United States indicates a dramatic change in opinion towards the LGB community. Support towards same-sex marriage has drastically increased in the last decade (See Hatzenbuehler et al. 2017; Masci et al. 2017). Moreover, there is also growing majority support for civil rights for the LGB community. By 2014, "89% of Americans say a gay person should be allowed to give a speech in their community. . . 83% allowed to teach and 81% support keeping a book written by a gay person" (NORC 2014).

Recent research shows that nonresponse rates to sexual orientation questions have declined over time among older adults (Fredriksen-Goldsend and Kim 2014; Fredriksen-Goldsen et al. 2015). There are still age cohort differences, as it is true that compared to younger cohorts, 65+ adults are still more likely to select "don't know/not sure" or refusals. However, even "don't know/not sure" and refusal rates have decreased over time (Saewyc 2011). These trends may be tied to sexual minorities feeling less ostracized in US society, with less stigma when self-identifying as gay, lesbian, or bisexual in survey data collections.

On the other hand, in the United States in the last few years, law officials and watchdog organizations have also reported a sharp increase in hate crimes toward the LGB community. Additionally, the current US administration has enacted policies that negatively impact sexual and gender minorities, including the proposed ban of transgender persons to serve in the military, new state level bans on same-sex couple adoptions, and the policy to end same-sex partner visas of diplomats (Durkin 2018). The outcome of recent and pending court decisions could also have negative outcomes for sexual and gender minorities (Law 2019; Allen 2018; Masterpiece Cakeshop versus Colorado Civil Rights Commission 2018). The result of such social climate events have spurred renewed protests and local activism by the LGB community. This may explain why a large portion of sexual minorities exhibited higher participation in activities such as voting, signing petitions, volunteering in organizations, and participating in a protest or rally.

The inclusion of a sexual orientation question and a sufficiently large sample size in the CBAMS enabled the U.S. Census Bureau to obtain research-based insights to specifically inform tailoring and targeting to sexual minorities for the first time. In terms of messaging, sexual minorities were more likely than straight respondents to select the following two reasons as most motivating: 'It determines how many elected representatives my state has in Congress' and 'It is used to enforce civil rights laws'. As such, messages in introductions, prenotices, survey invitations and paid ads specifically developed for sexual minorities should emphasize these themes. Additionally, sexual minorities were more likely to prefer online forms to paper ones, making mode targeting an important consideration – particularly in the 2020 Census, as the form can be submitted online. Finally, advertisements will be aimed at media specifically consumed by the LGB population with imagery that resonates (e.g., featuring same sex couples); partners and trusted voices closely connected to this community will also relay the importance of participation in the 2020 Census.

Given the scarcity of information about sexual minorities and surveys, we view our results as another step toward answering the question of whether sexual minorities are a hard-to-survey population. Given the polarized components of the social climate that affect sexual minorities, we cannot say definitively whether the stigma and prejudice has

decreased enough to rethink the assumption of LGBs as a hard-to-survey population. However, we did uncover clues that may help ensure their participation in future censuses and surveys.

6. Appendix

Table 7. Demographic distributions (%) by sexual orientation: CBAMS PUMS 2018 (standard errors in parentheses).

	Gay/lesbian/bi %	Straight %	Total %
Male	58.68 (2.412)	49.09 (0.460)	49.46 (0.452)
Female	41.32 (2.412)	50.91 (0.460)	50.54 (0.452)
Total	579	16016	16595
$X^2 = 22.7$, $df = 1$, $p = <.0001$			
18–24	10.80 (1.759)	3.57 (0.211)	3.85 (0.214)
25–34	26.09 (2.260)	13.88 (0.361)	14.35 (0.358)
35–44	19.60 (2.020)	16.89 (0.380)	16.99 (0.373)
45–54	16.43 (1.675)	16.47 (0.326)	16.47 (0.320)
55–64	16.66 (1.677)	23.01 (0.377)	22.76 (0.368)
65+	10.42 (1.197)	26.19 (0.363)	25.58 (0.353)
Total (unweighted)	579	16016	16595
$X^2 = 221.06$, $df = 5$, $p = <.0001$			
Hispanic	12.58 (1.538)	12.43 (0.285)	12.44 (0.280)
Non-Hispanic-White	66.00 (2.339)	68.77 (0.385)	68.66 (0.378)
Non-Hispanic-Black	13.18 (1.841)	11.99 (0.278)	12.03 (0.274)
Non-Hispanic-Asian	3.43 (0.749)	3.66 (0.142)	3.66 (0.139)
Non-Hispanic-Other	4.82 (1.067)	3.15 (0.149)	3.21 (0.149)
Total (unweighted)	579	16016	16595
$X^2 = 6.8$, $df = 4$, $p = 0.3272$			
Married	23.48 (1.918)	53.60 (0.457)	52.44 (0.449)
Not married	76.52 (1.918)	46.40 (0.457)	47.56 (0.449)

Table 7. Continued.

	Gay/lesbian/bi %	Straight %	Total %
Total (unweighted)	579	16016	16595
$X^2 = 223.9$, $df = 1$, $p = <.0001$			
Less than high school	8.21 (1.701)	9.89 (0.308)	9.82 (0.303)
High school graduates, no college	15.77 (2.030)	24.64 (0.418)	24.30 (0.410)
Some college or associate degree	32.71 (2.329)	31.21 (0.429)	31.27 (0.422)
Bachelor's degree or higher	43.32 (2.336)	34.26 (0.402)	34.61 (0.396)
Total (unweighted)	579	16016	16595
$X^2 = 36.8$, $df = 3$, $p = <.001$			
Own	41.26 (2.319)	65.29 (0.459)	64.36 (0.452)
Rent/Occupied without payment of rent	58.74 (2.319)	34.71 (0.459)	35.64 (0.452)
Total (unweighted)	579	16016	16595
$X^2 = 155.0$, $df = 1$, $p = <.0001$			
Less than USD34,999	42.71 (2.488)	33.30 (0.429)	33.66 (0.424)
USD35,000 – USD74,999	28.72 (2.160)	30.43 (0.427)	30.37 (0.419)
USD75,000 – USD149,999	19.03 (1.787)	24.47 (0.381)	24.26 (0.373)
USD150,000 and above	9.54 (1.158)	11.80 (0.269)	11.71 (0.262)
Total (unweighted)	579	16016	16595
$X^2 = 27.0$, $df = 3$, $p = <.0001$			
U.S. born	89.31 (1.463)	86.14 (0.301)	86.26 (0.294)
Foreign born	10.69 (1.463)	13.86 (0.301)	13.74 (0.294)
Total (unweighted)	579	16016	16595
$X^2 = 27.0$, $df = 3$, $p = <.0001$			

Source: 2018 CBAMS PUMS.

Table 8. Distribution of census knowledge index in 2020 census by sexual orientation (standard errors in parentheses).

	Gay/lesbian/bi	Straight	Total
High knowledge (8–11)	31.19	29.76	29.81
	(2.153)	(0.405)	(0.398)
Medium knowledge (4–7)	44.10	47.71	47.57
	(2.425)	(0.460)	(0.452)
Low knowledge (0–3)	24.71	22.54	22.62
	(2.293)	(0.397)	(0.392)
N (unweighted)	579	16016	16595
$X^2 = 3.4$, $df = 2$, $p < .1842$			

Source: 2018 CBAMS PUMS.

Table 9. Distribution of civic participation index in 2020 census by sexual orientation (standard errors in parentheses).

	Gay/lesbian/bi	Straight	Total
High civic engagement (7–10)	44.19	26.09	26.79
	(2.413)	(0.388)	(0.386)
Medium civic engagement (4–6)	31.56	34.66	34.54
	(2.282)	(0.434)	(0.427)
Low civic engagement (0–3)	24.25	39.25	38.669
	(2.179)	(0.450)	(0.441)
N (unweighted)	579	16016	16595
$X^2 = 112.7$, $df = 2$, $p < .0001$			

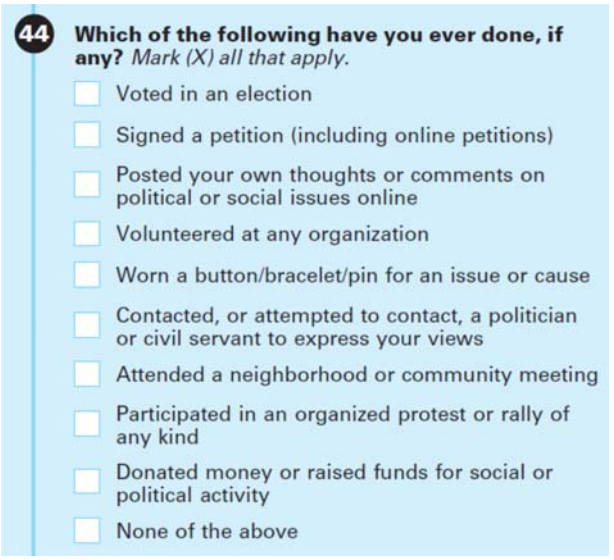
Source: 2018 CBAMS PUMS.

Table 10. Distribution of confidentiality index in 2020 census by sexual orientation (standard errors in parentheses).

	Gay/lesbian/bi	Straight	Total
No confidentiality concern	63.78	63.11	63.14
	(2.360)	(0.440)	(0.432)
	36.22	36.89	36.86
Confidentiality concern	(2.360)	(0.440)	(0.432
N (unweighted)	579	16016	16595
$X^2 = 0.1$, $df = 1$, $p < .7820$			

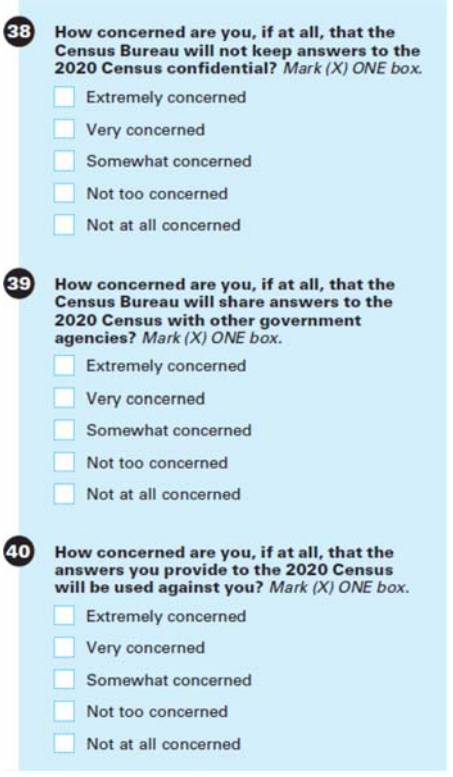
Source: 2018 CBAMS PUMS.

Question used to create Civic Engagement Index:



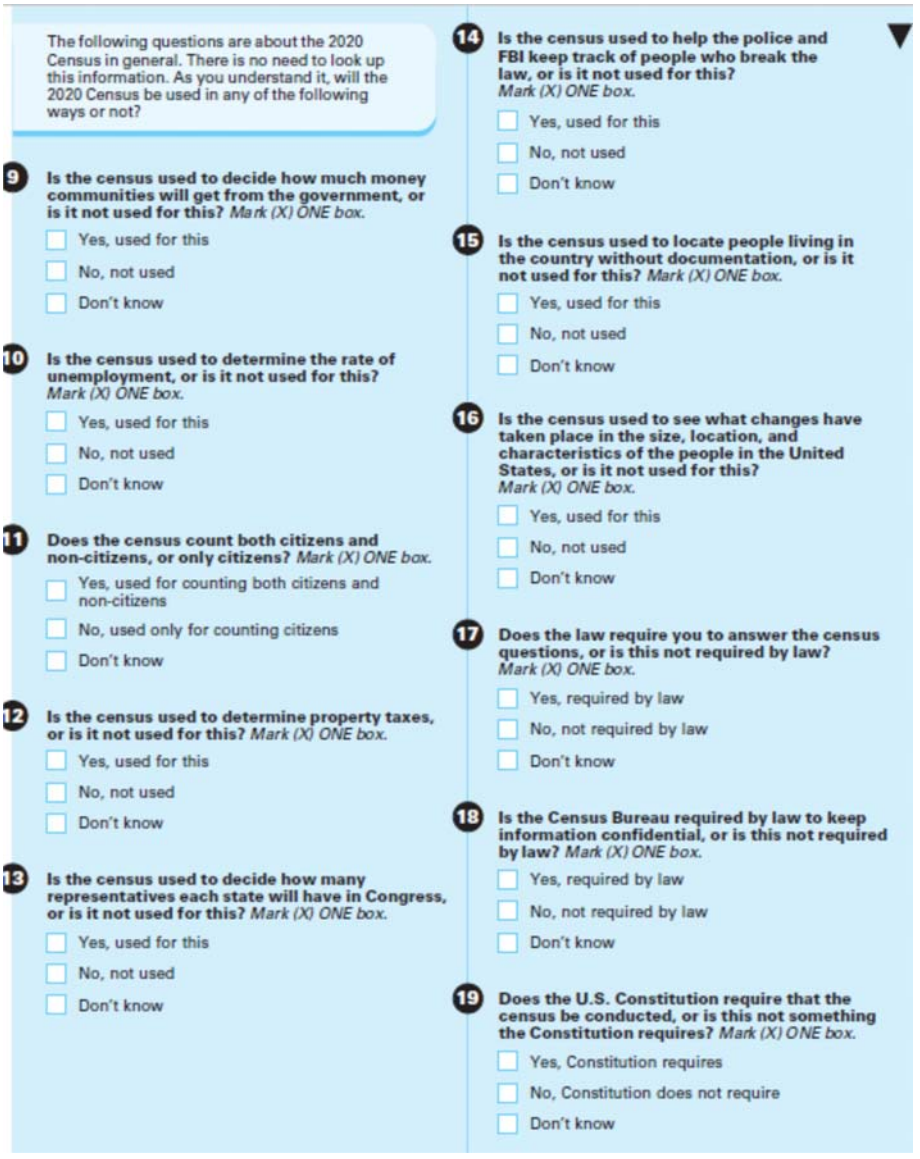
Source: 2020 CBAMS survey

Questions used to create the Concerns about Confidentiality Index:



Source: 2020 CBAMS survey

Questions used to create Knowledge Index:



The following questions are about the 2020 Census in general. There is no need to look up this information. As you understand it, will the 2020 Census be used in any of the following ways or not?

9 **Is the census used to decide how much money communities will get from the government, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

10 **Is the census used to determine the rate of unemployment, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

11 **Does the census count both citizens and non-citizens, or only citizens?** *Mark (X) ONE box.*
- ☐ Yes, used for counting both citizens and non-citizens
- ☐ No, used only for counting citizens
- ☐ Don't know

12 **Is the census used to determine property taxes, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

13 **Is the census used to decide how many representatives each state will have in Congress, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

14 **Is the census used to help the police and FBI keep track of people who break the law, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

15 **Is the census used to locate people living in the country without documentation, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

16 **Is the census used to see what changes have taken place in the size, location, and characteristics of the people in the United States, or is it not used for this?** *Mark (X) ONE box.*
- ☐ Yes, used for this
- ☐ No, not used
- ☐ Don't know

17 **Does the law require you to answer the census questions, or is this not required by law?** *Mark (X) ONE box.*
- ☐ Yes, required by law
- ☐ No, not required by law
- ☐ Don't know

18 **Is the Census Bureau required by law to keep information confidential, or is this not required by law?** *Mark (X) ONE box.*
- ☐ Yes, required by law
- ☐ No, not required by law
- ☐ Don't know

19 **Does the U.S. Constitution require that the census be conducted, or is this not something the Constitution requires?** *Mark (X) ONE box.*
- ☐ Yes, Constitution requires
- ☐ No, Constitution does not require
- ☐ Don't know

Source: 2020 CBAMS survey

7. References

AAPOR. 2016. *Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th edition. Available at: https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (accessed October 2019).

Ajzen, I. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50(2): 179–211. DOI: https://doi.org/10.1016/0749-5978(91)90020-T.

Allen, S. 2018. "These Five Court Cases Could Change the Future of LGBT Rights." *The Daily Beast*, July 7, 2018. Available at: https://www.thedailybeast.com/these-five-court-cases-could-change-the-future-of-lgbt-rights (accessed September 2019).

Atrostic, B.K., N. Bates, G. Burt, and A. Silberstein. 2001. "Nonresponse in US Government Household Surveys: Consistent Measures, Recent Trends, and New Insights." *Journal of Official Statistics* 17(2): 209–226.

Bates, N. and T.J. DeMaio. 2013. "Measuring Same-Sex Relationship." *Contexts* 12(1): 66–69. DOI: https://doi.org/10.1177/1536504213476251.

Bates, N. and M. Mulry. 2012. "Did the 2010 Census Social Marketing Campaign Shift Public Mindsets?" In proceedings of the annual conference of the American Association for Public Opinion Research, May 17, 2012. Orlando, FL: American Association for Public Opinion Research. Available at: http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2012/02_Bates-and-Mulry_A6-presentation.pdf (accessed September 2019).

Berg, N. and D. Lien. 2006. "Same-Sex Sexual Behaviour: US Frequency Estimates from Survey Data with Simultaneous Misreporting and Non-Response." *Applied Economics* 38(7): 757–769. DOI: https://doi.org/10.1080/00036840500427114.

Berg, N. and D. Lien. 2008. "Sexual Orientation and Self-Reported Lying." *Review of Economics of the Household* 7(1): 83–104. DOI: https://doi.org/10.1007/s11150-008-9038-1.

Berry, S.H. and P.P. Gunn. 2014. "Conducting research on vulnerable and stigmatized populations." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates. Cambridge: Cambridge University Press.

Beullens, K., G. Loosveldt, C. Vandenplas, and I. Stoop. 2018. "Response rates in The European Social Survey: Increasing, decreasing, or a matter of fieldwork efforts?" *Survey Methods: Insights from the Field*. DOI: https://doi.org/10.13094/SMIF-2018-00003.

Conron, K.J., S.K. Goldberg, and C.T. Halpern. 2018. "Socioeconomic Status of Sexual Minorities." Los Angeles: The Williams Institute. UCLA School of Law. Available at: https://williamsinstitute.law.ucla.edu/research/ses-sexual-minorities/ (accessed September 2018).

Dajani, A.N., A.D. Lauger, P.E. Singer, D. Kifer, J.P. Reiter, A.M. S, L. Garfinkel, S.A. Dahl, M. Graham, V. Karwas, H. Kim, P. Leclerc, I.M. Schmutte, W.N. Sexton, L. Vilhuber, and J.M. Abowd. 2017. "The modernization of statistical disclosure limitation at the U.S. Census Bureau." Washington, D.C.: U.S. Census Bureau. Available at: https://www2.census.gov/cac/sac/meetings/2017-09/statistical-disclosure-limitation.pdf?# (accessed September 2019).

Datta, A.R., T. Yan, D. Evans, S. Pedlow, B. Spencer, and R. Bautista. 2012. "2010 Census Integrated Communications Program Evaluation (CICPE)." 2010 Census Planning Memoranda Series. Suitland, MD: U.S. Census Bureau and NORC at the

University of Chicago. Available at: https://census.gov/content/dam/Census/library/publications/2012/dec/2010_cpex_167.pdf (accessed September 2019).

De Leeuw, E., J. Hox, and A. Luiten. 2018. "International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data." *Survey Insights: Methods from the Field*. DOI: https://doi.org/10.13094/SMIF-2018-00008.

DeMaio, T.J., N. Bates, and M. O'Connell. 2013. "Exploring Measurement Error Issues in Reporting Same-Sex Couples." *Public Opinion Quarterly* 77(S1): 145–158. DOI: https://doi.org/10.1093/poq/nfs066.

Dewaele, A., M. Caen, and A. Buysse. 2014. "Comparing Survey and Sampling Methods for Reaching Sexual Minority Individuals in Flanders." *Journal of Official Statistics* 30(2): 251–275. DOI: http://dx.doi.org/10.2478/jos-2014-0016.

Durkin, E. 2018. "US Begins Denying Visas for Same-Sex Domestic Partners of Diplomats." *The Guardian*, October 2, 2018. Available at: https://www.theguardian.com/world/2018/oct/02/us-visas-same-sex-domestic-partners-diplomats-un-staff-denied (accessed September 2019).

Dwork, C., F. McSherry, K. Nissim, and A. Smith. 2006. "Calibrating noise to sensitivity in private data analysis." In *Theory of Cryptography. TCC 2006. Lecture Notes in Computer Science*, edited by S. Halevi and T. Rabin, 265–284. Berlin, Heidelberg: Springer. DOI: http://dx.doi.org/10.1007/11681878_14.

Ellis, R., M. Virgile, J. Holzberg, D.V. Nelson, J. Edgar, P. Phipps, and R. Kaplan. 2017. "Assessing the Feasibility of Asking about Sexual Orientation and Gender Identity in the Current Population Survary: Results from Cognitive Interviews." Washington, D.C.: U.S. Bureau of Labor Statistics. Available at: https://www.bls.gov/osmr/research-papers/2017/html/st170210.htm (accessed September 2019).

Erdman, C. and N. Bates. 2017. "The Low Response Score (LRS) A Metric to Locate, Predict, and Manage Hard-to-Survey Populations." *Public Opinion Quarterly* 81(1): 144–156. DOI: https://doi.org/10.1093/poq/nfw040.

FBI. 2017. "2017 Hate Crime Statistics Released. Report Shows More Departments Reporting Hate Crime Statistics." *News*, November 13, 2018. Available at: https://www.fbi.gov/news/stories/2017-hate-crime-statistics-released-111318 (accessed September 2019).

Fredriksen-Goldsen, K.I. and H. Kim. 2014. "Count Me in: Response to Sexual Orientation Measures among Older Adults." *Research on Aging* 37(5): 464–480. DOI: https://doi.org/10.1177/0164027514542109.

Fredriksen-Goldsen, K.I., H. Kim, C. Shiu, J. Goldsen, and C.A. Emlet. 2015. "Successful Aging among LGBT Older Adults: Physical and Mental Health-Related Quality of Life by Age Group." *The Gerontologist* 55(1): 154–168. DOI: https://doi.org/10.1093/geront/gnu081.

Glasser, I., E. Hirsch, and A. Chan. 2014. "Reaching and enumerating homeless populations." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates, 180–200. Cambridge: Cambridge University Press.

Groves, R.M. and M.P. Couper. 1998. *Nonresponse in Household Surveys*. New York: Wiley.

Harkness, J., M. Stange, K.L. Cibelli, P. Mohler, and B. Pennell. 2014. "Surveying cultural and linguistic minorities." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates, 245–269. Cambridge: Cambridge University Press.

Hatzenbuehler, M.L., A.R. Flores, and G.J. Gates. 2017. "Social Attitudes Regarding Same-sex Marriage and LGBT Health Disparities: Results from a National Probability Sample." *Journal of Social Issues* 73(3): 508–528. DOI: https://doi.org/10.1111/josi.12229.

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." Washington, D.C.: U.S. Bureau of Labor Statistics. Available at: https://www.bls.gov/osmr/research-papers/2017/pdf/st170200.pdf (accessed September 2019).

Human Rights Campaign. 2018. "Violence Against the Transgender Community in 2018." *Human Rights Campaign*. Available at: http://www.hrc.org/resources/violence-against-the-transgender-community-in-2018/ (accessed May 2019).

Institute of Medicine (US) Committee on Lesbian, Gay, Bisexual, Transgender Health Issues and Research Gaps and Opportunities. 2011. *The Health of Lesbian, Gay, Bisexual and Transgender People: Building a Foundation for Better Understanding*. Washington D.C.: The National Academies Press. DOI: https://doi.org/10.17226/13128.

Kreuter, F. and K. Olson. 2013. "Paradata for Nonresponse Error Investigation." In *Improving Surveys with Paradata: Analytic Uses of Process Information*, edited by F. Kreuter, 11–42. New York: Wiley and Sons.

Law, T. 2019. "Trump Administration Asks Supreme Court to Permit Employment Discrimination Against Transgender Workers." *Time Magazine*, August 17, 2019. Available at: https://time.com/5654844/title-vii-trump-transgender-department-of-justice-supreme-court/ (accessed September 2019).

Lee, S., K.I. Fredriksen-Golden, and C. McClain. 2018. "Are Sexual Minorities Less Likely to Participate in Surveys? An Examination of Proxy Nonresponse Measures and Associated Biases with Sexual Orientation in the Population-based Health Survey." *Field Methods* 30(3): 208–224. DOI: https://doi.org/10.1177/1525822X18777736.

Lin, I. and N.C. Schaeffer. 1995. "Using Survey Participants to Estimate the Impact of Nonparticipation." *The Public Opinion Quarterly* 59(2): 236–258. DOI: https://doi.org/10.1086/269471.

Masci, D., A. Brown, and J. Kiley. 2017. "5 Facts about Same-Sex Marriage." Pew Research Center. Available at: https://www.pewresearch.org/fact-tank/2017/06/26/same-sex-marriage/ (accessed September 2019).

Massey, D.S. 2014. "Challenges to surveying immigrants." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates, 270–292. Cambridge: Cambridge University Press.

Masterpiece Cakeshop, Ltd. versus Colorado Civil Rights Commission. 2018. 584 U.S. Available at: https://www.supremecourt.gov/opinions/17pdf/16-111_j4el.pdf (accessed September 2019).

McGeeney, K., B. Kriz, S. Mullenax, L. Kail, G. Walejko, M. Vines, N. Bates, and Y. García. 2019. "2020 Census Barriers, Attitudes, and Motivators Study Survey Report." Suitland, MD: U.S. Census Bureau. Available at: https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/final-analysis/2020-report-cbams-study-survey.html (accessed May 2019).

Meiklejohn, J., J. Connor, and K. Kypri. 2012. "The Effect of Low Survey Response Rates on Estimates of Alcohol Consumption in a General Population Survey." *PloS One* 7(4): e35527. DOI: https://doi.org/10.1371/journal.pone.0035527.

Moohn, B. 2012. "2010 Census Integrated Communications Program Regional Partnership Assessment Report. 2010 Census Planning Memorandum Series, No. 217." Washington, D.C.: U.S. Census Bureau. Available at: https://www2.census.gov/programs-surveys/decennial/2010/program-management/5-review/cpex/2010-memo-217.pdf?# (accessed September 2019).

NHIS (National Health Interview Survey). 2015. "Sexual Orientation and Health among U.S. Adults: National Health Intervuew Survey." Hyatsville, MD: CDC/National Center for Health Statistics. Available at: https://www.cdc.gov/nchs/nhis/sexual_orientation/statistics.htm (accessed September 2019).

NORC. 2014. "Same-Sex Marriage and Gay Rights: A Shift in Americans' Attitudes." *The Associated Press-NORC Center for Public Affairs Research*. Available at: http://www.apnorc.org/projects/Pages/HTML%20Reports/same-sex-marriage-and-gay-rights-a-shift-in-americans-attitudes0305-8272.aspx (accessed May 2019).

OECD. 2019. "Society at a Glance 2019: OECD Social Indicators". Paris, France: OECD Publishing. DOI: https://doi.org/10.1787/soc_glance-2019-en.

Rossi, P.H., J.D. Wright, G.A. Fisher, and G. Willis. 1987. "The urban homeless: estimating composition and size." *Science* 235(4794): 1336–1341. DOI: https://doi.org/10.1126/science.2950592.

Saewyc, E.M. 2011. "Research on Adolescent Sexual Orientation: Development, Health Disparities, Stigma and Resilience." *Journal of Research on Adolescence* 21(1): 256–272. DOI: https://doi.org/10.1111/j.1532-7795.2010.00727.x.

Schwede, L., R. Terry, and J.H. Childs. 2014. "Ethnographic evaluations on coverage of hard-to-count minorities in the US decennial census." In *Hard-to-Survey Populations*, edited by R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter, and N. Bates, 293–315. Cambridge: Cambridge University Press.

Stoop, I.A.L. 2005. *The Hunt for the Last Respondent*. The Hague, the Netherlands: Social and Cultural Planning Office (SCP).

U.S. Census Bureau. 2013. "U.S. Census Bureau Statistical Quality Standards." Suitland, MD.: U.S. Census Bureau. Available at: https://www.census.gov/content/dam/Census/about/about-the-bureau/policies_and_notices/quality/statistical-quality-standards/Quality_Standards.pdf (accessed September 2019).

U.S. Census Bureau. 2016. National Advisory Committee on Race, Ethnic and Other Populations. "Administrative Records, Internet, and Hard to Count Population Working Group: Final Report." Suitland, MD.: U.S. Census Bureau. Available at:

https://www2.census.gov/cac/nac/reports/2016-07-admin_internet-wg-report.pdf (accessed September 2019).

Williams, D. and M. Brick. 2018. "Trends in US Face-to-Face Household Survey Nonresponse and Level of Effort." *Journal of Survey Statistics and Methodology* 6(2): 186–211. DOI: https://doi.org/10.1093/jssam/smx019.

Zauzmer, J. and T. McCoy. 2019. "D.C. Hate Crimes Nearly Double since 2016, with LGBTQ Community the Biggest Target." *Washington Post*, February 2, 2019. Available at: https://www.washingtonpost.com/local/social-issues/hate-crimes-in-dc-have-nearly-doubled-since-2016-according-to-city-data-and-incidents-based-on-sexual-orientation-top-the-list/2019/02/02/967aa46a-2651-11e9-90cd-dedb0c92dc17_story.html (accessed September 2019).

Received May 2019
Revised September 2019
Accepted November 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 731–752, http://dx.doi.org/10.2478/JOS-2019-0031

Test of a Hybrid Method of Sampling the LGBT Population: Web Respondent Driven Sampling with Seeds from a Probability Sample

Stuart Michaels[1], *Vicki Pineau*[1], *Becky Reimer*[1], *Nadarajasundaram Ganesh*[2], *and J. Michael Dennis*[3]

This article presents the results of a pilot feasibility study comparing two alternative recruitment approaches based on Respondent Driven Sampling using initial seeds selected from a US nationally representative panel, AmeriSpeak, to augment the number of lesbian, gay, bisexual, and transgender (LGBT) respondents to a short web survey on smoking, discrimination, and health. In the nomination condition after completing the survey both LGBT and non-LGBT seeds were invited to share the names and email address of up to four LGBT persons they knew. In the recruitment condition, seeds were given four unique PINs and links to the survey to distribute to LGBT persons. Both conditions were successful in producing new LGBT respondents. The recruitment condition was much more productive. LGBT seeds (and their recruits) were much connected to and willing to contact other LGBT people they knew to participate in a survey. Comparisons of characteristics and responses from the initial samples and the LGBT referrals as well as comparisons to LGB samples from a large national survey are presented. Results demonstrate the promise of this hybrid technique for increasing the number of LGBT respondents through referrals from an initial probability based sample.

Key words: LGBT population; respondent driven sampling; network sampling.

1. Introduction

Achieving large representative samples of the sexual and gender minority population is critical to advancing knowledge about these populations, their characteristics, the adversities they face, and potential negative social and health consequences of these adversities (IOM 2011). Before the 1990s, practically all research on lesbian, gay, bisexual, and/or transgender persons used convenience (i.e., nonprobability) samples (e.g., Kinsey et al. 1948, 1953; Bell and Weinberg 1978). With the advent of the HIV/AIDS epidemic, several large probability general population surveys that included questions about sexual orientation, which focused primarily on the sex of sexual partners, were fielded in the United States and in Europe (Laumann et al. 1994; Wellings et al. 1994; Spira et al. 1994). Since then, there has been a slow but steady increase in the number of

[1] NORC at the University of Chicago, 55 East Monroe Street, Chicago, Illinois 60603, U.S.A. Emails: michaels-stuart@norc.org, pineau-vicki@norc.org, and reimer-becky@norc.org
[2] NORC at the University of Chicago, 4350 East-West Highway, Bethesda, MD 20814, Maryland, U.S.A. Email: ganesh-nada@norc.org
[3] NORC at the University of Chicago – AmeriSpeak, 1250 Borregas Ave #135, Sunnyvale, California 94089, U.S.A. Email: dennis-michael@norc.org

national large scale probability surveys that include measures of sexual orientation and identity that produce relatively large samples of lesbian, gay, and bisexual persons (IOM 2011). Progress has been slower in the inclusion of measures of gender identity to population surveys, especially at the national level. The relative rarity of LGBT persons, estimated as approximately four percent of the US population (Gates 2014), makes it difficult and quite expensive for most researchers to obtain large representative samples of sexual and gender minority respondents.

This article describes the results of a pilot study of a hybrid sampling approach that combines an implementation of web-based Respondent-Driven Sampling (RDS) with seeds drawn from a probability-based panel of the US population that includes items on sexual and gender identity as part of panel members' profile. We were interested in developing a less expensive alternative to full probability sampling that could be used to generate large enough samples of sexual and gender minority persons to be able to address a wide range of research questions about these populations, including smaller subgroups within the LGBT populations, for example, specific age groups and/or racial and ethnic populations. The goal is to produce as representative a sample as possible, even when resources are limited, that will produce useful answers to important research questions, including the use of nonprobability sampling, what has been referred to as "fit for purpose" (Baker et al. 2013).

Our main goals were: 1) to test a method for generating an oversample of LGBT respondents to augment a sample selected from a national probability sample (specifically the NORC AmeriSpeak national panel) and 2) to assess the quality of the resulting oversample. Schonlau et al. (2014) describes the successful use of web-based RDS to recruit new internet panel members from among friends and acquaintances of an existing representative panel. As far as we are aware, this is the first publication describing an attempt to integrate techniques based on RDS using a probability sample as the starting point to recruit members of a hard-to-survey population (Tourangeau et al. 2014).

RDS is a form of snowball or chain referral sampling developed by Douglas Heckathorn (Heckathorn 1997; Heckathorn 2002; Salganik and Heckathorn 2004) to sample hidden populations that form a single interconnected network starting with a nonprobability, usually purposive sample of "seeds." The assumption that the population of interest is part of a single interconnected network insures that every member of the population has a non-zero probability of being included in the sample. Using network theory and methods, as well as strong assumptions, the goal of classical RDS is to be able to calculate estimates of characteristics of the population being sampled. Central to the method is to trace the chain of referrals, assess the number of potential referrals ("network size"), and generate long referral chains to reduce homophily effects between seeds and referrals. Heckathorn argues that when the assumptions are met, the realized sample is independent of the initial seeds. Most implementations of RDS have been in local areas involving face-to-face interactions, though web-based methods have been used (Wejnert and Heckathorn 2008; Bengtsson et al. 2012; Schonlau et al. 2014; Strömdahl et al. 2015). Each respondent, whether one of the seed cases or their referrals, who completes the survey is asked to refer a set number of other persons they know from among their acquaintances in the

population of interest. Respondents receive an incentive for their participation and for any of their referrals who successfully complete the survey. This process is repeated iteratively to produce multiple waves of referrals. RDS has been widely used in public health research internationally, in particular in the area of HIV/AIDS research to study hard-to-reach high risk populations such as sex workers, intravenous drug users, and men who have sex with men (McKnight et al. 2006; Bengtsson et al. 2012; Clark et al. 2014; Johnston et al. 2017; Schneider et al. 2017). Under a set of assumptions, and using information on network size and the linkage between respondents and referrals, RDS allows one to generate unbiased estimates of population parameters and generate standard errors for those estimates.

This project borrows the chain referral techniques of RDS for sampling hard-to-reach populations and applies them in a different context. Rather than starting with a set of seeds drawn from a local interconnected network of a hidden population, we apply these techniques to seeds that are drawn from a representative national panel. The initial sample of seeds have known non-zero probabilities of being selected from the target population since they were selected using probability-based methods from the full target population. In this sense, it is also related to network sampling (cf. Sirken 1998; Binson et al. 2007). Our primary focus in this research was not the statistical properties of our estimates, but the feasibility of augmenting the number of LGBT respondents in an efficient and cost-effective manner. In addition, we use this approach to compare the referral cases with seeds that are drawn from a true probability sample and another benchmark survey to make some assessment of how representative the resulting LGBT sample is.

2. Background

We conducted two pilot studies to develop and test this new method for sampling the LGBT population. The first pilot is described in more detail in an earlier publication (Pineau et al. 2017), but we give an overview here to provide background for the second pilot, the focus of the present work. These pilots combined aspects of probability sampling and RDS, seeking to take advantage of the strengths of both methodologies in conjunction with sample quality of the AmeriSpeak panel.

In both pilots, we selected a sample of LGBT and non-LGBT panelists from the probability-based AmeriSpeak Panel as seeds for the RDS referral methods. Though RDS does not require a probability-based sample as a starting point, use of a probability-based sample from AmeriSpeak Panel provided a less clustered and more randomly distributed starting point than the purposive sampling used in standard RDS studies. The AmeriSpeak Panel starting sample also provides a basis for evaluating the representativeness of the RDS referred sample on socio-demographic variables by comparing the RDS referred LGBT sample to the probability-based sample of LGBT seeds.

The seed sample of LGBT and non-LGBT panelists were invited to participate in an online survey that included a request to refer LGBT respondents they know to the survey. A brief online survey was developed for both pilots focused on smoking, as this is an area where prior research has found marked differences based on sexual orientation with important health consequences (Lee et al. 2009; Emory et al. 2016). In addition to

the substantive items about smoking, the survey asked about the number of LGBT persons who are part of their social network for RDS estimation purposes.

In traditional in-person RDS studies, seed sample respondents are given paper coupons to distribute to others in the population of interest so they can participate in the study. Our initial pilot used a variation of this RDS technique, which we refer to as the "nomination" approach. In this approach, seed respondents completed the substantive questionnaire first and then were asked to provide names and email addresses of up to three LGBT persons they knew to whom we would forward an invitation and a link to the survey. We selected this nomination approach at the time to avoid reprogramming our survey software, incorporating a manual process, where research staff added new survey records for the referral cases and sent survey invitations to them as needed.

The results of this first pilot indicated that web-based RDS can produce an LGBT oversample from seeds drawn from the AmeriSpeak Panel. We found that LGBT seeds (and referrals) were much more productive than their non-LGBT counterparts. Seeds and recruits knew many more LGBT friends and family that they said they could refer than the number of email addresses requested in the survey. However, overall the number of seeds (and referrals) providing referrals was disappointing. A total of 146 LGBT and 119 non-LGBT seeds provided 68 referrals (63 from LGBT seeds and five from non-LGBT seeds). None of the referrals nominated by non-LGBT seeds completed the survey, but 23 of the referrals from LGBT seeds did. After two more rounds of dwindling referrals and completes we were able to generate a total of 28 completed LGBT cases, which represented an increase of approximately 20% additional LGBT cases. Several seed sample respondents, who declined to provide email addresses for LGBT friends and family members in the referral section of the questionnaire, offered comments at the end of the survey indicating that they felt uncomfortable providing contact information for these people without first having the opportunity to get their permission. The comments from respondents in the initial pilot who were hesitant or unwilling to provide referrals led us to believe that allowing respondents to contact their friends directly and without sharing their email addresses with us could be more effective.

Based on these results, we decided to hone the design and carry out a second pilot, the focus of this article. We hypothesized that an alternative recruitment strategy closer to traditional RDS, in which respondents recruit others directly, could substantially increase the size of the final referral sample. Not only is this the standard in traditional RDS, it is also the technique used in web-based RDS surveys of men who have sex with men in Vietnam and Sweden (Bengtsson et al. 2012; Strömdahl et al. 2015).

In the second pilot, we used a split-sample design to compare the nomination approach from the first pilot as one experimental condition to an alternative condition where the panelist seeds were given the survey link to share with friends and family themselves. We also made improvements in the technology for prompting nonrespondents and we increased the number of RDS referrals from three to four. In RDS, the number of referrals is limited in order to create longer referral chains to reduce cumulatively the effects of homophily and fulfill the estimation assumptions. The exact number of referrals is not specified by the theory underlying RDS. It is common to use three, but many RDS studies use more. We settled on a small increase from three to four for the second pilot. The remainder of this article discusses methods and findings from this second pilot.

3. Methods

Data collection for the second pilot survey was conducted between August 18 and September 24, 2017. The source of the probability-based sample of seed cases was AmeriSpeak Panel, NORC's probability-based panel designed to be representative of the US household population. AmeriSpeak Panel was built using a rigorous sampling and recruitment methodology based on probability sampling techniques used by federally sponsored research. During the initial recruitment phase of the panel, randomly selected US households were sampled with known, non-zero probabilities of selection from the NORC National Sample Frame and then contacted by US mail, telephone, and field interviewers (face-to-face). The panel provides sample coverage of approximately 97% of the US household population. Those excluded from the sample include people with P.O. Box-only addresses, some addresses not listed in the USPS Delivery Sequence File, and some newly constructed dwellings.

3.1. Sample

When selecting panelists to act as seed cases for this study, our goal was to have roughly similar size samples of LGBT seeds and non-LGBT seeds. We used panelists' previous reports of LGBT status as a guide and then made a final determination of LGBT status by asking sexual orientation and gender identity (SOGI) questions within the pilot survey. The sample of adults drawn from the panel included a substantial oversample of adults who had previously identified as LGBT, with most of these having previously identified as LGBT in the AmeriSpeak Panel recruitment survey when joining the panel, and some having done so in another previous questionnaire fielded with AmeriSpeak panelists.

These previous self-reports varied in their recency and level of detail. Panel recruitment has occurred on an annual basis since 2014 (when the AmeriSpeak Panel was launched), and panelists are typically asked to complete several panel surveys per month. The SOGI items included in the recruitment survey were added gradually over the past few years starting with the sexual orientation question and then adding the two questions needed to determine gender identity. The questions were chosen with care, and are shown in Appendix A (Subsection 8.1). The sexual orientation question was taken from the National Health Interview Survey, which was developed after extensive cognitive testing by the questionnaire lab at the National Center of Health Statistics for use in general population surveys (Miller and Ryan 2011). The gender identity question selected was based on the recommendations from an experts panel (GenIUSS Group 2014) as used in the California Health Interview Survey.

Due to the changes to the SOGI questions used in the AmeriSpeak Panel recruitment survey over time, the sample included some panelists who had previously answered questions about both sexual orientation and gender identity, some who had answered questions about sexual orientation information but not gender identity, and some whose LGBT status was unknown. For example, it could have included panelists who had only been asked the sexual orientation item in 2014 without updates in the interim years, as well as panelists who had answered the SOGI questions as recently as the months leading up to the second pilot in 2017. Of the total 1,131 panelists sampled, 532 had previously identified as LGBT, 403 had previously identified as non-LGBT, and 196 had an unknown status. The latter were grouped with the non-LGBT seeds.

Panelists who had participated in the first pilot study were excluded from the sample. The seed sample was also limited to adults age 18 to 55. The decision to limit the age range was made based on other pending work at the time focusing on older LGBT adults. While this choice somewhat limits the generalizability of the experiment, it seemed prudent to avoid overburdening this rare population within the panel. In addition, even though AmeriSpeak is a multi-mode panel that incorporates both web and telephone interviewing, these pilot studies were conducted via web only, purely for budgetary reasons. Over 90% of the AmeriSpeak Panel responds via the web.

3.2. *Questionnaire*

Separate web surveys were created for the two conditions, which are described in Table 1. The structure of the survey was very similar between conditions, including substantive questions about smoking behavior, social media use, emotional well-being, and perceptions of discrimination, as well as demographic questions, and a request to nominate or recruit LGBT adults to take the survey. The smoking component focused on behaviors related to cigarettes, cigars, and cigarillos, as well as vaping and marijuana.

The demographics section included essentially the same SOGI questions used in previous rounds, shown in Table 1. The panelist seeds received the questions in the order listed here (sexual orientation identity, sex at birth, gender identity), but referral cases received them in a different order (Q14, Q15, Q13, i.e., sex at birth, gender, sexual orientation identity).

The structure of the survey differed for the panelist seeds compared to non-panelist referrals who were nominated or recruited, for two reasons. First, demographic

Table 1. Questionnaire.

Q13.
Which of the following best represents how you think of yourself?

1. [IF GENDER1=1 DISPLAY] Gay; [IF GENDER1<>1 DISPLAY] Lesbian or gay
2. [IF GENDER1=1 DISPLAY] Straight, that is, not gay; [IF GENDER1<>1 DISPLAY] Straight, that is, not lesbian or gay
3. Bisexual
4. Something else
5. I don't know the answer

Q14.
What sex were you assigned at birth, on your original birth certificate?

1. Male
2. Female

Q15.
How do you describe yourself today?

1. Male
2. Female
3. Transgender
4. I do not identify as male, female, or transgender

information needed to be collected for referrals, but these items did not need to be asked for panelist seeds since their demographic information was already known. Second, we chose to ask about LGBT status sooner in the survey for referrals than panelists for screening purposes. The questionnaire flows are shown in Appendix B (Subsection 8.2).

For panelist seeds, the questionnaire began with the substantive questions and then asked about sexual orientation and gender identity, to verify any information they had previously provided, and update it if necessary. The majority of panelists who completed the survey provided answers to the SOGI questions that confirmed their previously identified status as LGBT or non-LGBT. Of those who had provided this information in the past, 317 out of 347 (91%) across the two conditions confirmed their previous status, while 28 (8%) changed their identification from LGBT to non-LGBT, and two (less than 1%) changed their identification from non-LGBT to LGBT. A total of 63 panelists who had not answered SOGI questions in the past were sampled for the experiment. Among those, 55 identified as non-LGBT and eight identified as LGBT. In the analyses for this article, we use the responses to the SOGI questions in the survey as described below.

Panelists proceeded past the SOGI questions to an item assessing the size of individuals' LGBT networks. This question is presented in Appendix C (Subsection 8.3). It asked panelists to report the total number of LGBT adults within their social support and communication networks, including family members, friends, acquaintances, co-workers, classmates, and other people in one's community. The network size question is needed for RDS estimation, as persons who have larger networks have a greater chance of being referred to the study. Respondents who said they did not know any LGBT persons did not receive a request to refer others.

After the network size question, panelists were asked to nominate or recruit up to four of the LGBT adults they knew to take the survey. The survey text explained that if others completed the survey, the panelist would receive an incentive themselves, and those they referred to the survey would also receive incentives upon completion. The methods for this request differed between the two conditions and are described below. The last question for panelist seeds was an open-ended field seeking comments about the survey. Panelists received AmeriSpeak points after completing the survey.

When non-panelist referrals entered the survey, the first priority was to assess their LGBT status, to confirm their eligibility. The first few questions asked about basic demographics followed by the SOGI items. Those who identified as non-LGBT were screened out of the survey. Those who identified as LGBT proceeded to the substantive questions, then the network size question, referral request, a screen that collected contact information for an incentive to be mailed to them, and the survey comments.

3.3. Experimental Conditions

Sampled panelists were randomly assigned to one of the two experimental conditions: a *nomination condition* (used in the first pilot) and a *recruitment condition*. The conditions differed only in the methods by which panelists were asked to refer LGBT friends and family members at the end of the survey.

In the nomination condition, panelists were asked to nominate LGBT adults to participate by providing their names and email addresses. In the recruitment condition,

they were not asked to provide any contact information and were instead asked to personally invite LGBT adults by directly reaching out to these individuals themselves and sharing a survey link and unique PIN. In both conditions, the referred LGBT non-panelists who completed the survey were also asked to enlist other LGBT individuals to take the survey to generate multiple waves of referrals. The LGBT referrals were invited to refer others using the same technique (nomination or recruitment) as had been used to generate their participation.

A total of 1,131 AmeriSpeak panelists received an invitation to participate in the survey, 565 in the nomination condition and 566 in the recruitment condition. Based on panelist participation rates in previous pilot work, the goal was to have approximately 100 panelist seed cases complete the survey in each of four cells: LGBT nomination, non-LGBT nomination, LGBT recruitment, and non-LGBT recruitment.

3.3.1. Nomination Condition

In the nomination condition, the referral item included information about the study's methods and purpose, the importance of the research, confidentiality, and survey incentives. It requested that respondents nominate up to four LGBT people they knew to take the survey by providing their names and email addresses. Respondents who had indicated that their LGBT network size was zero skipped this question. Those who entered information to nominate others to the survey were asked to confirm that the information was correct on the following screen before email invitations were sent.

3.3.2. Recruitment Condition

In the recruitment condition, the referral item contained the same wording as the nomination condition about the study background, importance, confidentiality, and incentives. Again, respondents who had indicated that their LGBT network size was zero skipped this question. The item, which is shown in Appendix D (Subsection 8.4), requested that respondents invite up to four LGBT people they knew to take the survey by personally contacting them, rather than by providing their names and email addresses. It provided the survey link and four unique PINs for panelists to pass along to their referrals. It also had a sample survey invitation that they could copy and paste to use along with the unique PINs if desired. In addition, it allowed them to postpone sending invitations until a later, more convenient time by offering the opportunity to receive all of this information in an email for later use.

3.4. Incentives and Nominations/Referrals

Each respondent who completed the survey (in either condition) received the equivalent of USD 5.00 for doing so. Panelist seeds received this incentive in the form of 5,000 points, while non-panelist referrals received a USD 5.00 Amazon gift card. Non-panelists were asked to provide their mailing address to receive the gift cards. Those respondents who successfully referred others to the survey received the same incentive again for each referral who completed the survey.

3.5. Reminders

Multiple types of participation reminders were sent for this study. In both conditions, panelists who had not yet completed the survey received reminders twice a week. Panelists

who had completed the survey themselves, but whose referral cases had not yet completed received reminders once a week, to encourage them to remind their friends and family members about the survey. In the nomination condition, we were able to send these types of reminders to referral cases as well, because their email addresses were stored in the questionnaire data. For the recruitment condition, email addresses were not requested, so we were not able to send any reminders to referral cases.

3.6. LGBT Questions and Definition

The LGBT definition in this pilot (for both panelists and referral cases) used three questions: first, a question on sexual orientation identity developed for and used on the National Health Interview Survey and the two-step sex/gender identity questions used in the California Health Interview Survey. The exact wording and response categories for these questions is shown above in Subsection 3.2. For a respondent to be considered LGBT, they had to meet one of the following requirements:

- Identify as gay, lesbian, or bisexual at the sexual orientation question
- Identify as transgender on the gender identity question, or as a different sex on the gender identity question than on the sex assigned at birth question

Those who identified as "straight" and reported a gender identity matching their sex assigned at birth were categorized as non-LGBT. Those who reported that their sexual orientation was "something else" or that they did not know how to answer the question, or those who indicated that they did not identify as male, female, or transgender for the gender identity question were also considered non-LGBT.

This relatively restrictive definition of LGBT was made to facilitate the programming of the web survey, which had to determine whether referral cases could proceed to the substantive survey based on their responses to these questions. Mechanisms to allow persons who identify as for example "queer" or "gender non-binary" or some other designations that one might want to include as LGBT could be considered in future applications of this method.

Two panelists declined to answer at least one of the SOGI questions, meaning they had incomplete information, and were categorized as non-LGBT. A total of 11 panelists selected "something else" or "I don't know how to answer this question" for the sexual orientation item and were categorized as non-LGBT. Two of these eleven had a single successful referral to the survey, while the remaining nine had no referrals. Referral cases who fell into this category were screened out of the survey. A total of 19 referrals screened out of the survey as non-LGBT. This included five who identified as cisgender/straight and 14 who selected one of the other categories on the sexual orientation item and/or the gender identity item.

4. Results

As in the initial pilot, this pilot was successful in generating an oversample of LGBT respondents via outreach from probability-based panel members and their referrals. A total of 410 panelist seeds (182 LGBT and 228 non-LGBT) and 107 LGBT referrals completed

the survey. Table 2 summarizes the overall completed seed and referral cases under the two conditions based on whether the panelist seeds had identified themselves as LGBT or not during the survey.

Table 2 shows that the recruitment condition was nearly twice as productive as the nomination condition in yielding completed LGBT referral cases, with a total of 70 completed referral interviews compared to 37 in the nomination condition. This difference is also clear in the visual shown in Figure 1, which displays small squares for each panelist seed case that generated at least one referral, and small circles for each referral interview. Figure 1 details the number of referrals at each stage of the referral process: wave 1 are the referrals from the seeds drawn from the AmeriSpeak Panel, Wave 2 are the referrals made by the completed LGBT referrals in Wave 1, and so on. There are far more referral cases in the bottom half of the figure, representing the recruitment condition, than the top half, which represents the nomination condition.

Figure 1 shows the differences in the two conditions over the iterations of the RDS referral process for the productive LGBT and non-LGBT seeds. In the nomination condition, which required respondents provide a name and email address for referrals, 14 of the 91 LGBT seeds (15%) generated a total of 28 referral cases, whereas five of 117 non-LGBT seeds (4%) generated nine referral cases. Overall, the nomination condition resulted in 37 additional LGBT cases beyond the initial 91, a 41% increase in LGBT cases. In the recruitment condition where respondents received survey links and PINs to distribute to LGBT friends and family, again 14 of the 91 LGBT seeds (15%) generated referrals, but these referrals were more productive at each wave than those in the nomination condition. Cumulatively, the recruitment condition resulted in a total of 49 LGBT referrals (compared to only 28 in the nomination condition). In the non-LGBT group in the recruitment condition, eleven of the 111 seeds (10%) generated a total of 21

Table 2. Comparison of productivity of LGBT and non-LGBT seeds by condition.

	Nomination condition			Recruitment condition		
	LGBT sample	Non-LGBT sample	Total	LGBT sample	Non-LGBT sample	Total
Total sample released	268	297	565	264	302	566
Completed panelist seed interviews	91	117	208	91	111	202
Panelists with at least one completed referral	14	5	19	14	11	25
Total completed LGBT referral interviews	28	9	37	49	21	70
Percentage of additional completes via referrals	31%	8%	18%	54%	19%	35%

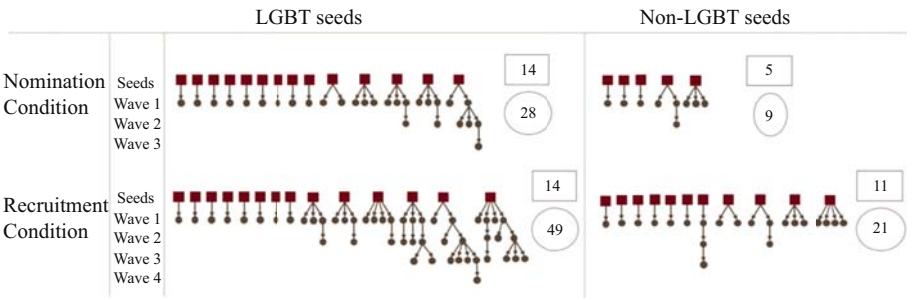


Fig. 1. Comparison of LGBT referrals by wave and condition for productive LGBT or non-LGBT seeds.

referral cases. Overall, the additional 70 LGBT referral completions in the recruitment condition represents a 77% increase over the initial group of 91 LGBT seeds. The total number of LGBT referral cases in the recruitment condition was higher than the nomination condition for several reasons: more non-LGBT seed cases referred at least one LGBT adult, seeds tended to refer more LGBT adults, and more of the referral cases, in turn, referred additional respondents.

In both conditions, LGBT seeds were more productive than non-LGBT seeds. As seen in Table 2 and Figure 1, independent of condition, the 182 LGBT seeds resulted in 77 LGBT referrals (28 from the nomination condition and 49 from the recruitment condition), whereas the 228 non-LGBT seeds only resulted in 30 LGBT referrals (nine from the nomination condition and 21 from the recruitment condition). In part, this may be because LGBT persons know and interact with a larger number of persons who are also sexual or gender minorities. In fact, we find strong evidence of this based on responses to the network size question (see Appendix C for wording of question).

Table 3 presents the distribution of the responses to the LGBT network size question for the LGBT and non-LGBT panelist seeds, collapsed across the nomination and recruitment conditions. Only 3% of the LGBT seeds report having no other LGBT people in their close network compared to almost one quarter of the non-LGBT seeds. Almost 60% of the LGBT seeds have more than ten LGBT people in their network compared to only 20% of the non-LGBT seeds.

Table 4 compares the actual numbers of successful LGBT referrals for the LGBT and non-LGBT panelist seeds across the two conditions. Even though almost all the LGBT

Table 3. LGBT network size for LGBT and non-LGBT seeds.

Number of LGBT adults known	LGBT seeds	Percent	Non-LGBT seeds	Percent
0	5	3%	55	24%
1	5	3%	9	4%
2	7	4%	28	12%
3	3	2%	18	8%
4–10	58	32%	72	31%
11+	104	57%	45	20%
Total	182		228	

Table 4. Successful referrals by LGBT and non-LGBT seeds.

Number of LGBT referred	LGBT seeds	Percent	Non-LGBT seeds	Percent
0	154	85%	212	93%
1	17	9%	10	4%
2	3	2%	3	1%
3	6	3%	2	1%
4	2	1%	1	<1%
Total	182		228	

seeds know other LGBT persons, 85% generated no referrals and only 15% successfully recruited at least one other LGBT person to the survey. Among the non-LGBT seeds, about half as many, 7%, recruited at least one LGBT respondent. LGBT seeds were also more likely to refer two or more other LGBT respondents.

4.1. Comparison of Productive and Non-Productive Seeds

One might ask whether there are demographic differences that account for whether some panelist seeds were productive while others were not. The seeds for this pilot were drawn from the AmeriSpeak Panel, which has a demographic profile of panel members. To address this question, we conducted a binomial logistic regression analysis on whether or not a seed case produced at least one successful LGBT referral by key socio-demographics while controlling for condition and whether or not the seed was LGBT. Table 5 presents the results of the regression. Based on the results from this model, the only variable that was significant at the .05 level was whether the seed was LGBT themselves. Otherwise, none of the demographic variables were significant predictors of seeds' successful recruitment of at least one LGBT referral. It is worth noting that the fact that condition is not significant ($p = .213$) may seem surprising, given the large differences in total LGBT

Table 5. Binomial logistic regression on seeds' referral success by demographics and condition (N = 366).

Dependent variables	B	S.E.	Sig.
Male (*ref. female*)	−.038	.342	.912
African American, non-Hispanic (*ref. white/missing*)	.174	.553	.753
Hispanic (ref. *white/missing*)	.506	.395	.200
Other race, non-Hispanic (ref. *white/missing*)	−.846	1.063	.426
Some college assoc. degree (r*ef. high school or less*)	−.129	.859	.881
Graduate degree (r*ef. high school or less*)	−.564	.820	.491
Post-graduate degree (r*ef. high school or less)*	.323	.844	.702
Referrer/seed is LGBT (r*ef. not LGBT*)	.902	.344	.009
Referral type is nomination (r*ef. recruitment*)	−.416	.333	.213
Household size	−.048	.117	.683
Income	−.021	.042	.616
Age	−.007	.017	.691
Constant	−1.823	1.129	.106

referrals by condition described above. However, this regression only looks at seeds and whether they produced at least one LGBT referral or not. As we saw in Figure 1, the overall success of the recruitment condition compared to the nomination condition was due to the fact that seeds in the recruitment condition were more likely to produce multiple referrals and their referrals were also more likely to be productive.

4.2. *Comparing the Resulting LGBT Referral Sample with Probability Sample Benchmarks*

The LGBT seeds were drawn from a national probability sample. Since the seeds can be considered a representative sample of all persons from the US population 18 to 55 years old, they are presumed to also be representative of the LGBT population from the US population 18 to 55 years old. The LGBT referrals are a nonprobability sample generated by an RDS snowball chain-referral sampling process. The ultimate goal is to be able to combine the seed and referral LGBT respondents into a larger, single analytic sample to study the LGBT population and/or compare it to the non-LGBT population. What biases might be introduced by simply combining all the LGBT cases? To investigate this question, we compare the probability and nonprobability components of the LGBT sample to each other and to an independent probability sample of the LGB population in the United States. Unfortunately, there is no national probability survey that asks about gender identity that would allow us to identify LGBT cases. In our sample there were a total of 14 transgender cases: three were seeds in the nomination condition; six referral cases in the nomination condition; and five referral cases in the recruitment condition. There were no transgender seeds in the recruitment condition.

Table 6 compares both the probability sampled Non-LGBT and LGBT seeds and the LGBT referrals from the RDS process from this pilot on key demographic variables. In addition, we provide estimates from the LGB and non-LGB adults, 18 to 55 years old, from the large representative National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III). The NESARC-III is a household-based survey of US adults carried out in 2012–2013 by the National Institute on Alcohol Abuse and Alcoholism. It includes a sexual orientation question that allows one to distinguish lesbian, gay, and bisexual respondents, but no gender identity question. Note that sexual orientation in NESARC-III is based on responses to the question: "Which of the categories on the card best describes you? 1) Heterosexual (straight), 2) Gay or lesbian, 3) Bisexual, 4) Not sure." The wording differs somewhat from the NHIS question used in AmeriSpeak found in Appendix A, which may affect estimates. The percentages presented here are based on weights that take into account the complex sample design. Full documentation of NESARC-III is available on line (National Institutes of Health 2013).

The percentages for the Non-LGBT and LGBT seeds (drawn from the AmeriSpeak Panel) were weighted based on their probabilities of inclusion and adjusting for nonresponse, exclusion of non-web panelists and other non-sampling error via post-stratification to socio-demographic population benchmarks. No weights are applied to the LGBT referrals. The percentages in Column D (All LGBT) takes into consideration the adjustments for the seeds described above and no adjustment applied to the referrals by adding the expected number of seeds based on their weighted percentages to the raw count

Table 6. Comparison of distributions of deeds, LGBT referrals on age, gender, race/ethnicity, and education and distributions of LGB and non-LGB respondents from the 2012–2013 NESARC-III general population survey.

	A	B	C	D	E	F
	Non-LGBT seeds	LGBT seeds	LGBT referrals	All LGBT	Gen pop LGB	Gen pop Non-LGB
Age		C	E			
18–34	50%	55%	77%	63%	61%	44%
35–49	33%	31%	18%	26%	27%	39%
50–64	16%	14%	4%	10%	13%	17%
65+	0%	0%	2%	1%		
Gender						
Male	50%	34%	46%	38%	41%	49%
Female	50%	61%	42%	54%	59%	51%
Transgender	0%	5%	7%	6%		
None of the above	0%	1%	5%	2%		
Race/ethnicity			E	E		
NH white	58%	68%	64%	66%	63%	61%
NH black	13%	10%	6%	8%	16%	13%
Hispanic	20%	16%	24%	19%	16%	18%
NH other	10%	7%	7%	7%	5%	8%
Education		C,E	E	E		
Less than HS	8%	6%	7%	6%	11%	12%
HS or equivalent	30%	31%	7%	23%	24%	25%
More than HS	62%	63%	86%	71%	65%	63%
Total Raw N	228	182	107	289	991	24,406

Notes: General Population 1 estimates from 18–55 year olds from NESARC-III. Superscripts indicate that the Chi-Square was significant at the .05 level for a variable when compared to the indicated column. E.g., Column B, LGBT Seeds differ significantly from Column C LGBT Referrals. Distribution of C (LGBT Referrals) differs from Column E (LGB in NESARC-III). The absence of a superscript means the Chi-Square was not significant at the .05 level. For example, the Non-LGBT Seeds (Column A) never differed from Column F. The distribution of the LGBT Seeds differed significantly from the distribution of the LGB cases in the general population NESARC-III sample for Education ($p = .042$), but did not differ for Age and Race/Ethnicity.

of LGBT referrals. A series of Chi-Square tests based on estimated counts were computed comparing distributions of age, race/ethnicity, and education of seeds with the corresponding weighted distributions obtained in the NESARC-III sample, as well as comparison of the LGBT seeds with the LGBT referrals. Chi-Square comparisons that were significant at the .05 level are indicated by superscripts with the column from which any column differs for the demographic variable. For example, the "C" under column B (LGBT Seeds) for Age indicates that the age distribution of LGBT Seeds and LGBT

referrals did differ significantly. The LGBT referrals tend to be younger than the seeds. The absence of an "E" under Column B for age, gender, and race/ethnicity indicates that distribution of the LGBT seeds for those variables did not differ from the distribution of the LGB sample in NESARC. On the other hand, the LGBT seeds did differ from the NESARC-III sample on education ($p = .042$), with fewer LGBT seeds who had not finished high school than in the NESARC-III sample and more high school graduates.

Overall, the LGBT and non-LGBT (seed) respondents drawn from the AmeriSpeak Panel are quite similar. There are more women among the LGBT than among the non-LGBT panelists. There is a similar tendency in the NESARC-III data. This pattern has also been noted in other recent US surveys, such as the National Survey of Family Growth, in which there are more women than men who report being LGB, especially at younger ages (Chandra et al. 2011). In addition, there is a higher proportion of whites (68%) among the LGBT seeds than among the non-LGBT seeds (58%). There is a similar tendency in the NESARC-III data, though not quite as marked. Chi-Square tests comparing the comparable parts of the seeds samples and the general population samples were not significant at the .05 level, implying that one cannot reject the null hypothesis that the samples were drawn from the same distributions.

The sample sizes are quite small, but it is important to note that there does appear to be a greater degree of difference between the probability sample of LGBT panelists and the LGBT referrals than between the LGBT and non-LGBT components of the probability sample. The LGBT referrals are younger, more evenly split between men and women, and more likely to have post-high school education. The LGBT referrals (and the combination of the LGBT seeds and referrals) differ from the national sample of LGB cases from NESARC-III on race/ethnicity and education. The former appears to be due to the smaller proportion of non-Hispanic Black LGBT referrals, which contributes to the significant difference in race/ethnicity when compared to the larger nationally representative sample of LGB cases in NESARC-III. The tendency to refer more educated LGBT cases also produces a skewed distribution on education for both the referrals and the combination of referrals and seed LGBT cases when compared to a national sample of LGB cases.

5. Summary and Discussion

Building on an earlier pilot, we carried out a pilot test of two versions of a web-based RDS protocol using seeds drawn from a probability sample panel of the US population. Using samples of identified LGBT and non-LGBT members of the AmeriSpeak Panel, we tested two types of respondent-driven recruiting: a nomination process where respondents were asked to provide a name and email address of people they knew who were LGBT and a referral process that asked respondents to distribute up to four unique PINs to LGBT friends. We found that while both techniques worked, the second approach, which we have called the recruitment condition, was much more effective in generating new LGBT cases. The recruitment condition, starting from the 202 panelist seed sample, split between LGBT and non-LGBT seeds, generated 70 additional LGBT respondents, a 35% increase. Most of that increase, 49 referrals, was generated by the 91 seeds who were LGBT, which generated an approximately 53% increase in LGBT cases. This represents a lower bound estimate due to the limited field period for the pilot, which only achieved a maximum of

four waves. In a longer field period, some of the respondents may have referred other LGBT respondents.

We found strong evidence that a referral process that allows seeds and referrals to carry out the recruitment process on their own has distinct advantages over a nomination process that requires respondents to share their acquaintances' contact information. We surmise that this was largely due to a reluctance on the part of respondents to give out the names and emails of persons they knew to a third party without the accord of those acquaintances, even when given assurances about confidentiality and when there was some monetary incentive for them and for their friends. This recruitment approach is closer to typical RDS and takes fuller advantage of the respondent driven aspect of this sampling technique, which enlists respondents in identifying and recruiting other members of a hidden population of interest (Heckathorn 1997).

The LGBT respondents were much more connected to and willing to contact other LGBT people they knew to participate in a survey. However, it is worth noting that, especially in the recruitment condition, non-LGBT seeds were also successful in recruiting LGBT referrals. In a probability panel that is representative of the general population, the number of LGBT panel members will be small, on the order of four percent, and if one were seeking to scale up the pilot and achieve a larger sample of LGBT cases, one may need to draw both LGBT and non-LGBT members to achieve the number of LGBT referrals sought. This could include a much larger number of non-LGBT than LGBT seeds.

While demonstrating that this method can be used to augment the number of LGBT cases in a resulting analytic sample, the comparisons between the LGBT seeds and the LGBT cases referred and the comparison of the resulting sample with an independent benchmark probability sample of LGB cases did show demographic bias in referrals who tended to be younger and more educated than the seeds. This may be due to younger and more educated LGBT persons being more open and comfortable with their sexual and/or gender minority status, making them more susceptible to be "out" to their friends and acquaintances and more likely to be willing to participate in a survey focused on them. This suggests caution regarding the representativeness of the resulting sample of LGBT persons. While the sample of seeds is a strict probability sample, the final LGBT sample is a combined probability and nonprobability sample. However, we would argue that it has definite advantages compared to a purely nonprobability convenience sample. In addition, the combination of a probability and nonprobability component allows one to measure and potentially adjust for differences in the latter.

6. Limitations

There are a number of limitations to this pilot study. For reasons of cost and efficiency, we restricted this implementation to a web-based mode only. We restricted the sample to panelists who prefer to complete surveys on the internet, whereas in usual AmeriSpeak surveys, responding by telephone is also available and is used by just under 10% of the 18–55 year olds in the panel. Though weights were applied to the AmeriSpeak seed sample to account for exclusion of the non-web panelists in the seed sample, the differences observed in the comparisons with the household-based NESARC-III data (e.g., age, race, and education) may be due in part to this design choice. Additionally, the characteristics of the referred samples may also be impacted by the exclusion of non-web panelists.

One of the difficulties in all RDS implementations due to its use of incentives to encourage recruitment is that it also produces an incentive for some respondents to attempt to "game the system" by either attempting to self-nominate or encouraging others to pretend to be eligible to do the survey. These problems are exacerbated in web-based surveys. At the same time, there are techniques to prevent such gaming (e.g., detecting and preventing multiple entries on the same device).

Ultimately, the final sample generated with our hybrid method is a combined probability-based and nonprobability-based sample. There are various methods for weighting and estimation that can be used, such as RDS estimation, propensity weighting approaches, super-population model approaches, and small area estimation approaches. As we continue to test and implement this new hybrid method for sampling rare populations, our intent is to concurrently explore and develop the most appropriate weighting and estimation methods.

7. Conclusion

While at one time it was widely believed that it would be impossible to even ask randomly selected respondents about their sexual and gender identity, the existence of a growing number of large representative population surveys that achieve that goal on a regular basis, as well as methods and techniques such as large scale probability panels and innovative nonprobability methods such as RDS prompted us to explore mixing the two in the pilot project presented here. We believe that the results are encouraging and show that this is a promising avenue for producing relatively cost-effective larger samples of LGBT respondents using a hybrid method that mixes probability and nonprobability sampling. More research and experience is needed to improve the methods used here and to better understand the properties of the resulting sample. The results from this pilot clearly indicate that a recruitment approach is preferable to a nomination approach. In future implementations, we would improve the protocol for identifying LGBT respondents to allow persons who are sexual and/or gender minority who use rarer terms to describe their sexual or gender identity, such as queer or gender non-binary to be considered eligible. Future research could also explore methods to increase the effectiveness of the RDS referrals by better explanations of the purpose and importance of recruitment, as well as testing higher incentives while monitoring any signs of increases in gaming the system. In addition, further research is needed on methods for adjusting and combining the nonprobability referral cases with the probability portion of the sample. The approach described here does not necessitate an existing probability panel. It could be implemented with any probability sample by adding an RDS recruitment component, for example, inviting and incentivizing respondents to recruit LGBT acquaintances to participate.

8. Appendix

8.1. Appendix A: Sexual Orientation and Gender Identity Questions

The sexual orientation and gender identity questions included in the AmeriSpeak registration and recruitment survey have been added over time. Below the variable name

for each item below, we have noted the years in which these items were included in the survey.

GENDER
[Asked in 2014 thru early 2017]
Are you . . ..

1. Male
2. Female

GENDER1
[Added in 2017]
What sex were you assigned at birth, on your original birth certificate?

1. Male
2. Female

GENDER2
[Added in 2017]
How do you describe yourself?

1. Male
2. Female
3. Transgender
4. Do not identify as male, female, or transgender

[IF GENDER2 = 3]
GENDER3
[Added in 2017]

Would you say you are?

1. Transgender, male to female
2. Transgender, female to male
3. Transgender, gender non-conforming
4. Other (please specify)----------

[IF (GENDER1= 1 AND GENDER2= 2) OR (GENDER1= 2 AND GENDER2= 1)]
GENDER4
[Added in 2017]

Just to confirm, you were assigned <GENDER1> at birth and now describe yourself as <GENDER2>. Is that correct?

1. Yes
2. No

LGBT
[Added in 2016]
This next question is about sexual orientation. Which of the following best represents how you think of yourself?
1. [IF GENDER1= 1 DISPLAY] Gay; [IF GENDER1<>1 DISPLAY] Lesbian or gay
2. [IF GENDER1= 1 DISPLAY] Straight, that is, not gay; [IF GENDER1<>1 DISPLAY] Straight, that is, not lesbian or gay
3. Bisexual
4. Something else
5. I don't know the answer

8.2. Appendix B: Questionnaire Flow for Panelists and Referrals

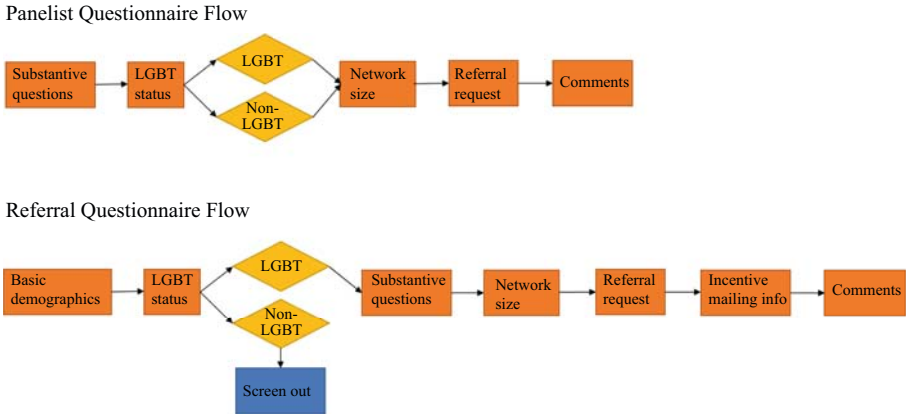


Fig. 2. Panelist questionnaire flow.

8.3. Appendix C: Network Size Question

We are interested in learning about people's social support and communication networks. By social support and communication network, we mean the set of other people that each person knows, including family members and friends, with whom they communicate on a fairly regular basis and with whom they exchange information and ideas. For this study, we're particularly interested in the number of people 18 and older who are lesbian, gay, bisexual, or transgender (LGBT) in your social network. Regardless of how you describe yourself, we're interested in the number of LGBT people you know.

If you had to estimate, about how many people 18 or older do you know and communicate with who are lesbian, gay, bisexual, or transgender (LGBT)?

Please include:

- family members
- friends
- acquaintances
- co-workers
- classmates
- other people you know in your community

Please do NOT include:

- people who live outside of the U.S.
- people who you haven't talked to or contacted (via email, text, Facebook, etc.) within the past two years
- people who you no longer have a way of contacting if you wanted to get in touch with them

If you aren't sure of the total number of LGBT people you know, please provide your best guess.

8.4. Appendix D: Wording of Referral Request in Recruitment Condition

We're studying LGBT health and we need your help!

One of the goals of this important health study is to contact and include enough lesbian, gay, bisexual or transgender (LGBT) people to be able to better represent their experiences and compare their experiences with their non-LGBT peers' experiences.

We'd like you to invite [[network= 1] the LGBT person you know who is/[network= 2-3] any of the {NUM} LGBT people you know who are /[network>=4] up to four of the {NUM} LGBT people you know who are] 18 years old or older and who live in the U.S. to take this survey. Each LGBT person you invite to take this survey will receive a USD 5 Amazon gift card for their participation, if they have not already taken the survey. For each person you invite who has not already been invited to take the survey and who successfully completes the survey, you will receive [[AMERISPEAK PANELIST: an additional 5,000 AmeriSpeak points]/[REFERRAL: another USD 5 Amazon gift card]] in addition to the one you will receive for finishing this survey.

In order to receive the maximum of [[AMERISPEAK PANELIST: 20,000 additional AmeriSpeak points]/[REFERRAL: four additional USD 5 Amazon gift cards]], you should invite up to four different LGBT people, and invite only those people who you are sure will take and complete the survey.

The information that the people you invite provide is strictly confidential and will be kept private. We will not associate their names or email addresses with their answers. This information will help us study health patterns in LGBT and non-LGBT populations. If you would like to learn more about this research, click here.

9. References

Baker, R., J.M. Brick, N.A. Bates, M. Battaglia, M.P. Couper, J.A. Dever, K.J. Gile, and R. Tourangeau. 2013. *Report of the AAPOR Task Force on Non-Probability Sampling*. Oakbrook Terrace, IL: AAPOR. Available at: https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/NPS_TF_Report_Final_7_revised_FNL_6_22_13.pdf (accessed August 2019).

Bell, A. and M.S. Weinberg. 1978. *Homosexualities: A Study of Diversity Among Men and Women*. New York: Simon and Shuster.

Bengtsson, L., X. Lu, Q.C. Nguyen, M. Camitz, N.L. Hoang, T.A. Nguyen, F. Liljeros, and A. Thorson. 2012. "Implementation of Web-Based Respondent-Driven Sampling among Men Who Have Sex with Men in Vietnam." *PLoS ONE* 7(11): e49417. DOI: https://doi.org/10.1371/journal.pone.0049417.

Binson, D., J. Blair, D.M. Huebner, and W.J. Woods. 2007. "Sampling in Surveys of Lesbian, Gay, and Bisexual People." In *The Health of Sexual Minorities: Public Health Perspectives on Lesbian, Gay, Bisexual, and Transgender Populations*, edited by I.H. Meyer and M.E. Northridge, 375–418. New York: Springer.

Chandra, A., W.D. Mosher, C. Copen, and C. Sionean. 2011. *Sexual Behavior, Sexual Attraction, and Sexual Identity in the United States: Data from the 2006–2008 National Survey of Family Growth*. Hyattsville, MD: National Center for Health Statistics.

Available at: https://www.cdc.gov/nchs/data/nhsr/nhsr036.pdf (accessed September 2019).

Clark, J.L., K.A. Konda, A. Silva-Santisteban, J. Peinado, J.R. Lama, L. Kusunoki, A. Perez-Brumer, M. Pun, R. Cabello, J.L. Sebastian, L. Suarez-Ognio, and J. Sanchez. 2014. "Sampling Methodologies for Epidemiologic Surveillance of Men Who Have Sex with Men and Transgender Women in Latin America: An Empiric Comparison of Convenience Sampling, Time Space Sampling, and Respondent Driven Sampling." *AIDS and Behavior* 18(12): 2338–2348. DOI: https://doi.org/10.1007/s10461-013-0680-0.

Emory, K., Y. Kim, F. Buchting, L. Vera, J. Huang, and S.L. Emery. 2016. "Intragroup Variance in Lesbian, Gay, and Bisexual Tobacco Use Behaviors: Evidence That Subgroups Matter, Notably Bisexual Women." *Nicotine and Tobacco Research* 18(6): 1494–1501. DOI: https://doi.org/10.1093/ntr/ntv208.

Gates, G.J. 2014. *LGB/T Demographics: Comparisons among population-based surveys*. Los Angeles, CA: The Williams Institute, UCLA School of Law. Available at: http://williamsinstitute.law.ucla.edu/wp-content/uploads/lgbt-demogs-sep-2014.pdf (accessed April 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. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/geniuss-report-sep-2014.pdf (accessed August 2019).

Heckathorn, D.D. 1997. "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations." *Social Problems* 44(2): 174–199.

Heckathorn, D.D. 2002. "Respondent-Driven Sampling II: Deriving Valid Population Estimates from Chain- Referral Samples of Hidden Populations." *Social Problems* 49(1): 11–34.

IOM (Institute of Medicine). 2011. *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press. Available at: https://www.ncbi.nlm.nih.gov/books/NBK64795/ (accessed August 2019).

Johnston, L.G., M.M. Mon, M. Steinhaus, and J. Sass. 2017. "Correlates of Forced Sex Among Young Men Who Have Sex with Men in Yangon and Monywa, Myanmar." *Archives of Sexual Behavior* 46(4): 1001–1010. DOI: https://doi.org/10.1007/s10508-016-0761-z.

Laumann, E.O., J.H. Gagnon, R.T. Michael, and S. Michaels. 1994. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: University of Chicago Press.

Lee, J.G.L., G.K. Griffin, and C.L. Melvin. 2009. "Tobacco Use Among Sexual Minorities in the USA, 1987 to May 2007: A Systematic Review." *Tobacco Control* 18(4): 275–282. DOI: http://dx.doi.org/10.1136/tc.2008.028241.

McKnight, C., D. Des Jarlais, H. Bramson, C. McKnight, D. Des Jarlais, H. Bramson, L. Tower, A.S. Abdul-Quader, C. Nemeth, and D. Heckathorn. 2006. "Respondent-Driven Sampling in a Study of Drug Users in New York City: Notes from the Field." *Journal of Urban Health* 83(Suppl 1): 54. DOI: https://doi.org/10.1007/s11524-006-9102-1.

Miller, K.S. and J.M. Ryan. 2011. *Design, Development and Testing of the NHIS Sexual Identity Question*. Atlanta, GA: Questionnaire Design Research Laboratory. Office of

Research and Methodology, National Center for Health Statistics. Available at: https://www.impactprogram.org/wp-content/uploads/2011/11/Miller-2011-HHS-report-on-measuring-sexual-orientation.pdf (accessed August 2019).

National Institutes of Health. National Institute on Alcohol Abuse and Alcoholism. 2013. National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III). Available at: https://www.niaaa.nih.gov/research/nesarc-iii (accessed August 2019).

Pineau, V.J., J.M. Dennis, S. Michaels, S. Emery, and N. Ganesh. 2017. "Surveying Rare or Hidden Populations Using a Probability-Based Household Panel." In Joint Statistical Meetings Proceedings, Survey Research Methods Section, 3733–3746. Alexandria, VA: American Statistical Association. Available at: http://www.asasrms.org/Proceedings/y2017/files/594086.pdf (accessed August 2019).

Salganik, M. and D.D. Heckathorn. 2004. "Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling." *Sociological Methodology* 34: 193–239. DOI: https://doi.org/10.1111/j.0081-1750.2004.00152.x.

Schneider, J.A., B. Cornwell, A. Jonas, R. Behler, N. Lancki, B. Skaathun, S. Michaels, A.S. Khanna, L.E. Young, E. Morgan, R. Duvoisin, S. Friedman, P. Schumm, and E.O. Laumann. 2017. "Network Dynamics and HIV Risk and Prevention in a Population-based Cohort of Young Black Men Who have Sex with Men." *Network Science* 5(3): 381–409. DOI: https://doi.org/10.1017/nws.2016.27.

Schonlau, M., B. Weidmer, and A. Kapteyn. 2014. "Recruiting an Internet Panel Using Respondent-Driven Sampling." *Journal of Official Statistics* 30(2): 291–310. DOI: https://doi.org/10.2478/jos-2014-0018.

Sirken, M.G. 1998. "A Short History of Network Sampling. In Joint Statistical Meetings Proceedings, Survey Research Methods Section, 1–6." Alexandria, VA: American Statistical Association. Available at: http://www.asasrms.org/Proceedings/papers/1998_001.pdf (accessed October 2019).

Spira, A., N. Bajos, and the ACSF Group. 1994. *Sexual Behaviour and AIDS*. Aldershot, England: Avebury.

Strömdahl, S., X. Lu, L. Bengtsson, F. Liljeros, and A. Thorson. 2015. "Implementation of Web-Based Respondent Driven Sampling Among Men Who Have Sex with Men in Sweden." *PloS one* 10(10): e0138599. DOI: https://doi.org/10.1371/journal.pone.0138599.

Tourangeau, R., Edwards, B., Johnson, T.P., Wolter, K.M., and Bates, N. (Eds.). 2014. *Hard-to-Survey Populations*. Cambridge, England: Cambridge University Press.

Wejnert, C. and D.D. Heckathorn. 2008. "Web-Based Network Sampling Efficiency and Efficacy of Respondent-Driven Sampling for Online Research." *Sociological Methods and Research* 37(1): 105–134. DOI: https://doi.org/10.1177/0049124108318333.

Wellings, K., J. Field, and A.M. Johnson. 1994. *Sexual Behaviour in Britain: The National Survey of Sexual Attitudes and Lifestyles*. London: Penguin Books.

Received January 2019
Revised August 2019
Accepted August 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 753–776, http://dx.doi.org/10.2478/JOS-2019-0032

Surveying Persons in Same-Sex Relationships in a Probabilistic Way – An Example from the Netherlands

Stephanie Steinmetz[1] *and Mirjam Fischer*[2]

In the last decade, the call for improved estimates of lesbians, gay men and bisexual (LGB) populations has grown steadily. This is related to the increasing visibility of same-sex unions and the rapidly evolving changes in the legal and normative institutional frameworks regarding same-sex relationships in Western countries. The aim of this article is to present the sampling strategy and discuss the quality of a recently conducted probability-based survey in the Netherlands that targeted mixed-sex and same-sex couples with and without children. The core questions addressed are (1) whether the sampling strategy paid off in terms of identifying same-sex households and (2) whether the collected sample is representative of the target population. While the sampling strategy has success in identifying same-sex households, the question of representativeness remains a challenging task in surveying LGB populations and couples in particular. Especially, aspects related to the sampling strategy, the survey mode and the covered topic of the research are central to understanding observed selection patterns in the examined mixed- and same-sex samples.

Key words: Sampling LGB populations; sampling same-sex couples; selection biases, representativeness; probability-based sampling.

1. Introduction

In the last decade, the call for improved estimates of lesbians, gay men and bisexual (LGB) populations has grown steadily (Valfort 2017; OECD 2019). This is related to the increasing visibility of same-sex unions and the rapidly evolving changes in the legal and normative institutional frameworks regarding sexual diversity in Western countries. While the study of LGB persons in general, and same-sex couples in particular, has a strong tradition in qualitative methodological approaches (e.g., Barrett and Tasker 2001; Goldberg and Allen 2007; Hicks 2005; Patterson 1992), scholars continuously make efforts to overcome methodological challenges associated with the quantitative study of this population (Umberson et al. 2015). A prominent challenge in this regard is the establishment of a large representative sample of the LGB population by means of probability sampling.

[1] University of Lausanne, Institute for Social Sciences, Quartier UNIL-Mouline, Batiment Géopolis, 1015 Lausanne, Switzerland. Email: Stephanie.Steinmetz@unil.ch
[2] German Institute for Economic Research (DIW Berlin), Survey Methodology and Survey Management (SOEP Survey), Möhrenstraße 58, 10117 Berlin, Germany. Email: mfischer@diw.de
Acknowledgments: The UNICON project was funded by the Amsterdam Centre for Inequality Studies (AMCIS); the research group Institutions, Inequalities and Life courses (IIL) at the University of Amsterdam; Fonds Stimulering en Ontwikkeling by the Nederlandse Vereniging voor Seksuologie (NVVS) and the ERC project 'Family complexity' by Matthijs Kalmijn (FP7/2007-2013 / ERC grant agreement no. 669334).

By addressing the question how LGB populations can be sampled (considering LGBs as one group (persons in same-sex couples), which we compare to the population of persons in same-sex couples), the aim of this article is twofold. First, we want to know how well a recent probability-based survey in the Netherlands did in terms of identifying households where same-sex couples with and without children live (Unions in Context Study (UNICON); Fischer et al. 2017). Second, we want to know to what extent the collected sample is representative of the target population of partnered LGBs in the Netherlands. To this end, we use two available probability-based national surveys, the Netherlands Longitudinal Life Course Study (NELLS, Tolsma et al. 2014) and the Dutch Safety Monitor (VM), as benchmarks for the comparison of core socio-demographic characteristics of both mixed- and same-sex couples. The article contributes to the debate on how a specific subset of the LGB population, namely persons in same-sex relationships, can be sampled in a probabilistic and efficient way, leading to surveys which allow for reliable estimation of statistical parameters.

2. Challenges of Studying LGBs and Persons in Same-Sex Relationships in a Probabilistic Manner

Until the 1990s, convenience (non-probability) samples dominated research on LGB populations (see, for example in the United States, Kinsey et al. 1948; Bell and Weinberg 1978). And even though with the onset of the HIV/AIDS crisis in the early 1980s in the United States, several large probability-based population surveys started to include measures of sexual orientation, in particular health surveys in the United States and Europe (Galupo et al. 2016; Laumann et al. 1994; Spira et al. 1994; Wellings et al. 1994; Wolff et al. 2017), they remain often limited in scope (health, victimization or LGB-specific topics) and target population (only LGBs). This makes the type of substantive knowledge we have about LGB populations rather narrow.

As underlined by the study of the Institute of Medicine Committee on Lesbian, Gay, Bisexual and Transgender Health Issues (2011), there are three important challenges when trying to collect data on LGB populations in general and same-sex couples in particular: (1) defining and measuring sexual orientation, (2) overcoming the reluctance of some LGB individuals to identify themselves to researchers, and (3) obtaining high-quality samples of relatively small population segments.

One of the continuously discussed issues when collecting data on LGB populations is the optimal strategy of measuring sexual orientation (e.g., Almazan et al. 2009; Haseldon and Joloza 2009). Commonly measured dimensions of sexual orientation include self-assignment to identity labels (lesbian, gay, bisexual, etc.) or same-sex attraction and behavior (dating or sexual) (e.g., Dewaele et al. 2014; Meyer and Wilson 2009). While self-identification is often applied in topical surveys (in particular health-related surveys), an alternative measurement strategy that is frequently applied in bigger cross-national and/or multi-disciplinary surveys is the identification of LGBs via the self-reported gender and the self-reported relationship between two respondents of a household (partnership inference). The strategy of partnership inference to measure sexual orientation in such surveys bears great potentials in terms of topic coverage and cross-national comparisons, yet it only covers the partnered subset of the LGB population.

A further challenge is related to the sensitivity of the topic (such as sexual orientation, income, etc.) and the reluctance of some participants to disclose accurate information about themselves. In the case of sexual minorities, the disclosure bias might be generally related to privacy concerns or, more particularly, to avoid stigma and discrimination. When confronted with such a sensitive question, respondents may decline to answer or may intentionally give an inaccurate response (item nonresponse). In some cases, respondents may decide not to participate in the study at all (nonparticipation), thereby reducing the overall response rate and possibly making the sample less representative of the larger population. All of these outcomes have important implications for data quality (Lee 1993; Tourangeau and Yan 2007) as those who do not disclose their sexual orientation accurately, or decline to participate altogether, differ in relevant ways from those who do disclose and participate. Besides the importance of the cultural competence of the researcher studying hard-to-survey populations (e.g., Dillman et al. 2009; Tourangeau 2014), a number of techniques have been used to improve response rates to questions relating to sensitive topics. For instance, modes of data collection (such as the web) that foster participants' sense of confidentiality or anonymity may yield higher rates of disclosure (e.g., Valfort 2017; Villarroel et al. 2006).

Finally, as indicated by Meyer and Wilson (2009), as well as by Binson et al. (2007), careful sampling begins with the ability to enumerate (i.e., identify and count) the population of interest. If a population can be easily enumerated, it can be sampled in a fairly straightforward manner. However, sampling LGB populations is a challenging endeavor as they are what survey methodologists would describe as 'hard-to-survey' (Tourangeau 2014; Sudman et al. 1988; Solarz 1999). In particular, probability-based sampling presents an obstacle as there is no conceivable list of every LGB person or same-sex couple in a country that could serve as a sampling frame. As a consequence, it is hard to sample LGB respondents randomly (screening presents the only option, but this procedure is very costly and ineffective). Instead, research often relies on non-probability (convenience) sampling, which is problematic in terms of representativeness. For instance, studies that rely on (on- or offline) advertisement strategies to recruit respondents have persistent problems with the non-representativeness of their samples. This is due to the fact that volunteering for a survey – in particular an online survey – is often done by a rather selective group in terms of socio-demographic characteristics and motivations (Bolding et al. 2007; Meyer and Colten 1999; Rosser et al. 2009). For example, studies conducted among men who have sex with men based on web surveys for the United States and the United Kingdom found that certain variables were associated with higher rates of questionnaire noncompletion, such as nonwhite ethnicity, concealment of sexual orientation, self-identifying as heterosexual or bisexual, and in some cases, younger age (Evans et al. 2008; Ross et al. 2004). A further drawback of these studies is that they only focused on LGBs without including heterosexuals as a comparison group (European Union Agency for Fundamental Rights 2012 or for the United States, see Herek et al. 2010). As a consequence, findings based on such data are not comparative and cannot be generalized to a broader target population. In fact, such studies rather provide insights into very specific LGB subgroups, like persons who are recruited via health facilities, subscribers to certain magazines or visitors of LGB party venues. It is obvious that these subgroups may differ from the general LGB population with respect to health issues (when

recruited via health facilities) or a more outgoing lifestyle (when recruited at party venues). Researchers who want to generalize their findings to the population they are researching are in need of probability samples and appropriate sampling frames. In this respect, the use of population registers and census data to identify officially registered partnerships can be appealing (e.g., Andersson et al. 2006; Schwartz and Graf 2009 for the United States). Registers have the advantage that they provide a list of a country's entire population of officially registered same-sex couples (married or in a civil union). This can serve as sampling frame for a random sample, or the entire population can be analyzed. The large number of observations offers promising possibilities for quantitative studies of same-sex couples. The drawback of register data is that the range of phenomena that can be studied is often limited (for example to demographic information, such as income, employment and living situation) compared to social surveys, which cover larger thematic areas. Moreover, LGBs in register data constitute a selective group, namely those who are married to or in a civil union with a partner of the same sex. Cohabiting same-sex couples who are not married or in a civil union cannot be reached in this way.

When looking for probability-based large-scale surveys, which allow for the identification of LGBs (either in a direct or indirect way), the options tend to be limited and often tailored towards topics such as victimization, health and lifestyle. When searching for broader probability-based social survey data covering LGB populations in Europe, for instance, only two large-scale surveys come to the fore: the European Social Survey, ESS (2018), (http://www.europeansocialsurvey.org/) and the Generations and Gender Programme (GGS Waves 1, and 2, http://www.ggp-i.org/). These surveys allow for an indirect identification of the LGB population via partnership inference. A big drawback of population-based probability surveys is that in such data, the proportions of men and women who identify as non-heterosexual are often small – around 1% (Rothblum 2007). On the one hand, this creates statistical power issues for the estimation of reliable parameters. On the other hand, otherwise harmless mistakes, such as recording errors (e.g., recording one partner's gender wrong and hence misclassifying the couple), can cause distorted findings in the numerically small group (see Banens and Le Penven 2016; Cortina and Festy 2014; DeMaio et al. 2013; Fischer 2016; Regnier-Loilier 2018).

3. Targeting Cohabiting Same-Sex Couples in the Netherlands – The UNICON Project

Given the lack of quality representative survey data on the LGB population and persons in same-sex relationships (OECD 2019), the UNICON project collected representative survey data among lesbian, gay and heterosexual couples and families in the Netherlands in 2016. In particular, it aimed at creating a carefully crafted data set, which (1) purposely includes persons in same-sex couples in a sampling frame, (2) uses a random sample, (3) oversamples persons in same-sex couples to achieve a large number of observations, (4) includes a suitable comparison group of persons in mixed-sex couples, and (5) covers a wide range of LGB-specific and more general topics (see Fischer et al. 2017). Note that the joint study of LGB and heterosexual respondents poses unique challenges in terms of the framing of the study to potential respondents. However, in order to study structural

disadvantage along sexual orientation lines, it is indispensable to study both groups jointly (Umberson et al. 2015; Valfort 2017). Therefore, the UNICON data is unique in the way it incorporates both LGB and heterosexual respondents and covers both general and LGB-specific topics.

3.1. Sampling Design

For the project, a two-stage sampling design was used. Firstly, 36 municipalities were selected that vary in size, degree of urbanization and geographical region. The Netherlands is divided into 380 administrative units in 2016 – the municipalities. Often a municipality equals a large city and its suburbs. Outside the highly urbanized areas, a municipality typically encompasses multiple smaller towns or a cluster of villages. The stratified selection of municipalities is important, given the stark differences regarding population size, ranging from more than 840,000 inhabitants in Amsterdam to just short of 1,000 in the smallest municipality (Statistics Netherlands 2018). The municipality selection for the UNICON study was based on a different survey, the NELLS, which has sampled the municipalities according to an urbanization- and region-stratified sampling method (De Graaf et al. 2010). The UNICON is based on the same sample of municipalities in order to allow a joint study of the two surveys. This provides many advantages, such as obtaining detailed data on attitudes toward gender and family values and homosexuality from the NELLS to study individual-level phenomena in the UNICON in relation to the attitudinal context (see Fischer 2019). The stratified municipality sample ensures geographic distribution and variation in the way people think about topics such as homosexuality (Kuyper 2016).

Secondly, the local authorities in each municipality were contacted by a phone call inviting them to cooperate with the research project and follow-up letters with detailed information about the project. In the Netherlands, municipalities can decide whether they want to cooperate with scientific institutes by sharing information from the population registers. Those municipalities which agreed to participate proceeded to draw a random sample of households from their registers. The sample of households was stratified according to three household types: (1) cohabiting mixed-sex couple households, (2) cohabiting same-sex couple households without children, and (3) cohabiting same-sex couple households with children. The main motivation for choosing these groups was the intent to oversample same-sex couple households, in particular those with children, so that the group would be large enough to make statistical comparisons between them. For mixed-sex couples, it could be expected that a substantial share would have children without making this an explicit sampling criterion. Among same-sex couples, however, the groups have been differentiated, since same-sex couples have children far less often than mixed sex couples (Andersson et al. 2006; Gates 2013; Kuyper 2016).

Based on the Dutch population registers, it is possible to identify persons who are in legally registered unions, such as marriage or civil partnerships. As the aim of the project was to include all couples, also those who cohabitate without being married or in a civil union, an approximation strategy was applied. That is, we used a number of known characteristics from the registers in order to make an estimate of households where it is likely that couples live. The list of these approximated households then

served as the list of addresses (the population of interest) from which the probability sample was drawn.

We approximate couple households in the following manner: we selected households where two persons live, who (1) are between the age of 30–65; (2) are of the same or different sex, (3) do not have a parent–child relation with each other, (4) are not each other's siblings, and (5) live with a child under the age of 18. The age limit was implemented in order to avoid contacting respondents who are (a) not in our target group (i.e., student households where young people of the same sex often live together as flat mates rather than romantic partners) and (b) who are less likely to participate in web mode surveys (i.e., persons older than 65 years). We deemed this strategy appropriate in order to maximize the final sample size given limited financial funds of the project. Criteria 3 and 4 were central to ensuring that two persons were likely a cohabiting couple. For type 3 households – same-sex couples with children – the last criterion (5) applies. We intentionally did not rely on registered parent–child relationships in order to allow for (legally) complex parenthood constellations among same-sex parents to be captured.

The contact with respondents was established by sending a letter to each household with an invitation for both partners in the couple to participate in our web survey. We decided to not feature the LGB topic prominently in the invitation letter to the participants. Instead, we emphasized diversity in living arrangements and families in the hope that this would be attractive enough for same-sex couples to feel encouraged to take part, while reducing possible threats to (selective) response for the mixed-sex couples. In the survey, we confirmed whether the contacted persons were indeed a couple. To enhance the response rate among same-sex couple households – the main target group in this study – they received incentives entailing a prepaid voucher worth EUR 5 for a popular Dutch online retailer. Depending on participation, households received up to three reminders via mail over the course of multiple weeks (Fischer and Steinmetz 2018).

3.2. Did the Sampling Strategy Pay Off in Terms of Reaching the Target Population?

Twenty-two municipalities of the 36 municipalities provided a sample of addresses from their population registers for the UNICON study. This translates to a total of 22 participating municipalities and a response rate of 61% at the municipality level. The final sample is based on 20 municipalities under exclusion of two with too low participation. Table 1 shows the distribution of the realized sample. The numbers in parentheses show the number of municipalities that was sampled. What stands out is the observation that almost all strongly urban municipalities that were sampled also participated, whereas participation was lower in the marginal and moderate ones. This is likely related to the fact that larger municipalities are used to receiving research proposals like ours and they have the capacity to handle them quickly. Smaller municipalities sometimes do not have a statistics department and have to spend more time and resources into fulfilling such a request. There are no notable patterns regarding the regional nonresponse among municipalities.

At the individual level, the UNICON data collected information of 1,353 individual respondents (in 880 households), including 510 persons in mixed-sex couples and 843 persons in same-sex couples. Of those 843 persons in same-sex couples, 267 indicated that

Table 1. Distribution of the realized sample of municipalities in the UNICON survey; original sample in parentheses.

Region	Degree of urbanization			Total
	marginal	moderate	strong	
North/East	2 (3)	3 (4)	1 (3)	6 (10)
West	1 (3)	2 (4)	6 (10)	9 (17)
South	1 (3)	0 (4)	4 (2)	5 (9)
Total	4 (9)	5 (12)	11 (15)	20 (36)

they lived with one or more children. Same-sex households with children were almost exclusively women. The response rate at the household level was 24.5%. It was highest among same-sex households with children (34%), followed by same-sex households without children (27%) and mixed-sex households (20%). This is likely a reflection of the topic of the survey, which was advertised as a study about diversity in living arrangements. These differences can be observed in all three geographical regions (see Fischer et al. 2017). In municipalities of moderate urbanization, same-sex couples had a slightly lower response than mixed-sex couples. In all other municipalities, same-sex couples participated more often. This is in line with another Dutch study among lesbian and heterosexual parents in the Netherlands, which uses a wide range of recruitment strategies, and also observed a higher response among lesbian families compared to heterosexual families (Bos 2004). The overall participation was highest in the North/East (27%), followed by the West (23%) and the South (21%). Overall, the response rate in the UNICON is modest, yet reasonable for a web-based survey in times of declining response rates and survey fatigue. Even a response rate below 10% is not uncommon for web surveys (Conrad et al. 2010; Lozar Manfreda et al. 2008; Munoz-Leiva et al. 2010; Shih and Fan 2008; Smyth and Pearson 2011).

As indicated, it was possible (but not strictly required) for both partners in the household to respond to our survey. In half of the cases, we do have information from both partners. As the sample was stratified by region and degree of urbanization, and the obtained sample sizes and response rates varied from municipality to municipality, weights were constructed to make the sample nationally representative. To do so, we first obtained from Statistics Netherlands the population of couples in each combination of region and urbanization separately for the three types of couples (mixed-sex, same-sex without children, same-sex with children).

To answer the question whether our approximation strategy was successful in identifying the target population, we calculated the accuracy of the sampling strategy. Table 2 shows how many households that were sampled as one of the three household types confirmed their status in the survey. The calculations show that 99.1% of the households that were sampled as mixed-sex couples via the approximation strategy turned

Table 2. Sampled from registers (expected) and respondent self-reports (observed).

	Expected		
Observed	*Household type 1* Mixed-sex with and without children	*Household type 2* Same-sex without children	*Household type 3* Same-sex with children
Mixed-sex with and without children	**317 (99.1%)**	20 (4.4%)	5 (3.6%)
Same-sex without children	1 (0.3%)	**352 (78.2%)**	3 (2.1%)
Same-sex with children	0 (0.0%)	38 (8.4%)	**130 (92.9%)**
Single	2 (0.6%)	40 (8.9%)	2 (1.4%)

Note: Single = respondent without a partner. Calculations based on N = 923 households in 19 municipalities as one municipality did not want to include the household type in the sample and could therefore not be included in this calculation.

out to be mixed-sex couples. Among same-sex households with children the accuracy was high (92.9%). The approximation strategy was somewhat weaker for same-sex households without children, where the accuracy lies at 78.2%.

A number of households that were sampled as same-sex households without children turned out to be same-sex couples with children (8.4%) and they remain in the dataset. This relatively high number of children in households where we did not expect them may further point towards the fact that parenthood in same-sex couples can be legally complex. An equally large number of households turned out to be singles of the same sex who share a household but who have not reported a romantic relationship with each other. Our lower age bound of 30 is not a foolproof criterion for excluding shared flats, as young working professionals often remain in these living situations well into their thirties (Kenyon and Heath 2010). Finally, between 3% and 4% of the couples that have been sampled as same-sex couples with and without children turned out to be mixed-sex couples. Registration mistakes at the municipality level are one possible explanation. Another could be that people deliberately provide false information to the authorities in order to creatively navigate the extremely tensed housing market. All these cases come from highly urban areas, which makes such an explanation plausible.

Overall, the applied approximation strategy to identify same-sex couple households in the register data proved to be successful in obtaining a sizable group of same-sex households. The proposed research design payed off and proved to be an innovative strategy to create a sampling frame for a population, for which creating a sampling frame is otherwise simply not possible. As such, the UNICON data are an advancement in collecting probability-based survey data for LGB populations. It is the biggest probability-based survey on same-sex couples and families, which allows studying a broad range of topics (such as well-being, social integration and relationship quality) in comparison with heterosexual couples in the Netherlands. As previously stated, this comparative aspect is crucial when trying to reveal the structural inequality of one social group in comparison with the majority.

4. Evaluating the Selectivity and Representativeness of the Collected Sample

To evaluate the question of representativeness of the collected data, it would be ideal to compare characteristics of the collected survey data to those in the population. However, as previously discussed, the availability of probability data on persons in same-sex relationships and LGBs is limited in the Netherlands and elsewhere (OECD 2019). Hence, it is nearly impossible to find an existing representative data source that includes LGBs in the Netherlands that could be used as benchmark for our comparison. This illustrates again the important gap that the UNICON survey addresses for LGB studies in the Netherlands.

We have applied two strategies, which can give us some insights on the representativeness of the UNICON data. First, we considered the survey, which formed the basis for the municipality selection (NELLS) and compared the respondents in mixed-sex couples to each other. Second, we use the only nationally representative survey in the Netherlands that allows the identification of partnered LGBs to compare them to the persons in same-sex relationships in the UNICON.

The former is an attempt at evaluating how well the UNICON reached one of our three target groups, namely persons in mixed-sex couples. The latter comparisons is an attempt to examine this LGB-specific response obstacles more directly. Here, the respective ways of identifying persons in same-sex relationships differ, which is not ideal. Yet, the totality of these comparisons can provide an initial sense of how representative the UNICON data might be, in the absence of true population data. In the following, we elaborate on these two national benchmark surveys we used for the comparison and our analytical strategy.

4.1. Two National Benchmark Surveys

First, we turned to the NELLS in order to compare demographic characteristics of persons in mixed-sex couples. The NELLS is a nationally representative, large-scale survey of the Dutch population aged 15–45 (for more details, see De Graaf et al. 2010). It was conducted between December 2008 and May 2010, partly through face-to-face interviews and partly through self-completion questionnaires. The data contain information on 5,312 individuals from 35 municipalities in the Netherlands. The survey yielded an overall response of 52%. The comparison with the NELLS is an obvious choice, since the selection of municipalities in the UNICON is based on the municipality sample of the NELLS. It is therefore possible to use the two surveys for comparing respondents within the same municipalities. Another advantage of this comparison is the fact that persons in mixed-sex relationships were identified in the same way in both surveys, namely by means of partnership inference (i.e., the combination of the respondent's own gender and the gender of the partner reported by the respondent). One drawback is that the age range covered in the NELLS (15–45 years) and the UNICON (30–65 years) only overlaps by 25 years. Therefore, we were forced to limit our comparison to persons in mixed-sex relationships between the ages of 30 and 45 years.

Second, we used data from the *Dutch Safety Monitor* (VM, pooled annual data 2012–2015) for the comparison between persons in same-sex relationships. The VM is an annual large-scale survey among individuals aged 15 years and older focusing on various aspects related to safety, experiences of violence, and crime. The nationally representative survey covers many observations, with a minimum of 65,000 respondents annually.

Response rates typically lie at 44% and the questionnaires of the VM are primarily completed online (Statistics Netherlands 2016). If respondents do not participate right away, they receive a reminder via mail that includes the option to complete a paper and pencil version of the questionnaire. The VM is among the only representative surveys in the Netherlands that allows an identification of LGBs via self-reported sexual attraction. In order to render the surveys as comparable as possible, we have included a number of restrictions on the VM data. We selected respondents who were married or in a civil union to ensure that we are dealing with partnered persons. We also restricted the sample to those who reported same-sex attraction. Respondents who indicated that they are attracted to both sexes could be in mixed-sex relationships. Of course, this risk remains a possibility also with those reporting same-sex attraction since the different dimension of sexual orientation – identity, attraction and behavior – do not necessarily overlap (Sell 1996). However, we minimized the problem by excluding those who are attracted to both sexes. Finally, we imposed the UNICON age range (30–65 years) to the VM sample. Table A3 in the appendix (Section 6) provides an overview of how we have accomplished a comparison between the surveys.

4.2. *Measures*

We compared socio-demographic characteristics, such as gender (women versus men), age in cohorts, educational level (low (ISCED 0–2), medium (ISCED 3–4), and high (ISCED 5–6)) based on the International Standard Classification of Education (UNESCO 2012), whether a person has children (yes/no), is married (yes/no), and has a paid job (yes/no). To explore possible interesting outcome variables on which same-sex and mixed-sex couples could potentially differ, we also compared, in the case of mixed-sex couples, church attendance (once a month or more versus rest) and political part preference (Dutch left versus rest). For the comparison of same-sex couples, the variables of neighborhood cohesion (index ranging from 0 indicating low cohesion to 5 indicating high cohesion) and of having been a victim of violence in the last five years (yes/no) have been used. For a detailed overview of the operationalization, see Table A4 in the appendix.

4.3. *Method and Strategy*

As mentioned before, in a first step of evaluating the comparability between the UNICON data and the reference surveys we focus on mixed-sex couples, assuming that if the differences between the samples are not significant, this is at the least a good sign for the same-sex couples in the UNICON data as well. We start with mean comparisons (percent in the UNICON minus percent in the benchmark survey), using the weighted benchmark surveys as a reference for the population we were targeting. On the basis of a t-test it is determined whether the observed difference in means is also statistically significant. This serves as a first crude exploration of possible selectivity of the UNICON sample. However, to determine not only the selectivity patterns of the samples, but also the magnitude of the differences in more detail, we used *Average Relative Differences* (ARD) (for a comparable application, see Steinmetz et al. 2014). The *ARD* for the covariate x (with p categories) is defined by $\frac{1}{p}\sum_{j=1}^{p}|d_j|$, where d_j is the relative difference for category j, which is defined by

$d_j = \frac{Perc_j^{UNICON} - Perc_j^{POP}}{Perc_j^{UNICON}}$. The ARD can assume any value between 0 and $+\infty$ In a final step, we examined bivariate associations between possible outcomes of interested and core socio-demographic variables. This is done on the basis of correlation analyses to determine whether similarities or differences (in terms of significance and sign) between the samples could be determined.

4.4. Results

4.4.1. Comparison of Mixed-Sex Couples

Figure 1 shows the differences in the weighted means for core socio-demographic characteristics, as well as some potential comparable outcome variables (church attendance and party preferences) for mixed-sex couples.

A few differences between the two samples stand out. In the UNICON data, women, people aged 35–39, people who have paid work and people who are married are significantly overrepresented, while people aged 30–34, people with medium education and people with children are underrepresented. A comparison with the ARD (see Table A1 in the appendix) partly confirms the findings of Figure 1. The magnitude of the selectivity seems to be particularly high for age (0.56), education (0.31) and church attendance (0.26), whereas it is lowest for left party preference (0.05).

In a next step, we examined whether similarities and differences between the two samples can be observed when looking at bivariate relationships between possible outcome variables and socio-demographic characteristics (Table 3). This is motivated by the fact that researchers might be interested in using the data beyond purely descriptive purposes (and even extending it to a multivariate framework).

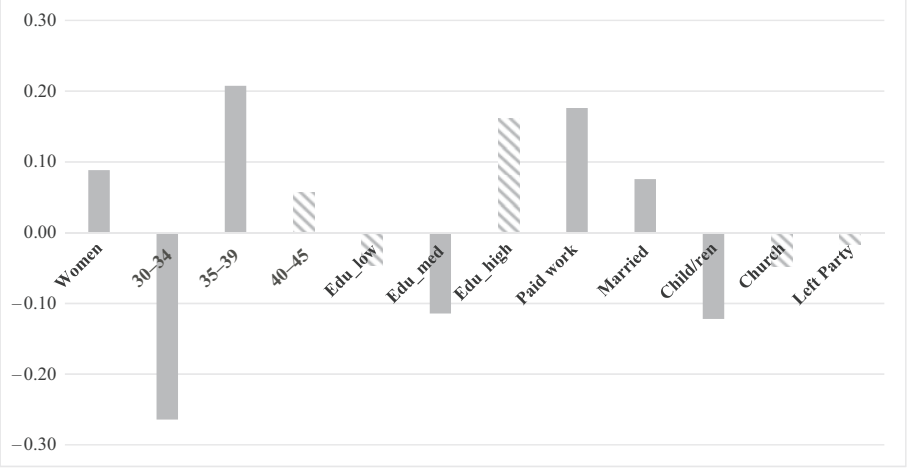


Fig. 1. Weighted mean differences in percent, UNICON versus NELLS data – individuals in mixed-sex couples.
Source: UNICON 2016 (N 452) and NELLS 2009 (N 1495), 21 municipalities.
Note: Age is restricted to 30–45 years. Positive values indicate an overrepresentation of the selected characteristics in the UNICON data, whereas a negative values indicates an underrepresentation. The grey bars indicate a significant difference, while the striped bars represent insignificant differences.

Table 3. Correlation analyses for church attendance and left-party preference, UNICON versus NELLS, persons in mixed-sex couples.

	Church attendance		Left party preference	
	UNICON	NELLS	UNICON	NELLS
Women	−0.006	−0.178***	0.084	−0.030
30–34	−0.028	−0.027	0.125	0.005
35–39	0.103	−0.002	−0.089	0.001
40–45	−0.081	0.028	−0.014	−0.005
Edu_low	−0.048	0.127***	−0.111*	−0.015
Edu_med	0.069	0.007	−0.074	−0.026
Edu_high	−0.031	−0.130***	0.140**	0.042
Paid work	−0.084	−0.064	0.063	−0.004
Children	0.053	0.138***	−0.118*	−0.023
Married	0.171***	0.256***	−0.160**	0.015
N	452	1495	452	1495

Source: UNICON 2016 (N = 452) and NELLS 2009 (N = 1495), 21 municipalities.
*p < 0.05, **p < 0.01, ***p < 0.001.
Note: Age is restricted to 30–45 years.

Starting with the correlations for church attendance, besides being married, none of the variables shows a significant correlation within the UNICON sample, whereas in the NELLS data, being a woman, having a low level and high level education, having children and being married is significantly associated with church attendance. The lack of significant correlations in the UNICON data might be attributed to the smaller sample size in comparison to the NELLS data. In addition, only for the variable low education can we observe a divergent sign (negative in UNICON and positive in NELLS) that could hint towards a bias in the UNICON data, where we have a lower share of low-level educated respondents (see Figure 1). Turning to left party preferences, a slightly different picture emerges. While none of the correlations in the NELLS is significant, we can observe four significant correlations within the UNICON (namely for low-level and high-level education, having children and being married). However, as before only one correlation (being married) deviates strongly also in terms of the sign (negative and significant in the UNICON and positive but not significant in the NELLS). Again, this might indicate an underrepresentation of non-married people in the UNICON, which is no surprise given that being a couple was one of our sample criteria.

4.4.2. Comparison of Same-Sex Couples

Continuing with the comparison, Figure 2 shows the weighted mean differences for core socio-demographic characteristics, as well as two potential outcome variables (neighborhood cohesion and victim of violence) for same-sex couples. Overall, the results indicate that only for four out of 14 comparison groups, significant differences can be observed. It appears that people with low-level education are underrepresented in the UNICON data, while women, highly educated people and people who have paid work are overrepresented. Also, here a comparison with the ARD (see Table A2 in the appendix)

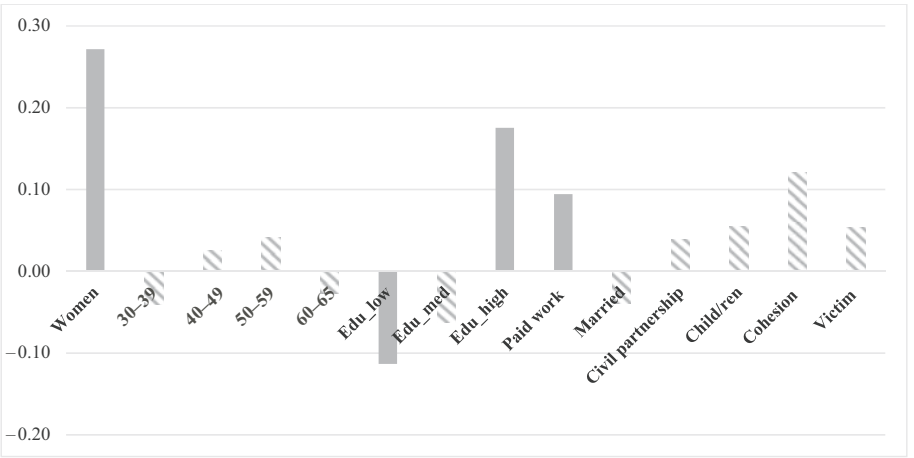


Fig. 2. Weighted mean differences in percent, UNICON versus VM data – persons in same-sex couples.
Source: UNICON 2016 (N 452) and VM pooled 2012–2015 (N 355), 16 municipalities.
Note: Positive values indicate an overrepresentation of the selected characteristics in the UNICON data, whereas a negative values indicates an underrepresentation. The grey bars indicate a significant difference, while the striped bars represent insignificant differences.

partly confirms the findings of Figure 2. The magnitude of the selectivity among same-sex couples appears to be particularly high for gender (0.76), being a victim of violence (0.60), and education (0.38), while it is lowest for neighborhood cohesion (0.04).

For the comparison between persons in same-sex relationships, in a final step, we also examined which similarities and differences between the two samples come to the fore when looking at bivariate relationships (Table 4). Starting with the correlation between being a victim of violence and socio-demographic variables, in none of the samples a

Table 4. Correlation analyses for being a victim of violence, UNICON versus. VM, persons in same-sex couples.

	Victim of violence last 5 years		Neighborhood cohesion	
	UNICON	VM	UNICON	VM
Women	0.006	−0.005	0.180***	0.011
30–39	0.061	0.050	−0.043	−0.001
40–49	−0.053	0.086	0.015	−0.078
50–59	0.023	−0.077	0.000	0.087
60–65	−0.034	−0.070	0.032	−0.006
Edu_low	0.007	−0.040	−0.035	−0.144**
Edu_med	−0.036	0.015	−0.143***	−0.078
Edu_high	0.029	0.015	0.149***	0.176**
Paid work	0.019	0.012	−0.037	0.098
Child/ren	−0.033	0.019	0.220***	−0.043
Married	0.030	0.040	−0.045	−0.099
Civil partnership	−0.030	−0.040	0.045	0.099
N	585	355	585	355

*p < .05, **p < .01, ***p < .001.

significant correlation can be observed. We can see some differences in signs; however, due to the small samples sizes in both cases and the very weak correlations, these deviations can be considered as irrelevant. However, for the variable neighborhood cohesion we find four significant correlations in the UNICON data for being a woman, having a medium-level and high-level education, and having children. For the VM data, we find only two significant correlations, namely for having a low-level and high-level education. This leaves us with only two deviations regarding the significance of the correlations, and we do not find any deviation with respect to the sign. The differences in significance in this case is less related to the difference in the sample size between the samples, but might hint towards selectivity patterns in the two surveys with respect to education (see also Figure 2).

Overall, when examining the representativeness of the UNICON, in particular for the same-sex sample, the observed selectivity seems to be less severe compared with the mixed-sex sample. However, the difference between the UNICON and the NELLS might be related to the fact that the response rate was lowest among persons in mixed-sex relationships in the UNICON (20%). Therefore, it seems reasonable that we observe some difference. However, it seems difficult to determine the reason for the observed patterns within a sample and the differences across the samples. Even though we have applied sample weights for all analyses, part of the observed selectivity might be the result of different intertwined issues: a) the targeted sampling strategy (with an approximation of household types and an overrepresentation of same-sex households), b) the topic of the survey, which might have been more appealing, for instance, to women and c) to the mode of the survey (web-based) that accommodates the participation of particular respondents.

5. Discussion and Conclusion

The present article aimed to answer two core questions in the evaluation process of the recently conducted broadly-oriented probability-based UNICON survey in the Netherlands targeting mixed-sex and same-sex couples with and without children. With respect to the first question – whether our sampling strategy paid off in terms of identifying same-sex households – we can conclude that the strategy was successful. Based on an innovative sampling strategy, we were able to collect a probability survey that allows the comparative analysis of 843 persons in same-sex relationships and 510 persons in mixed-sex relationships.

Regarding our second question – to what extent the collected sample is representative of the target population, we can conclude that, in particular for the UNICON same-sex couple sample – which was our main target – selectivity seems to be reasonably low. The observed difference between the surveys is likely related to the fact that the identification of sexual orientation differs between the two surveys. Regarding the bigger deviations for the mixed-sex couple sample in the UNICON and the NELLS, the low response rate likely has to do with the stronger selectivity pattern. For some of these selectivity patterns, different challenges might be closely related. First, general challenges arise from the web mode of the survey. People who do not have easy access to a computer and the Internet (although Internet penetration rates are high in the Netherlands, 98% in 2017, see CBS 2018), who are illiterate or do not fully

master the Dutch language, are by default excluded from participating in the study. Although a web survey might be a good opportunity to receive information on sensitive topics, such as sexual orientation (Villarroel et al. 2006; for a review, see Gribble et al. 1999), we could not reach these nonresponse populations with telephone or face-to-face interviews. Second, nonresponse might be related to the way in which we have framed the main purpose of the study. Participation depends on the personal interest of people in the survey topic (family complexity/diversity in living arrangements) and their motivation to support the research. In particular, LGB populations in the Netherlands are of high interest to researchers and politicians. This can lead to an alertness and sensitization of this group towards any kind of research (phenomenon of an over-researched population). In addition, we also have to recognize that the evaluation of a sample based on a benchmark survey can be challenging in particular when the different modes and measures are used (see Appendix, Table A3). However, in the absence of population information, this can be considered a first step.

Although, we are confident to have collected a quality probability-sample of mixed-sex and same-sex couples in the Netherlands it is important to be aware that those who declined to participate in the UNICON survey might differ in relevant ways from those who did participate (e.g., Tourangeau and Yan 2007). To correct for these biases, additional and more advanced weighting techniques could be applied to adjust the sample to the population of interest. However, new challenges arise with respect to available population benchmarks, particularly for same-sex couples. Biases might also affect the weighted benchmark surveys when it comes to the LGB population. Applied weighting models may be insufficient to alter these biases since they cannot correct the distributions to a sampling frame of LGBs. As long as reliable population information on LGBs and same-sex couples is missing, determining and correcting selection biases of hard-to-survey groups will remain a challenge.

6. Appendix

Table A1. Average relative differences (ARD) between the UNICON and the NELLS data, mixed-sex couples.

	ARD UNICON versus NELLS
Age	**0.56**
Education	**0.31**
Church	**0.26**
Paid work	0.25
Women	0.19
Child/ren	0.16
Married	0.11
Left Party	**0.05**

Source: Unicon 2016 (N = 452) and NELLS 2009 (N = 1495).
Note: Age is restricted to 30–45 years.

Table A2. Average relative differences (ARD) between the UNICON and the VM data, samesex couples.

	ARD UNICON versus VM
Women	**0.76**
Victim of violence	**0.60**
Education	**0.38**
Child/ren	0.18
Married	0.17
Age	0.15
Paid work	0.12
Neighbourhood cohesion	**0.04**

Source: Unicon 2016 (N = 585) and VM pooled 2012–2015 (N = 355).
Note: Age is restricted to 30–45 years, consideration only of couples who are married and in registered partnerships.

Table A3. Overview of the data sources.

Comparison 1: Mixed-sex couples			
	UNICON	NELLS	Comparison
Sexual orientation measure	Partnership inferred sexual orientation	Partnership inferred sexual orientation	Mixed-sex couples only
Age range	30–65	15–45	30–45
Municipalities	20	35	19 municipalities which overlap
Survey mode	Web mode	Face-to-face	
Comparison 2: Same-sex couples			
	UNICON	VM	Comparison
Sexual orientation measure	Partnership inferred sexual orientation	Attraction inferred (individuals not necessarily partnered)	Limited to same-sex couples in legally registered unions
Age range	30–65	15+	30–65
Municipalities	20	35	16 municipalities which overlap
Survey mode	Web mode	Web mode and paper-and-pencil	

Table A4. Operationalisation overview.

	Recoded variables used in comparison		Original variables and how they are recoded	
		UNICON	NELLS	VM
Gender	0 men, 1 women	What is your gender? 0 man, 1 woman	What is your gender? 0 man, 1 woman	What is your gender? 0 man, 1 woman
Education		What is the highest education you completed with a diploma?	What is the highest level of education you have followed? Have you obtained a diploma from this education? y/n;	What is the highest education you completed with a diploma?
	1 low = no education completed, early childhood education, primary education, lower secondary education (ISCED0- 2);	**1 low** = no schooling completed, elementary [lo], lower secondary vocational [lbo], lower secondary general [mavo];	**1 low** = lagere school, lbo, vmbo- kb/bbl, mavo, vmbo-tl;	**1 low =** geen opleiding, lagere school, Lager Beroepsonderwijs (LBO, LTS), VMBO basisberoepsgerichte of kaderberoepsgerichte leerweg, Mavo, VMBO theoretische of gemengde leerweg, ULO, MULO;
	2 medium = upper secondary and post-secondary non-tertiary education (ISCED3-4);	**2 medium** = intermediate vocational [mbo], higher secondary general [havo-vwo];	**2 medium** = havo, vwo/gymnasium, mbo-kort (kmbo), mbo-tussen/lang (mbo)	**2 medium** = Havo, VWO, Gymnasium, HBS, MMS, Middelbaar beroepsonderwijs (MBO, BOL, BBL);
	3 high = tertiary education such as Bachelor's or equivalent and higher (ISCED5-6)	**3 high** = higher vocational [hbo], university [wo], post-academic [wo+]	**3 high** = hbo, universiteit, promotietraject	**3 high** = Propedeuse, Kandidaats, Bachelor, Hoger Beroepsonderwijs (HBO), Doctoraal, Master, semi- Wetenschappelijk onderwijs

Table A4. Continued.

	Recoded variables used in comparison		Original variables and how they are recoded	
		UNICON	NELLS	VM
Child/ren	0 no, 1 yes	Respondents report how many living children they have (also counting stepchildren). For each child that is reported to live in the household and who is younger than 18 years the **value 1** is assigned.	Respondents report how many living children they have (also counting stepchildren). For each child that is reported to live in the household and who is younger than 18 years the **value 1** is assigned.	Information retrieved from population registers
Married	0 no, 1 yes	Are you married to your partner? y/n	Are you married to your partner? y/n	Information retrieved from population registers
Paid job	0 no, 1 yes	Do you currently have paid work? y/n	Do you currently have paid work? y/n	Which best describes you current situation? **0** = unemployed, volunteer, unable to work, student, homemaker, pensioner; **1** = doing paid work, selfemployed
Church attendance	0 less than once a month, 1 once a month or more	How often attend religious service (besides special occasions)? 0 = almost never, a few times a year; 1 = regularly (e.g., every month or week)	How often do you go to a service in a church, mosque, synagogue, temple or shrine? 0 = never, 1–2 per year, 3 – 11 times per year; 1 = once a month, 2–3 times per month, every week, several times a week	

Table A4. Continued.

	Recoded variables used in comparison	Original variables and how they are recoded	
		UNICON	NELLS	VM
Party preference	1 left parties, 0 all other parties	Which party do you currently prefer? **1 left** = PvdA, SP, GroenLinks, PvdD; **0 rest** = CDA, VVD, PVV, Christen Unie, D66, SGP, 50Plus, Other	Which party do you currently prefer? **1 left** = PvdA, SP, GroenLinks, PvdD; **0 rest =** CDA, VVD, PVV, Christen Unie, D66, SGP, Groep Verdonk, Other	
Victim of violence	0 no, 1 yes	During the past 5 years, were you ever attacked or threatened with violence at home or elsewhere? y/n		During the past 5 years, has anyone ever attacked you or threatened with violence? y/n
Neighbourhood cohesion	Mean score on index of 3 items; range 0 – 5; all items recoded so that a high score represents high neighborhood cohesion	(a) Most people know each other in this neighborhood. (b) I feel at home with the people who live in this neighborhood. (c) I have a lot of contact with the other residents in the neighborhood. *1 disagree strongly – 5 agree strongly*		(a) People in this neighborhood barely know each other. (b) I feel at home with the people who live in this neighborhood. (c) I have a lot of contact with the other residents in the neighborhood. *1 agree strongly – 5 disagree strongly*

7. References

Almazan, E. et al. 2009. "*Best Practices for Asking Questions about Sexual Orientation on Surveys?*" Los Angeles: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/SMART-FINAL-Nov-2009.pdf (accessed September 2019).

Andersson, G., T. Noak, A. Seierstad, and H. Weedon-Fekjaer. 2006. "The demographics of same-sex marriages in Norway and Sweden." *Demography* 43: 79–98. DOI: http://doi.org/10.1353/dem.2006.0001.

Banens, M. and E. Le Penven. 2016. "Sex miscoding in the census and its effects on the enumeration of same-sex couples." *Population* 71(1): 131–143. DOI: http://doi.org/10.3917/popu.1601.0135.

Barrett, H. and F. Tasker. 2001. "Growing up with a gay parent: Views of 101 gay fathers on their sons' and daughters' experiences." *Educational and Child Psychology* 18: 62–77.

Bell, A. and M. Weinberg. 1978. "*Homosexuality's: a study of diversity among men and women*." New York: Simon and Shuster.

Binson, D., J. Blair, D. Huebner, and W. Woods. 2007. Sampling in surveys of lesbian, gay and bisexual people. In *The health of sexual minorities: Public health perspectives on lesbian, gay, bisexual and transgender populations*, edited by I. Meyer and M. Northridge, 375–418. New York: Springer.

Blair, J. 1999. "A probability sample of gay urban males: The use of two-phase adaptive sampling." *The Journal of Sex Research* 36: 39–44. DOI: http://doi.org/10.1080/00224499909551965.

Bolding, G., M. Davis, G. Hart, L. Sherr, and J. Elford. 2007. "Where young MSM meet their first sexual partner: The role of the Internet." *AIDS & Behavior* 11(4): 522–526. DOI: http://doi.org/10.1007/s10461-007-9224-9.

Bos, H. 2004. "*Parenting in Lesbian Families*." Amsterdam: Amsterdam University Press.

CBS. 2018. "The Netherlands leads Europe in internet access" (press release). Available at: https://www.cbs.nl/en-gb/news/2018/05/the-netherlands-leads-europe-in-internet-access (accessed September 2019).

Committee on Lesbian, Gay, Bisexual, and Transgender Health Issues and Research Gaps and Opportunities. 2011. "*The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding. Institute of Medicine (US)*." Washington (DC): National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK64802/.

Conrad, F., M. Couper, R. Tourangeau, and A. Peytchev. 2010. "The impact of progress indicators on task completion." *Interacting with Computers* 22: 417–427. DOI: http://doi.org/10.1016/j.intcom.2010.03.001.

Cortina, C. and P. Festy. 2014. "Identification of same-sex couples and families in censuses, registers and surveys." *FamiliesAndSocieties Working Papers Series 8*. Available at: http://www.familiesandsocieties.eu/wp-content/uploads/2014/12/WP8CortinaFesty2014.pdf (accessed September 2019).

De Graaf, P., M. Kalmijn, G. Kraaykamp, and C. Monden. 2010. "*Design and content of the NEtherlands Longitudinal Lifecourse Study (NELLS).*" Netherlands: Tilburg University & Radboud University Nijmegen.

DeMaio, T., N. Bates, and M. O'Connell. 2013. "Exploring measurement error issues in reporting same-sex couples." *Public Opinion Quarterly* 77(Special Issue): 45–158. DOI: http://doi.org/10.1093/poq/nfs066.

Dewaele, A., M. Caen, and A. Buysse. 2014. "Comparing survey and sampling methods for reaching sexual minority individuals in Flanders." *Journal of Official Statistics* 30(2): 251–275. DOI: http://doi.org/10.2478/jos-2014-0016.

Dillman, D., J. Smyth, and L. Christian. 2009. "*Internet, mail, and mixed-mode surveys: The tailored design method (3rd ed.).*" New York: John Wiley & Sons.

ESS, European Social Survey Cumulative File, ESS 1-8. 2018. Data file edition 1.0. NSD - Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC. DOI: http://doi.org/10.21338/NSD-ESS-CUMULATIVE.

European Union Agency for Fundamental Rights. 2012. "EU LGBT survey – Technical report Methodology, online survey, questionnaire and sample." *Research report*. Available at: http://fra.europa.eu/en/survey/2012/eu-lgbt-survey (accessed September 2019).

Evans, A., R. Wiggins, G. Bolding, and J. Elford. 2008. "Characteristics of gay and bisexual men who drop out of a web survey of sexual behaviour in the UK." *AIDS & Behavior* 12(6): 957–963. DOI: http://doi.org/10.1007/s10461-008-9400-6.

Fischer, M. 2016. "Identifying same-sex couples in cross-national survey data: A comparison of same-sex couples' demographic and socio-economic traits in six European countries." In *Nederland in context: verschillen en overeenkomsten*, edited by R. Meuleman, G. Kraaykamp, and M. Wittenberg, 50–79. Den Haag: DANS.

Fischer, M. 2019. "*Free to Live their Lives as they Wish? The Social Well-Being Gap between Persons in Same-Sex and Mixed-Sex Relationships in Europe*" [Doctoral dissertation]. University of Amsterdam: The Netherlands. ISBN 978-94-92679-72-7.

Fischer, M. and S. Steinmetz. 2018. "Designing a Sampling Frame for Same-Sex Couples and Families: The Unions in Context Study" In *SAGE Research Methods*. Cases (Vol. 2). London: SAGE.

Fischer, M., M. Kalmijn, and S. Steinmetz. 2017. "Design, content and questionnaire of the UNICON study: Lesbian, gay and heterosexual couples and families in the Netherlands." Research report. University of Amsterdam, the Netherlands. Available at: https://dare.uva.nl/search?identifier=d7a835f9-7c1b-448b-8047-a9779f5710db (accessed September 2019).

Galupo, M., E. Lomash, and R. Mitchell. 2016. "All of my lovers fit into this scale: sexual minority individuals' responses to two novel measures of sexual orientation." *Journal of Homosexuality* 64: 145–165. DOI: http://doi.org/10.1080/00918369.2016.1174027.

Gates, G. 2013. "LGBT Parenting in the United States." Los Angeles, CA: Williams Institute, UCLA School of Law. Available at: http://williamsinstitute.law.ucla.edu/wp-content/uploads/LGBT-Parenting.pdf.

GGS Waves 1, and 2. DOI: http://doi.org/10.17026/dans-z5z-xn8g, and http://doi.org/10.17026/dans-xm6-a262.

Goldberg, A. and K. Allen. 2007. "Lesbian mothers' ideas and intentions about male involvement across the transition to parenthood." *Journal of Marriage and Family* 69: 352–365. DOI: http://doi.org/10.1111/j.1741-3737.2007.00370.x.

Gribble, J., H. Miller, S. Rogers, and C. Turner. 1999. "Interview mode and measurement of sexual behaviors: Methodological issues." *Journal of Sex Research* 36(1): 16–24. DOI: https://doi.org/10.1080/00224499909551963.

Haseldon, L. and T. Joloza. 2009. "*Measuring sexual identity: a guide for researchers*." Newport (UK): Office for National Statistics. Available at: https://www.ons.gov.uk/methodology/classificationsandstandards/sexualidentityguidanceandprojectdocumentation (accessed November 2019).

Herek, G., A. Norton, T. Allen, and C. Sims. 2010. "Demographic, psychological, and social characteristics of self-identified lesbian, gay, and bisexual adults in a US probability sample." *Sexuality Research and Social Policy* 7(3): 176–200. DOI: http://doi.org/10.1007/s13178-010- 0017-y.

Hicks, S. 2005. "Is gay parenting bad for kids? Responding to the 'very idea of difference' in research on lesbian and gay parents." *Sexualities* 8: 153–168. DOI: http://doi.org/10.1177/1363460705050852.

Institute of Medicine (US) Committee on Lesbian, Gay, Bisexual, and Transgender Health Issues and Research Gaps and Opportunities. 2011. "*The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*." Washington (DC): National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK64802/.

Kenyon, E. and S. Heath. 2010. "Choosing this life: narratives of choice amongst house sharers." *Housing Studies* 16(4): 619–635. DOI: http://doi.org/10.1080/02673030120080080.

Kinsey, A., W. Pomeroy, and C. Martin. 1948. "*Sexual Behaviour in the Human Male*." Philadelphia: Saunders.

Kuyper, L. 2016. "LHBT-monitor 2016: opvattingen over en ervaringen van lesbische, homoseksuele, biseksuele en transgender personen." Den Haag: Sociaal en Cultureel Planbureau.

Laumann, E., J. Gagnon, R. Michael, and S. Michaels. 1994. "*The social organization of sexuality: Sexual practices in the United States*." Chicago: The University of Chicago Press.

Lee, R. 1993. "*Doing research on sensitive topics*." London: Sage Publications.

Lozar Manfreda, K., M. Bosnjak, J. Berzelak, I. Haas, and V. Vehovar. 2008. "Web surveys versus other survey modes: A meta-analysis comparing response rates." *International Journal of Market Research* 50(1): 79–104. DOI: http://doi.org/10.1177/147078530805000107.

Meyer, I. and M. Colten. 1999. "Sampling gay men: Random Digit Dialing versus sources in the gay community." *Journal of Homosexuality* 37(4): 99–110. DOI: http://doi.org/10.1300/J082v37n04_06.

Meyer, I. and P. Wilson. 2009. "Sampling lesbian, gay, and bisexual populations." *Journal of Counseling Psychology* 56(1): 23–31. DOI: http://doi.org/10.1037/a0014587.

Muñoz-Leiva, F., J. Sánchez-Fernández, F. Montoro-Ríos, and J. Ibáñez-Zapata. 2010. "Improving the response rate and quality in Web-based surveys through the

personalization and frequency of reminder mailings." *Quality & Quantity* 44(5): 1037–1052. DOI: http://doi.org/10.1007/s11135-009-9256-5.

OECD. 2019. "*Society at a Glance 2019: OECD Social Indicators*." Paris: OECD Publishing. DOI: https://doi.org/10.1787/soc_glance-2019-en.

Patterson, C. 1992. "Children of lesbian and gay parents." *Child Development* 63: 1025–1042. DOI: http://doi.org/10.2307/1131517.

Régnier-Loilier, A. 2018. "Are the Generations and Gender Surveys well suited for studying same-sex couples?" *European Journal of Population* 34(4): 567–578. DOI: http://doi.org/10.1007/s10680-017-9440-6.

Rosser, B., M. Miner, W. Bockting, M. Ross, J. Konstan, L. Gurak, J. Stanton, W. Edwards, S. Jacoby, A. Carballo-Dieguez, R. Mazin, and E. Coleman. 2009. "HIV risk and the Internet: Results of the Men's Internet Sex (MINTS) Study." *AIDS & Behavior* 13(4): 746–756. DOI: http://doi.org/10.1007/s10461-008-9399-8.

Ross, M., B. Rosser, and J. Stanton. 2004. "Beliefs about cybersex and Internet-mediated sex of Latino men who have Internet sex with men: Relationships with sexual practices in cybersex and in real life." *AIDS Care* 16(8): 1002–1011. DOI: http://doi.org/10.1080/09540120412331292444.

Rothblum, E. 2007. "From science fiction to computer-generated technology: Sampling lesbian, gay, and bisexual individuals." In *The health of sexual minorities: Public health perspectives on lesbian, gay, bisexual and transgender populations*, edited by I. Meyer and M. Northridge, 441–454. New York: Springer.

Schwartz, C. and N. Graf. 2009. "Assortative matching among same-sex and different-sex couples in the United States, 1990–2000." *Demographic Research* 21: 843–878. DOI: http://doi.org/10.4054/DemRes.2009.21.28.

Sell, R. and C. Petrulio. 1996. "Sampling homosexuals, bisexuals, gays, and lesbians for public health research: A review of the literature from 1990 to 1992." *Journal of Homosexuality* 30(4): 31–47. DOI: http://doi.org/10.1300/J082v30n04_02.

Smyth, J. and J. Pearson. 2011. "Internet survey methods: A review of strengths, weaknesses, and innovations." In *Social and behavioral research and the internet. Advances in applied methods and research strategies*, edited by M. Das, P. Ester, and L. Kaczmirek, 11–44. New York, NY: Routledge.

Solarz, A. (ed.). 1999. "*Lesbian health: current assessment and directions for the future*." Washington, DC: National Academy Press.

Sudman, S., M. Sirken, and C. Cowan. 1988. "Sampling rare and elusive populations." *Science* 240: 991–996. DOI: https://doi.org/10.1126/science.240.4855.991.

Spira, A., N. Bajos, and the ACSF Group. 1994. "*Sexual Behaviour and AIDS*." Aldershot, England: Avebury.

Statistics Netherlands. 2016. "Veileigheidsmonitor 2015." Research report. Den Haag, the Netherlands. Available at: https://download.cbs.nl/pdf/veiligheidsmonitor-2015.pdf (accessed November 2019).

Statistics Netherlands. 2018. "*Population dynamics; birth, death and migration per region*." Available at: https://opendata.cbs.nl/statline/#/CBS/en/dataset/37259eng/table?dl=1310C (accessed November 2019).

Tolsma, J., G.L.M. Kraaykamp, P.M. de Graaf, M. Kalmijn, C.M. Monden. 2014. Netherlands Longitudinal Lifecourse Study - NELLS Panel Wave 1 2009 and Wave 2 2013 - versie 1.2. DANS. DOI: https://doi.org/10.17026/dans-25n-2xjv.

Tourangeau, R. and T. Yan. 2007. "Sensitive questions in surveys." *Psychological Bulletin* 133(5): 859–883. DOI: http://doi.org/0.1037/0033-2909.133.5.859.

Tourangeau, R. 2014. "Defining hard to survey populations." In *Hard to Survey Populations*, edited by B. Edwards, R. Tourangeau, T. Johnson, K. Wolter, and N. Bates, 3–21. Cambridge University Press.

Shih, T.-H. and X. Fan. 2008. "Comparing response rates from Web and mail surveys: A metaanalysis." *Field Methods* 20(3): 249–271. DOI: http://doi.org/10.1177/1525822X08317085.

Umberson, D., M. Thomeer, P. Kroeger, A. Lodge, and M. Xu. 2015. "Challenges and opportunities for research on same-sex relationships." *Journal of Marriage and Family* 77(1): 96–111. DOI: http://doi.org/10.1111/jomf.12155.

UNESCO. 2012. "*International Standard Classification of Education ISCED 2011*." UNESCO Institute for Statistics, Canada. Available at: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf (accessed September 2019).

Wellings, K., J. Field, A. Johnson, and J. Wadsworth. 1994. "*Sexual Behaviour in Britain: The National Survey of Sexual Attitudes and Lifestyles*." London: Penguin Books.

Wolff, M., B. Wells, C. Ventura-DiPersia, A. Renson, and C. Grov. 2017. "Measuring sexual orientation: a review and critique of U.S. data collection efforts and implications for health policy." *The Journal of Sex Research* 54: 507–553. DOI: http://doi.org/10.1080/00224499.2016.1255872.

Valfort, M.-A. 2017. "LGBTI in OECD Countries: A Review." *OECD Social, Employment and Migration Working Papers*, OECD Publishing, Paris. Available at: https://doi.org/10.1787/d5d49711-en (accessed June 2019).

Villarroel, M., C. Turner, E. Eggleston, A. Al-Tayyib, S. Rogers, A. Roman, P. Cooley, and H. Gordek. 2006. "Same-gender sex in the United States: Impact of T-ACASI on prevalence estimates." *Public Opinion Quarterly* 70(2): 166–196. DOI: http://doi.org/10.1093/poq/nfj023.

Received June 2018
Revised June 2019
Accepted September 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 777–805, http://dx.doi.org/10.2478/JOS-2019-0033

Comparing Self-Reported and Partnership-Inferred Sexual Orientation in Household Surveys

Simon Kühne[1], *Martin Kroh*[1], *and David Richter*[2]

Research comparing heterosexuals with bisexuals and homosexuals in economics and the social sciences typically relies on two strategies to identify sexual orientation in existing survey data of general populations. Probing respondents to self-report their sexual orientation is generally considered the preferred option. Since self-reports are unavailable in most large multidisciplinary surveys, often researchers infer sexual orientation from the gender-constellation of a respondent's partnership instead. Based on German Socio-Economic Panel (SOEP) data, this article reviews both strategies empirically in the context of a household panel survey. The analysis shows that self-reported and partnership-inferred sexual orientation are not mutual substitutes, instead leading to substantively different conclusions about differences between heterosexuals and LGBs (Lesbian, Gays, and Bisexuals). The article discusses problems of non-coverage in partnership-inferred sexual orientation and also investigates measurement error in self-reported sexual orientation, finding notable mode and interviewer effects.

Key words: Surveys; sexual orientation; measurement error; interviewer effects; survey methodology.

1. Motivation

The concept of sexual orientation received not only increasing public, but also academic attention in the past decades. While early scholarly proponents of research on sexual orientation often came from clinical psychology, public health research, and social psychology, the concept has been increasingly adopted by quantitatively oriented scholars from economics and social scientists alike. Sexual orientation is thereby acknowledged to be an important dimension of inequality, structuring societies at large and affecting individuals' lives comprehensively similar to the inequality dimensions of race, gender identity, age, and class. The emergence of the academic field of *Queer Studies* reflects this view that the concept of sexual orientation is multidisciplinary in nature. The cross-cutting nature of the concept as well as its relevance for various disciplines constitute the need for a measurement to become an established part of questionnaires in multidisciplinary surveys, similar to the other inequality dimensions.

The concept of sexual orientation is conceived of by many scholars as a durable sexual attraction to either the different sex (heterosexuals), the same sex (homosexuals), both

[1] Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany. Emails: simon.kuehne@uni-bielefeld.de and martin.kroh@uni-bielefeld.de
[2] Socio-Economic Panel Study (SOEP), DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany. Email: drichter@diw.de
Acknowledgments: We would like to thank the Guest Editors Stephanie Steinmetz, Mirjam Fischer, and Nancy Bates, as well as the anonymous reviewers for their valuable comments and time and effort devoted to improving the quality of the article.

sexes (bisexuals), or none of them (for an overview of the debates on the conceptualization of sexual orientation, see, for example, Williams Institute 2009; Haseldon and Joloza 2009; Moliero and Pinto 2015). The term sexual orientation is thereby used rather broadly compared to the more specific concepts of sexual attraction ("I feel attracted to. . ."), sexual behavior ("I had sexual experience with . . ."), and sexual identity ("I am heterosexual, homosexual, . . ."). For instance, Geary et al. (2018) describe sexual attraction, behavior, and identity as subdimensions constituting sexual orientation (see also Chandra et al. 2013; Galupo et al. 2016; Wolff et al. 2016). They show that these facets of sexual orientation do not fully overlap empirically, nor do they serve the same research interests: public health research may be interested more strongly in aspects of sexual behavior, clinical psychology may focus more on emotional attraction, and economics and social sciences more on the social identity aspect of sexual orientation.

While topical surveys often include multiple-item measures of different facts of sexual orientation in their questionnaires, the paucity of any survey items on sexual orientation in multidisciplinary surveys is astonishing. An inquiry of codebooks of the Inter-university Consortium for Political and Social Research (ICPSR) data repository suggests that in only 90 out of the 10,443 archived surveys, that is, less than 1%, do study descriptions include the key words "sexual orientation". Although alternative search terms, such as "lesbian" (525 studies), "gay" (1,873 studies), "bisexual" (356 studies), and "homosexual" (912 studies) lead to higher incidences, these studies often include items on the acceptance of sexual minorities in the general population, rather than identifying respondents by sexual orientation. Browsing these lists suggests that the larger the sample size and the broader the scope of surveys thematically, the less likely that they include measures of sexual orientation. Hence, researchers interested in utilizing existing social surveys, election studies, and census data for research on sexual orientation will either evade to smaller topical studies or they will draw on alternative strategies to identify sexual orientation in respondents.

One popular alternative strategy to directly probing sexual orientation is to indirectly infer an indicator of sexual orientation in existing multidisciplinary surveys from the reported gender-constellation of respondent's current and previous partnership(s). For instance, in household surveys (e.g., the Panel Study of Income Dynamics (PSID), the Survey of Health, Ageing and Retirement in Europe (SHARE), the Socio-Economic Panel (SOEP) – utilized in the present article – and Census Studies of the United States, as well as many other countries), researchers can draw on self-reported gender and the self-reported relationship between two respondents of a household in order to identify same-sex and opposite-sex couples. Although the prevalence of empirical studies on sexual orientation using partnership-inferred measures of sexual orientation clearly varies across disciplines and is low in public health research and queer studies in general, it represents the predominant empirical strategy in other fields. Highly valuable research on poverty rates in sexual minorities, occupational segregation, and the pay gap between heterosexuals and homosexuals relies in large parts on partnership-inferred sexual orientation based on Census data and data from the General Social Survey, for instance (see an overview, Klawitter 2014).

As survey-based research comparing heterosexuals with LGBs is often constrained to using gender constellations of partnership as a substitute for direct measures of sexual

orientation, findings rely on the assumption of the functional equivalence of alternative ways to measuring sexual orientation. This article aims at comparing both strategies of measuring sexual orientation in survey-based research: a) self-reported sexual identity, and b) a partnership-inferred indicator of sexual orientation. For our analysis, we rely on Socio-Economic Panel Study data (SOEP), see Goebel et al. (2018). The SOEP is a nationally representative longitudinal survey of about 15,000 private households in Germany with annual interviews since 1984. All members of the selected households aged 18 years and older are asked to participate in annual interviews. The SOEP consists of multiple subsamples in order to maintain a reasonable panel size and representativeness of the German population over time. In addition to general population subsamples, some subgroups are specifically boosted, including migrants (samples B and M) and families (sample L), see Kroh et al. (2018). Moreover, new household members (e.g., new partners or grown-up children) are invited to join the study. To minimize attrition, individuals are followed even if a household splits or moves. On average, a respondent in wave 2016 already participated for 11 years in the SOEP (Min: 1, Max: 33, Median: 7).

The household structure of the SOEP permits studying partnership-inferred sexual orientation from its first year of interviewing in 1984. Moreover, in 2016, self-reported sexual orientation was surveyed for the first time by SOEP. More specifically, SOEP respondents report their sexual identity by categorizing themselves as heterosexuals, homosexuals, or bisexuals.

Before comparing both strategies of measuring sexual orientations in Section 4, we review the advantages and disadvantages of the two approaches separately. Section 2 discusses self-reported sexual orientation with a focus on measurement error issues. More precisely, we investigate mode and interviewer effects on self-reported sexual orientation and item nonresponse. Section 3 examines partnership-inferred sexual orientation and discusses potential measurement errors such as misclassification. Section 4 firstly investigates the extent to which the two strategies generate the same classification of respondents with respect to sexual orientation. Secondly, we examine differences in the sub-populations classified by each method across a variety of socio-demographics and traits. Finally, Section 5 summarizes the results and offers practical implications for researchers conducting or relying on surveys.

2. Self-Reported Sexual Orientation

Different items have been developed in order to measure sexual orientation by self-reporting (Sell 2007; Gates 2011; Wolff et al. 2016). In the following paragraphs, we will focus on the social identity dimension of sexual orientation. Reviewing the existing literature, the most common approach to obtain self-reports on the identity facet of sexual orientation in (large scale) surveys is to outright ask respondents whether they identify themselves as heterosexual, homosexual, or bisexual (e.g., Wooden 2014; Uhrig 2014).

In the 2016 wave, for the first time, respondents of the SOEP were asked about their sexual orientation. The question is worded, “In the context of relationships, the question of sexual orientation arises. Would you describe yourself as...?” The available answers were “Heterosexual or straight (that is, attracted to the opposite sex),” “Homosexual (gay or lesbian, that is, attracted to the same sex),” “Bisexual (attracted to both sexes),” “Other,”

Table 1. Distribution of responses to the question on self-reported sexual orientation.

Response	Percent (unweighted)
Heterosexual	86.11
Homosexual	0.65
Bisexual	0.54
Other	7.10
Prefer not to say	4.41
No answer (nonresponse)	1.19

n = 24,287.

"Prefer not to say" and "No answer (nonresponse)". A similarly worded question is used by the UK Understanding Society Study (Booker et al. 2017) and the Australian Household, Income and Labour Dynamics Survey (HILDA, Wooden 2014). The explanations in parentheses were added, as earlier pretesting showed that the terms homosexual and heterosexual alone caused misunderstandings among some respondents. Table 1 displays the distribution of responses.

About 1.2% of respondents report identifying as either homosexual or bisexual. Applying cross-sectional weights, which compensate for the disproportionate sampling, as well as nonresponse (Kroh et al. 2018), we estimate that the share of adults in Germany who openly identify as lesbian, gay, or bisexual is 1.9%.

The high number of answers avoiding self-categorization on the dimension of sexual orientation reported in Table 1 shows that, in line with previous research, the strategy of directly probing sexual orientation is not without problems.

A number of previous studies analyzing measurement error in questions about sexual orientation, for instance, showed that a particularly high number of respondents chooses to not answer the question at all (*item nonresponse*), either by outright refusing to answer or by selecting the answer "Other" (e.g., Jans et al. 2015). This response behavior may reflect uncrystallised views on the self-identification, with sexual orientation on the one hand, and the perception of information on sexual orientation as sensitive on the other hand. Incorrect information may be provided intentionally by respondents in order to meet presumed societal expectations (*social desirability bias*, see Krumpal (2013) for an overview) and to obtain social approval (from the interviewer). Moreover, some interviewers may try to avoid a seemingly awkward interview situation by pre-quoting the item nonresponse option. Both social desirability bias and interviewer behavior point to the importance of contextual factors of the interview situation (e.g., Tourangeau and Yan 2007; Kühne 2018; Hilgert et al. 2016).

In the upcoming section, thus, we investigate possible effects of the mode of data collection, and the interviewer, on response behavior to the direct question on sexual orientation in the SOEP. Our analysis aims to provide guidance for researchers who plan to collect data on self-reported sexual orientation of respondents in surveys.

2.1. *Mode Effects*

An important choice that researchers make in designing a survey is the mode of data collection. As a multi-mode survey, SOEP allows testing of the effects of interviewer-

administered interviewing versus self-administered interviewing in surveying self-reported sexual orientation.

In the SOEP, data collection is largely based on personal, face-to-face interviews. Since 1998, the SOEP has been gradually replacing paper and pencil interviewing (PAPI) with computer-assisted personal interviewing (CAPI) as the predominant mode of data collection. For instance, in the 2016 wave, 72% of the more than 24,000 respondents who were asked for their sexual orientation were interviewed by an interviewer face-to-face, with 90% of those interviews conducted via CAPI. The remaining 28% of respondents in the 2016 wave used a self-administered mode based on a printed or digital version of the questionnaire and without an interviewer present.

Numerous studies have shown that the mode of data collection is decisive for measurement error in survey questions (e.g., Dillman et al. 2009; Schwarz et al. 1991). For instance, item nonresponse rates tend to be lower in interviewer-administered surveys. There are three main reasons for this. First, in cases where respondents have difficulties in understanding a question, interviewers can clarify questions and answer options, thus helping respondents to provide a valid answer. Second, interviewers may actively probe and argue in order to obtain valid responses (Kuha et al. 2014). Third, many respondents likely perceive that not providing an answer is an undesired behavior in front of the interviewer because it runs counter to the main purpose of the interview of collecting valid information.

These advantages of face-to-face interviewing suggest that respondents are more likely to provide responses in interviewer-administered interviewing on the one hand. However, on the other hand, the privacy of the self-interview may have positive effects on measurement, as respondents may feel more comfortable divulging sensitive information on sexual orientation.

Table 2 displays responses to the sexual orientation question across modes. 4.41% of respondents refused to answer the question by actively stating “Prefer not to say”. With 5.40%, the share is higher in the Mail/CAWI (Computer-Assisted Web Interview) mode compared to interviewer-administered interviews (4.02%). Moreover, in the Mail/CAWI mode, 3.29% did not provide any answer at all (nonresponse). Missing information in the CAPI mode amounts to 0.38% only – most likely due to the fact that the interviewers manage the CAPI system and aim at low nonresponse rates. In total, 5.60% of all respondents choose to not provide a valid answer, with 8.69% in Mail/CAWI and 4.40% in the CAPI mode. Item nonresponse rates in nearby questions in the questionnaire are much lower, suggesting that the observed response behavior is specific to the question on sexual orientation and does not reflect a general tendency of respondents to not provide valid answers (e.g., as a type of satisficing behavior, see Krosnick et al. 1996). The observed item nonresponse rate is quite high compared to other existing studies surveying sexual orientation (3.4% [Wooden 2014], 3.2% [Uhrig 2014], 2.78% [Frederiksen-Goldsen and Kim 2014], 0.93% [Dahlhamer et al. 2014]).

The presence of an interviewer may not only affect the propensity of item nonresponse, but also the selection of valid answer options. In fact, socially desirable response behavior and impression management in interviewees tend to be more prevalent when interviewers (and others) are present (Krumpal 2013). Social desirability bias is shown to be most prevalent in questions that are perceived to be sensitive and stigmatizing (e.g., Tourangeau

Table 2. Distribution of self-reported sexual orientation across survey modes.

Response	F2F %	Mail/CAWI %	Total %
Heterosexual	87.46	82.60	86.11
Homosexual	0.56	0.88	0.65
Bisexual	0.36	1.02	0.54
Other	7.21	6.81	7.10
Prefer not to say	4.02	5.40	4.41
No answer (nonresponse)	0.38	3.29	1.19

In percent. Unweighted. n = 24,287.

and Yan 2007), such as sexual orientation for lesbian, gay, or bisexual respondents. Some may experience difficulties in disclosing their sexual orientation to another person, particularly in cases where (parts of) the social environment, such as other household members, neighbors, and colleagues are unaware of the respondent's sexual orientation. Hence, we expected a lower share of respondents reporting an LGB sexual orientation in the interviewer-administered mode than in the self-administered mode, which is perceived as more anonymous. The results in Table 2 match our expectations, with 1.90% reporting a homosexual or bisexual orientation in the self-administered, compared to 0.92% in the interviewer-administered survey mode.

Related to this, an unexpected result is the large share of respondents choosing the answer category "Other" (7.10%). We initially implemented this category for respondents that do not identify as either heterosexual, homosexual, or bisexual, but prefer other forms of sexual self-categorization, such as asexual, pansexual, and queer. While there is no reliable population estimate of the number of, for example, persons identifying as asexual and pansexual available for Germany, Aicken et al. (2013) estimate the prevalence of asexuals in the UK at 0.4%. This is supported by results of the SOEP Innovation Sample in 2015 (see Richter and Schupp 2015) in which only 0.16% of over 5,000 respondents reported an asexual identity.

Hence, it is unlikely that the 7.10% of respondents who choose the "Other" category represent asexuals, for instance. Although follow-up write-ins for those who report "Other" have not been collected in the main survey, we surveyed this information in 2006 in a pretest of 1,057 respondents. Here, only eight respondents provided write-ins and none of them reported a queer sexual orientation, such as asexual or pansexual; rather, on the contrary, seven mentioned variations of "normal" and a single respondent wrote "I don't know these words". Hence, we interpret the high number of respondents reporting the "Other" category in most cases as an unanticipated form of item nonresponse to the sexual orientation question (including 'refusal', 'don't know', and 'does not apply').

In Table 3, we report the results of multilevel logistic regression models with respondents nested in households and households nested in interviewers assessing mode effects on response behavior. Although interviewers do not actively collect information in the self-administered mode, they may be involved, nonetheless, in the process of contacting these households and thus promoting participation. Thus, we use the multilevel data structure with respondents nested in interviewers even for households with self-administered interviews. Hierarchical regression modelling allows to address the

Table 3. Mode effects on response propensity, reports of homosexual and bisexual orientation, and reporting "other".

	Response (yes/no)		Homosexual/bisexual | response = yes		Reporting "other" | response = yes	
Variable	OR	95% CI	OR	95% CI	OR	95% CI
Mode						
Face to face (Ref.)	–	–	–	–	–	–
Mail/CAWI	0.38***	0.29– 0.49	4.60***	2.37–8.89	2.79***	1.64–4.72
Third person present						
No third person (ref.)	–	–	–	–	–	–
Partner	0.76*	0.59–0.97	0.77	0.38–1.56	1.27	0.82–1.96
Other person	0.71*	0.51–0.98	2.83**	1.34–5.96	1.51	0.82–2.78
No information	0.41***	0.30–0.56	2.92*	1.25–6.84	3.69***	2.10–6.48
Respondent panel experience						
1–2 years (ref.)	–	–	–	–	–	–
3–5 years	1.84**	1.24–2.73	0.87	0.40–1.88	0.71	0.33–1.57
6–10 years	2.35***	1.52–3.63	0.83	0.37–1.83	0.90	0.39–2.05
11+ years	1.12	0.72–1.75	0.66	0.27–1.64	2.40	0.99–5.78
Subsample						
Cross–sectional (ref.)	–	–	–	–	–	–
Migrants	0.70	0.48–1.00	0.18**	0.07–0.48	1.03	0.49–2.18
Families	1.41	0.96–1.07	0.81	0.39–1.66	2.29*	1.14–4.57
Other	0.92	0.66–1.28	0.32*	0.12–0.83	0.62	0.32–1.20
Respondent gender						
Male (ref.)	–	–	–	–	–	–
Female	0.80**	0.68–0.93	1.30	0.87–1.94	1.54***	1.20–1.97
Respondent age						
<29 (ref.)	–	–	–	–	–	–
30–39	1.33	0.91–1.94	0.38**	0.18–0.78	0.96	0.48–1.90
40–49	0.88	0.63–1.25	0.28***	0.14–0.57	1.44	0.77–2.69

Table 3. Continued.

Variable	Response (yes/no)		Homosexual/bisexual | response = yes		Reporting "other" | response = yes	
	OR	95% CI	OR	95% CI	OR	95% CI
50–59	0.55**	0.40–0.78	0.12***	0.06–0.30	3.09***	1.64–5.79
60–69	0.43***	0.29–0.62	0.06***	0.02–0.17	14.41***	6.88–30.17
70–79	0.28***	0.19–0.42	0.03***	0.01–0.12	51.17***	22.28–117.57
80+	0.18***	0.11–0.29	0.01***	0.00–0.04	67.48***	25.15–181.08
In a relationship?						
No (ref.)	–	–	–	–	–	–
Yes	2.89***	2.26–3.65	0.87	0.52–1.46	0.28***	0.18–0.43
Household size						
Single (ref.)	–	–	–	–	–	–
2	1.18	0.82–1.66	0.53	0.25–1.13	1.27	0.61–2.64
3	1.12	0.76–1.64	0.24***	0.06–0.35	4.24***	1.93–9.32
4	0.88	0.61–1.30	0.05***	0.02–0.14	3.97**	1.78–8.86
5	0.74	0.48–1.13	0.05***	0.02–0.16	2.60*	1.07–6.32
6+	0.70	0.44–1.10	0.05***	0.02–0.20	1.80	0.70–4.67
German region						
North (ref.)	–	–	–	–	–	–
East	0.57*	0.32–0.99	1.09	0.44–2.69	0.51	0.14–1.82
South	0.55*	0.32–0.96	0.36*	0.14–0.93	0.68	0.19–2.40
West	062	0.36–1.05	0.77	0.33–1.82	0.63	0.19–2.14
Constant	0.00***	0.00–0.01	0.00***	0.00–0.01	0.00***	0.00–0.00
n_{resp}	24,287		22,928		22,928	

*p < .05, **p < .01, ***p < .001.

Multilevel logistic regression with respondents nested in households and households nested in interviewers.

See Table A1 in the Appendix (Section 6) for a comparison of characteristics across modes.

hierarchical nature of the survey data at the level of respondents due to specifics of the sampling and survey design, and thereby improves the estimation of coefficients and their standard errors. We estimate mode effects on three outcomes independently. First, whether there is a valid response at all (yes/no) and, second, whether a homosexual or bisexual orientation is reported (yes/no), conditional on response, and third, whether "Other" is reported (yes/no), conditional on response. Please note that in the SOEP, the assigned mode is not randomly allocated across households and respondents. Thus, we add a number of potential confounding factors to each model. The most influencing factor of allocating mode is the predominant field work routine in a subsample of SOEP. For instance, from wave 2011 onwards, all refreshment and enlargement samples are exclusively interviewed in the CAPI mode. We control for predominant field work routine by adding a subsample identifier into the analysis. Other factors that correlate with mode in the SOEP are respondent's gender, age, and years in the panel. Moreover, we control for correlating household characteristics including whether another person was present in an interview, household size and geographic region in Germany. Finally, we control for current relationship status (yes/no). As an additional robustness check, we replicated our analyses for each subsample and separated for paper-and-pencil and computer-assisted mode. No substantial differences are observed.

The first model estimates the effect of mode on response propensity (response versus item nonresponse). Compared to the face-to-face mode (reference), and in line with the results in Table 2, the odds of providing a valid response in the self-administered mode are less than half of the odds in the interviewer-administered mode (Odds Ratio (OR) = 0.38, $p < .001$). In addition, there are a number of other interesting results when turning to the respondent level characteristics. We find, for instance, lower response propensities for females and older respondents. Moreover, those currently in a relationship are more likely to respond. A possible explanation could be that some respondents understand the survey question as only relating to sexual orientation in a currently ongoing relationship, rather than general sexual orientation. This seems plausible, as the question was placed right after questions about family and relationship status.

In the second model, we estimate mode effects on the probability of reporting a homosexual or bisexual orientation given a valid response. Respondents participating via Mail/CAWI are associated with a 4.60 times higher chance (odds) of reporting a homosexual or bisexual orientation ($p < .001$) rather than a heterosexual orientation or answering "Other". Again, the results match expectations for reporting a homosexual or bisexual orientation; respondents are much more likely to report sensitive and potentially undesired responses in the self-administered – and likely more private – interview mode. In addition, effects of respondent-level characteristics match expectations. The propensity of reporting a homosexual or bisexual orientation decreases dramatically with the increasing age of respondents. This is plausible, as older cohorts are less likely to openly identify and live as non-heterosexual. Moreover, the larger the household size, the less likely it is to obtain a non-heterosexual response. This is most likely due to the fact that LGBs more often live in smaller households compared to heterosexual households. Note that our estimation is associated with comparatively large uncertainty, as only few respondents report a non-heterosexual relationship; as a consequence, the 95% confidence interval for the estimated odds ratio is comparatively wide.

In the third model, we estimate mode effects on respondents selecting the "Other" category as an answer. The odds ratio of respondents using the "Other" option in self-administered versus interviewer-administered mode is estimated at 2.79 ($p < .001$). Hence, in persons who do not explicitly refuse an answer on sexual orientation, it is much more likely that they choose "Other" when not being interviewed by an interviewer in person. Turning to respondent level characteristics, and matching the results of Model 1, female and older respondents are more likely to respond "Other".

The latter two models reported in Table 3 on homosexual and bisexual responses as well as the "Other" response option consider the first stage of the answering process, that is, refusal or response, to be uninformative for the second stage. It may very well be, however, that both stages of the response process are correlated. Heckman selection models allow modelling correlated residuals at stages one and two (Heckman 1976, 1979, Puhani 2000). Estimating Heckman models (probit, not reported in the form of a table) suggest that the first stage of response versus refusal and the second stage of LGB response are only moderately correlated and thus, estimates of Table 3 are hardly affected by a change in model specification. However, the first stage of response versus refusal and the second stage "Other" option are positively and strongly correlated, suggesting that some respondents use the "Other" option as a substitute for directing refusing a response. Also, the Heckman specification suggests that the effect of self-administered interviewing turns negative for the "Other" option; most other estimated remain largely unchanged. The change of the direction of the mode effect suggests that interviewer-administered interviews increase refusals both in the form of direct refusals and in the form of the hidden refusal using the "Other" option.

To sum up, observing a valid response is more likely in face-to-face interviewing than in self-administered modes. However, more respondents are willing to share a non-heterosexual orientation in the self-administered mode compared to interviewer-administered interviews. The latter inflates the propensity of a hidden item nonresponse using the "Other" response option. This has implications for measuring sexual orientation in surveys. Our results suggest a strategy of relying on face-to-face interviewing in order to minimise item nonresponse, but to switch to computer-assisted self-interviewing approach (CASI) as a more private mode of interviewer-administered data collection.

2.2. *Interviewer Effects*

Some interviewers may generally be more efficient in obtaining valid responses than others. Thus, in addition to mode effects, we also investigate interviewer effects (see West and Blom 2017 for an overview) on the prevalence of item nonresponse and the selection of the "Other" category. As there are only 290 respondents that report a non-heterosexual orientation, and these cases are often clustered within households and interviewers, we did not investigate potential interviewer effects (intra-interviewer clustering) in LGB responses.

Consequently, in this section, we only analyze interviewer-administered interviews, which reduces the number of observations by about one third. On average, each SOEP interviewer conducts 68 interviews in wave 2016 (Median $= 59$, Min $= 1$, Max $= 207$). Item nonresponse rates in the question on sexual orientation vary greatly across interviewers. For the upcoming analysis, we only analyze interviewers that conducted at least five personal interviews (416 interviewers). On average, an interviewer obtains 6%

non-valid responses (Median = 0.00, SD = 13.74). While 59% of the interviewers do not collect any invalid information for the question about sexual orientation at all, many of the interviewer staff did, showing large variation with up to 100% item nonresponse rates (two interviewers; $p_{75} = 4.30\%$; $p_{90} = 16.67\%$). Turning to participants reporting "Other" as an answer, the interviewer-average amounts to 9% (Median = 0.00, SD = 21.74). Again, the majority of interviewers did not collect any "Other" answer (64%). However, there are 13 interviewers that exclusively (100%) obtained "Other" as responses ($p_{75} = 4.35\%$; $p_{90} = 26.67\%$). These 13 interviewers are slightly older and more experienced compared to the rest of the interviewer staff. They do not differ in terms of gender and education.

Are these differences across interviewers driven by characteristics of the interviewers or by respondent, household, or other confounding factors? We use the widely accepted intra-interviewer variance coefficient ρ_{int} proposed by Kish (1962) to quantify interviewer variance in item nonresponse. Interviewer variance relates the interviewers' contribution to the total variation in a survey variable, resulting from the individual biases introduced by each interviewer. The more homogeneous the responses collected by individual interviewers, compared across interviewers, the higher the share of variance that is due to the interviewers. There is a large body of literature on measures of intra-interviewer correlation in survey variables (e.g., Groves 2004, 365; Schnell and Kreuter 2005; West and Olson 2010). They show that interviewer effects are present across all survey topics and question types.

Multilevel cross-classified linear mixed models (see Rasbash and Goldstein 1994) are used to estimate the intra-interviewer variance for item nonresponse and the replying of "Other". These models acknowledge that respondents are nested in households and households are nested in a cross-classified structure of geographic areas (German counties) and interviewers. Using cross-classified models can allow separating interviewer from area effects, a general problem in many large-scale surveys as interviewers are allocated to a specific geographic area, that is, single or few sample clusters only. In these cases, estimated interviewer effects are likely confounded with area effects (Schnell and Kreuter, 2005; Campanelli and O'Muricheartaigh, 1999; Durrant and D'Arrigo, 2014); answers observed by a single interviewer may be more homogeneous not because of the interviewers' biasing effects on responses, but due to the homogeneity of individuals living in the same geographic area. In the SOEP, on average, each interviewer is assigned to five German counties (Median = 5, Min = 1, Max = 17). In each county, on average, six interviewers are conducting interviews (Median = 5, Min = 1, Max = 36).

Whether a respondent provides a non-valid, that is, missing answer (y/n), functions as the dependent variable in the first model. In the second model, and similar to the above analysis of mode effects, we estimate the probability of respondents replying "Other" (yes/no). Adding a number of geographic area covariates and respondent characteristics in our models further minimizes potential problems due to confounding with area effects. Controls include respondent socio-demographics (gender, age, education), current partnership status (yes/no), as well as area characteristics at neighborhood level (street type, age distribution, share of Turkish migrants, move turnout, socio-economic status), at municipality level (size, voting results of the 2013 German Federal Election, age distribution), as well as at county level (unemployment rate, share of foreigners, share of higher educated employees, share of students). The intra-class correlation for the

interviewers is then derived as the share of the interviewer-level variance compared to the total variance that is decomposed into interviewer-specific variance, the area-specific variance, the household-specific variance, and the individual respondent residual variance.

The estimated share of variance in item nonresponse (yes/no) that is due to the interviewer is 0.43 (or 43%). Thus, almost half of the variance in item (non)response originates from the interviewer level. In other words, the interviewer largely influences the propensity of a respondent to provide a valid answer or not. A possible explanation is that some interviewers may feel uncomfortable asking the question because they perceive it as too sensitive, and thus, choose to not ask the question at all. With respect to the selection of "Other" as an answer, the interviewer variance is estimated even higher at 0.88%. Thus, almost 90% of the variance in choosing the other category can be explained by knowing which respondent is allocated to which interviewer. A possible explanation could be that some interviewers choose to not read out the answer option "Other" at all. On the other hand, some interviewers may even suggest to respondents to answer "Other" because they perceive other answers as being too sensitive. Moreover, some interviewers may just not read out the question at all and just answer "Other" for the respondent. While this clearly violates the interview protocol, it is usually not detected by fieldwork management as a) the response is not flagged as item "nonresponse", and b) there is no chance for inconsistencies in responses, since "Other" is compatible with any gender and partnership-constellation. Finally, interviewers themselves may misinterpret the question as relating to a person currently in a relationship only.

Do specific interviewer characteristics explain the observed nonresponse rates? For instance, older interviewers may feel less comfortable asking about sexual orientation, thus achieving lower response rates. To test this, we add a number of interviewer characteristics into the models. This includes socio-demographics (gender, age, education), work experience, personality traits (Big Five, see McCrae and John 1992), as well as political attitudes.

Table 4 displays the results of two multilevel logistic regressions. Although interviewers differ considerably in item nonresponse rates, few of the interviewer characteristics tested exert a statistically significant effect on the observed response behavior. Interviewers with many years of experience (21+ years) obtain more item nonresponse compared to interviewers with up to ten years of work experience (OR = 2.82). Moreover, interviewers with a higher workload are more successful in obtaining valid responses (OR = 0.43 and 0.37). In this regard, the workload itself likely does not have a direct effect on responses. Rather, the workload reflects interviewer skills and experience, as more experienced interviewers are usually given higher workloads. Older interviewers collect more "Other" responses (OR = 1.09) while higher educated interviewers achieve less "Other" responses (OR = 0.32, and = 0.31). Finally, there is evidence that the interviewers' self-reported personality is associated with response: interviewers that describe themselves as comparatively extroverted are more likely to obtain "Other" as an answer (OR = 1.61).

3. A Partnership-Inferred Proxy of Sexual Orientation

Presumably, the most frequent form of data that researchers apply to operationalizing same-sex and opposite-sex couples is the 'household grid' (or household-matrix). This

Table 4. Explaining interviewer effects.

	Nonresponse (yes/no)		Reply "other"	
Variable (Interviewer)	Odds ratio	95%–CI	Odds ratio	95%–CI
Gender				
Male (ref.)	–	–	–	–
Female	0.80	0.47–1.38	0.66	0.32–1.33
Age	1.01	0.98–1.04	1.09***	1.05–1.15
Education				
Basic (ref.)	–	–	–	–
Secondary	0.70	0.34–1.42	0.32*	0.13–0.77
Tertiary	0.73	0.34–1.55	0.31*	0.12–0.81
Work experience (years)				
1–10 (ref.)	–	–	–	–
11–20	1.50	0.82–2.76	0.75	0.35–1.62
21+	2.82**	1.38–5.74	0.57	0.21–1.58
Number of personal interviews in wave 2016				
1–49	–	–	–	–
50–99	0.43**	0.24–0.77	0.69	0.33–1.46
100+	0.37*	0.17–0.79	0.94	0.36–2.40
Personality score (Big five)[1]				
Openness	1.04	0.79–1.35	1.12	0.78–1.64
Conscientiousness	1.37	0.96–1.95	1.41	0.89–2.21
Extraversion	0.91	0.65–1.26	1.61*	1.03–2.50
Agreeableness	0.98	0.70–1.39	1.42	0.91–2.21
Neuroticism	1.19	0.96–1.49	1.11	0.84–1.48
Party identification				
No party (ref.)	–	–	–	–
SPD	0.72	0.33–1.57	0.62	0.22–1.75
CDU/CSU	0.59	0.31–1.14	1.47	0.65–3.34
Greens	0.65	0.23–1.80	0.37	0.08–1.67
The left	0.44	0.15–1.35	1.71	0.44–6.71
Other	0.76	0.24–2.52	0.98	0.21–4.48
Constant	0.00**	0.00–0.24	0.00***	0.00–0.00
$n_{respondents}$	14,521		13,389	
$n_{interviewers}$	432		401	

*p < .05, **p < .01, ***p < .001.

[1]Range: 1 to 5 with lower values representing less open, agreeable, etc.

Multilevel logistic regression with respondents nested in interviewers.

Controls: Respondent gender, age, education, years in the panel, geographic region (north, east, south, west), whether another person was present (none, partner, other, no information), whether a respondent was in a relationship (y/n), household size, municipality size as well as an initial subsample identifier.

See Table A2 in the Appendix (Section 6) for the full model coefficients.

grid, which is typically used in census surveys, enlists all residents of a household. The responding householder reports, among others, gender of other household members and their relationship to these persons (e.g., partner, mother, child). In some other cases, questionnaires include items on the gender of partners, irrespective of whether the respondent and the partner currently cohabitate. In addition, longitudinal household surveys (e.g., SOEP, PSID, and Understanding Society) provide not only information on current partnerships, but also on past relationships, as long as they fall into the period of observation. In principle, biographical questionnaires implemented in longitudinal, as well as cross-sectional, surveys allow an extension of this period to any prior partnership (Bates and DeMaio 2013). However, a cursory search of partnership biographies in different surveys suggests that only in a minority of cases, the gender of the previous partner was included among the surveyed items. Conversely, in many questionnaires the introductory text presumes that partners are opposite-sex or restricts previous partnership to marriage, which in many countries excludes same-sex partnerships.

3.1. *Prevalence of Same-Sex Couples*

To provide an overview of the incidence of same-sex couples in the SOEP, Figure 1 compares the proportion of same-sex couples among all couples in the SOEP in 2016 with corresponding estimates from the 2015 German Microcensus (MZ), the 2010 U.S. Census, as well as cross-national estimates of the European Social Survey (ESS), pooling its 2002–2016 waves. The ESS cross-national estimates range from zero same-sex couples in the samples of Poland and Russia to almost 2% of all couples in Denmark. According to the ESS estimates, Germany is among the countries with above-average numbers of same-sex couples in Europe. Figure 1 also displays sizable differences between surveys within Germany. While the 2016 German Microcensus reports only 0.46% same-sex couples and the German ESS samples 1.7%, the SOEP estimate is 0.9% of all couples in Germany.

3.2. *Using Partnership Information as a Proxy of Sexual Orientation*

Many surveys ask their respondents about partnership and (typically, binary) gender of the partner. This allows distinguishing between respondents not in a partnership, respondents in a partnership with a person of the opposite sex, and respondents in a same-sex partnership.

Researchers use partnership information to infer a proxy of sexual orientation in case a direct measurement is not available (Black 2000; Black et al. 2007, Antecol et al. 2008; Leppel 2009; Klawitter 2011; Liu et al. 2013). This holds true for many household surveys and censuses in particular. In these studies, partnership information is the only information available for analyzing the life of (parts of the) LGB population and comparing them (again, with parts of the) heterosexual population.

There are a number of measurement issues with regard to a partnership-inferred proxy of sexual orientation. Firstly, it does not allow a distinction between a bisexual orientation and a homosexual or heterosexual orientation. A respondent who reports being in a same-sex partnership may be homosexual or bisexual, just like a respondent who reports being in an opposite-sex relationship may be heterosexual or bisexual.

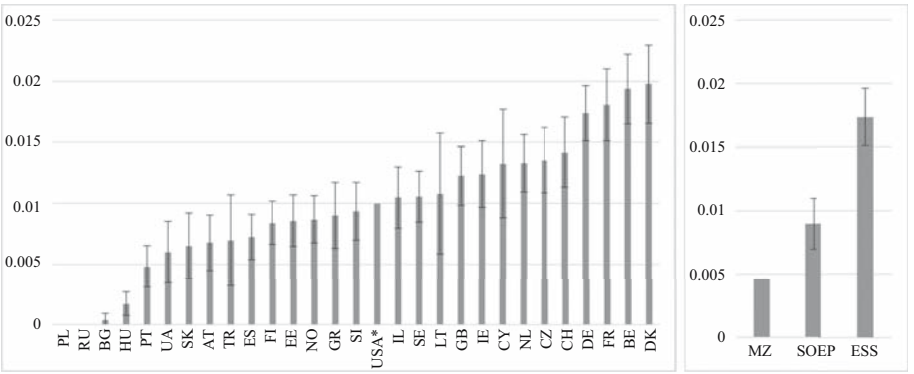


Fig. 1. Prevalence of same-sex couples across 29 countries (left) and three surveys in Germany (right).
Note. Population estimates of the rate of same-sex couples among all cohabitating couples.
Source. European Social Survey in waves 2002 to 2016 (Rounds 1–8). Socio-Economic Panel 2016 (SOEP.v33.1), U.S. Census (2010), 2015 German Microcensus (https://www.destatis.de/EN/FactsFigures/InFocus/Population/SameSexCouples.html). For estimates from the German Family Panel (pairfam), see Hank and Wetzel (2018). The estimated prevalence of SOEP uses information on the gender constellation of cohabiting partners only. All estimates weighted applying cross-sectional weights.
*2010 U.S. Census (Lofquist et al. 2012).

Secondly, the strategy relates to couples only. Thus, any comparisons between the LGB and heterosexual sub-populations are restricted to those who state that they are currently in a partnership, which excludes singles and couples living apart together.

Thirdly, respondents may be in a same-sex relationship, but identify as heterosexual and vice versa. As sexual orientation is subjective, the adequacy of 'objective' (observable) characteristics and measures, such as current partnership, is generally limited. Thus, partnership information is a valid proxy for some individuals, but not for others.

Finally, reporting error in the – often complex – survey instrument can result in the misclassification of same-sex and opposite-sex couples. Cortina and Festy (2014) provide an overview of different sources of reporting error, as well as the presence of this problem in census data of different countries (see also Festy 2007). Research on U.S. Census data shows that reporting error considerably inflates the number of same-sex couples. Kreider and Lofquist (2015), for instance, identify misreporting by matching couples from the U.S. Census with Social Security Administration records. A sizable number of U.S. Census same-sex couples appear to be opposite-sex in administrative records. Lengerer (2017), analyzing German Micro Census data, shows that item nonresponse in the household grid is a source of underestimation of the number of same-sex couples. DeMaio et al. (2013), classifying household member names by gender, also find considerable mismatches between reported gender in the household grid and the name-based gender of household members. These studies also identify ways to minimize reporting error by adapting the data collection mode and questionnaire design, for instance (see also Lewis et al. 2015).

4. Comparing Self-Reported and Partnership-Inferred Sexual Orientation

As previous research and this article suggest, both self-reported and partnership-inferred sexual orientation are potentially plagued by bias attributable to coverage, as well as

measurement. In the absence of information of the "true" sexual orientation of respondents, we are unable to compare the relative quality of both strategies directly. Thus, the following paragraphs provide some evidence on the extent to which both strategies lead to similar substantive conclusions about the differences between heterosexuals, on the one hand, and homosexuals and bisexuals, on the other hand.

For the partnership-inferred proxy of sexual orientation, we not only use information provided in the SOEP wave 2016, but any partnership-information available in the current and previous waves 2011 to 2015, including partnership-biographies. With this approach, we aim to minimise the risk of misclassifications due to bisexuals incorrectly classified as heterosexuals. In this regard, we refrain from coding sexual orientation if there is only a single piece of information on an opposite-sex partnership. Thus, respondents were classified as LGB in cases where they report a same-sex partnership once in the waves from 2011 to 2016. They are classified as heterosexual if they report an opposite-sex partnership in at least two waves (and never report a same-sex partnership). As a consequence, for many respondents, a classification is not possible because they have not been in a partnership within the observation period (referred to as "no partnership information" in the following).

Table 5 reports a cross-table of the number of SOEP respondents in 2016 classified as LGB by partnership-information, as well as their self-identification. Among the over 24,000 SOEP respondents who participated in 2016, on the basis of partnership-information, we categorize 153 as LGBs and 18,452 as heterosexuals (compared to 290 LGBs and 20,914 heterosexuals based on self-reported sexual orientation). In 5,682 cases (23%), no classification based on partnership-information was possible because a respondent has not been in a partnership over the last years. In comparatively few cases (747; 3%), a respondent provides a non-substantial answer *and* no information on current and previous partnerships.

As expected, for some respondents, the strategies of estimating sexual orientation lead to conflicting results. Thirty-two respondents out of more than 16,000, who self-report as heterosexual, were coded as LGB on the basis of partnership data. We analyzed whether these 32 cases are more recent additions to the panel, thus, pointing to the possibility that respondents may need more time to openly identify as Lesbian, Gay, or Bisexual in a survey. However, they do not differ significantly from the rest of the sample in terms of panel experience. Conversely, 66 out of 290 respondents, that is, 23%, who self-reported being homosexual or bisexual were coded as heterosexuals on the basis of available partnership data. As expected, the majority of these respondents identify as bisexual. Moreover, the conflicting classification results from the fact that both measures capture different aspects of sexual orientation. While the self-reports relate to the self-identification and self-perception of each individual, the partnership-inferred measure reflects acting sexual orientation in relationships – and these aspects can, but do not necessarily match.

Rather surprisingly, only 111 respondents (102 lesbian/gay, 9 bisexual) have consistent information indicating homosexual and bisexual orientation both in partnership and in self-reported data.

What can we learn from these results for survey practice? First of all, the cross-table emphasizes the fact that each method is accompanied by measurement error. While

Table 5. Self-reported and partnership-inferred sexual orientation.

	Partnership-inferred			
Self-Reported	Opposite-sex partnership (Heterosexual)	Same-sex partnership (LGB)	No partnership information	Total
Heterosexual	16,060	32	4,822	20,914
Lesbian/gay	11	102	45	158
Bisexual	55	9	68	132
Other	1,337	1	386	1,724
Refused	989	9	361	1,359
Total	18,452	153	5,682	24,287

Source. SOEP.v33, Wave 2016.

self-reported sexual identity is associated with substantial item nonresponse, partnership-inferred sexual orientation faces problems due to the non-coverage of singles and potential misclassification of bisexual respondents. Second, for many respondents, we find one measure of sexual orientation to substitute missing information on the other measure. Turning to LGBs, in 113 cases, absence of partnership data coincides with homosexual and bisexual self-reports, while in ten cases, absence of self-reported data coincides with same-sex partnership data. Using partnership-inferred sexual orientation may be used to fill in missing information in self-reported sexual orientation. Of the 3,083 respondents who either refused to answer or answered "Other", 2,336 respondents (76%) can be categorized as (likely) heterosexual or LGB, based on their present and previous partnership information. Thus, while some respondents (and their interviewers) may feel uncomfortable talking directly about their sexual orientation, in many cases, they are willing to provide information on the gender of their previous and current partners. Another plausible explanation relates to current relationship status; those currently not in a relationship were more likely to answer "Other" in wave 2016. However, they might have been in a relationship in previous years, thus, can be categorized based on past partnership information.

The cross-table also provides insights on the nature of item nonresponse in self-reports. The vast majority of respondents who provided a non-substantial answer when directly asked about their sexual orientation are classified as heterosexual based on their partnership information. Only one individual reported "Other" but is classified as LGB based on partnership information. Thus, refusing to answer or choosing "Other" seems to be a strategy applied by individuals in opposite-sex partnerships who either feel uncomfortable being asked about their sexuality or do not understand the question and terms (correctly). This is supported by our previous results on mode and interviewer effects on responding "Other".

In a next step, we had a look at potential consequences for applying one or the other measurement strategy when analyzing survey data, for instance, performing comparisons of LGBs and heterosexuals. Table 6 compares the distribution of demographic and socio-economic characteristics across (i) respondents categorized as LGBs by partnership information; (ii) respondents identifying as LGBs by self-reports; (iii) respondents

categorized as heterosexuals by partnership information, and finally (iv) respondents categorized as heterosexuals by self-reports. We only include respondents who report being in a partnership in the wave 2016. Please note that the four groups are not disjunctive, that is, in the case of two consistent measures of sexual orientation, respondents enter both (i) and (ii) or (iii) and (iv); but in the case of inconsistent (= conflicting) measures, respondents enter both (i) and (iv) or (ii) and (iii). Asterisks indicate mean differences between LGBs and heterosexuals at $p < 0.05$, that is, either between columns (i) and (iii) or between columns (ii) and (iv). We investigate whether there are significant mean differences based on one measurement strategy, but not the other.

The first – largely trivial – finding in Table 6 is that respondents whose sexual orientation was coded on the basis of current and previous partnership information are much more likely to currently live in the same household with their partner (85% and 97%). If we draw on self-reported sexual orientation, about 70% of LGBs and 88% of the heterosexuals cohabitate with a partner.

Based on partnership information, LGBs have higher employment rates (77%) than heterosexuals (59%). However, the division of labor in same-sex partnerships differs from opposite-sex partnerships, in that dual-income households are more frequent among same-sex partnerships (Kroh et al. 2017). If we use the self-reported information, we do not find a significant difference in employment by sexual orientation (66% versus 59%) (for effects of the measurement on earning differences by sexual orientation, see Klawitter 2014).

Turning to other characteristics, such as the distribution of gender, highest level of education, and political party identification, both strategies of measuring sexual orientation result in similar estimates. Compared to heterosexuals, not only do LGBs report university entrance degrees ("Abitur", German high school diploma) more frequently, but they also identify more frequently with the Green Party and less frequently with the Christian Democrats.

Table 6. Descriptives of LGBs and heterosexuals in a partnership in 2016 by measures of sexual orientation.

	(i) LGB Partnership N = 136	(ii) LGB self-report N = 198	(iii) Heterosexual partnership N = 16,872	(iv) Heterosexual self-report N = 16,631
Age 17–29	0.16	0.26	0.07*	0.13*
Age 30–45	0.45	0.38	0.33*	0.33
Age 45–60	0.29	0.26	0.33	0.31
Age 60+	0.10	0.10	0.28*	0.24*
Female	0.60	0.61	0.61*	0.52*
Partner in HH	0.85	0.69	0.97*	0.88*
Employed	0.77	0.66	0.59*	0.59
Municipality < 20K inh.	0.26	0.27	0.43*	0.41*
Municipality 20–100K inh.	0.22	0.21	0.27	0.28*
Municipality > 100K inh.	0.52	0.51	0.29*	0.30*
Univ. Entrance Degree	0.50	0.45	0.37*	0.38*
Support CDU/CSU	0.17	0.16	0.39*	0.38*
Support Green Party	0.36	0.34	0.14*	0.15*

Source. SOEP.v33, Wave 2016. Asterisks indicate mean differences between LGBs and heterosexuals at $p < 0.05$, that is, either between columns (i) and (iii) or between columns (ii) and (iv).

While item nonresponse in self-reported sexual orientation is particularly high among respondents age 65 and older (see Table A2 in the Appendix, Section 6), missing information on the partnership-inferred proxy of sexual orientation decreases by age. This is because younger respondents often lack biographical information on previous partnerships and are still single. Thus, the age distribution of respondents with valid information differs considerably between measures of sexual orientation. In self-reported data, both heterosexuals and LGBs are considerably younger than in the partnership-inferred data. Hence, if differences between heterosexuals and LGBs are present in certain periods of the life cycle, but not in others, the choice of the sexual orientation measure also affects substantive estimates.

5. Conclusions

As diversity is an emerging topic in many Western societies and in politics, there is a growing need for data that empirically describes diverse forms of living arrangements, both from academic and public policy perspectives. For instance, EU institutions repeatedly advise member states to monitor the equality of LGBITQ* in various areas of life (see e.g., The European Parliament 2014, "Resolution of 4 February 2014 on the EU roadmap against homophobia and discrimination on grounds of sexual orientation and gender identity"). Exemplifying this trend, official statistics in several countries now report statistics by sexual orientation and gender identity. For example, the U.S. Census publishes data by sexual orientation (U.S. Census Bureau 2019), as does the UK Office for National Statistics (2019), the Australian Bureau of Statistics (2016), and the Federal Statistical Office of Germany (2019), to name a few. However, generalizability of research on sexual orientation is hampered by the lack of consensus on its operationalization. While most researchers presumably would agree that the measurement of self-reported sexual orientation is the preferred option the omission of this measure in many surveys constrains secondary data users to operationalize sexual orientation alternatively using proxy information on the gender of respondents' partner.

The most obvious limitation of this partnership-inferred proxy of sexual orientation is the exclusion of singles and partners living apart, as well as the misclassification of bisexuals. However, the present article shows that implementing self-reports of sexual orientation in surveys also comes with the potential of error. The analyses show that self-reporting is sensitive to mode of data collection and interviewer characteristics. Survey practitioners may choose to rely on interviewers to minimize item nonresponse, but implement the question about sexual orientation within a (computer-assisted) self-interviewing module in order to minimize socially desirable answering behavior (see De Leeuw et al. 2003).

Moreover, we illustrate that the partnership-inferred and the self-reported measures of sexual orientation may result in dissimilar conclusions about the differences between heterosexuals and LGBs. In particular, research linked to with characteristics of partnership, such as occupation, earnings, and social networks is to be interpreted with caution, as is likely to be contingent on the choice of measure of sexual orientation. Therefore, we strongly argue in favour of also collecting data on self-reported sexual orientation in multidisciplinary survey data to enhance the reliability and relevance of LGB research.

Based on our results, there are a number of practical implications when collecting and analyzing survey data on sexual orientation:

- Mode of Data Collection
  - With respect to mode selection, we recommend interviewer-administered self-interviewing by, for instance, CASI mode. Our results suggest that the presence of the interviewer reduces item nonresponse and the privacy of the self-interview increases the prevalence of LGB identification.
  - Question wording and placement need to make sure that respondents understand the question as not only relating to current partnerships.
  - Instead of providing a closed-ended "Other" response option in surveying sexual orientation, which seems to be used as a substitute for refusals by some respondents, we recommend using an open-ended response option with follow-up write-ins.
- Interviewer Effects
  - The large interviewer effects reported in the article point to the necessity of extensive interviewer training. Interviewers should be specifically informed about the importance of the question and how to handle potential concerns, uncertainties and discomfort of respondents (and themselves). Also, interviewer monitoring during fieldwork may be used to identify interviewers who struggle with asking the question appropriately.
  - We suggest acknowledging interviewer effects in the analysis of the survey data by estimating multi-level regression models. This prevents researchers from underestimating standard errors and minimizes the likelihood of type 1 errors, that is, the rejection of a true null hypothesis.
- Self-reported versus partnership-inferred sexual orientation
  - Wherever possible, researchers should obtain self-reports on sexual orientation rather than information on partnership constellation only. Self-reports allow to address the general population instead of the subpopulation of cohabitating couples, and they allow to classify bisexual (and e.g., pansexual) respondents. Also, from the perspective of research ethics, directly probing sensitive information, including an open-ended response option appears superior to inferring the information from other sources.
  - In case both self-reports and partnership information is available, we suggest combining both sources of information, for instance, by analyzing current and previous partnerships for respondents that refused to provide self-reports. In the case of the SOEP sample, combining both ways of measuring sexual orientation allows classifying 97% of all adult respondents as either LGB or heterosexual. Although partnership-inferred information used as a replacement for self-reported information may be marred by misclassification error, the combination of data lends itself for estimating the bounds or error. Alternatively, we may use partnership-inferred information and self-reported information with a large set of additional variables within a multiple-imputation framework, replacing missing information. Again, the partnership-inferred information will be a strong predictor of sexual orientation and this approach lends itself for assessing the uncertainty of combining information.

6. Appendix

Table A1. Composition of respondents across modes.

	Face-to-face		Self-administered	
Variable	N	%	n	%
Third person present during interview				
No information	9,382	53	5,266	78
Partner	4,314	25	1,151	17
Other	3,587	20	280	4
No third person	262	2	45	1
Respondent Panel Experience				
1–2 years	1,634	9	286	4
3–5 years	5,022	29	501	7
6–10 years	6,285	36	1,528	23
11+ years	4,604	26	4,427	66
Subsample				
Early cross-sectional	8,452	48	3,479	52
Migrants	4,085	23	304	5
Families	3,869	22	1,360	20
Other	1,139	7	1,599	23
Respondent Sex				
Male	8,051	46	3,068	46
Female	9,494	54	3,674	54
Respondent Age				
<25	1,888	11	775	12
26–35	2,602	15	766	11
36–50	5,689	32	1,968	29
51–65	3,856	22	2,002	30
66+	3,510	20	1,231	18
Household size				
Single	1,617	9	381	6
2	4,706	27	1,410	21
3	3,241	18	1,316	20
4	4,144	23	1,866	28
5	2,195	13	1,124	17
6+	1,672	10	645	10
Region				
West Germany	14,242	81	5,010	74
East Germany	3,303	19	1,732	26

Table A2. Full model – explaining interviewer effects.

	Nonresponse (y/n)		Rep. "other" (y/n)	
Variable (interviewer)	OR	95% CI	OR	95% CI
Gender				
Male (ref.)	–	–	–	–
Female	0.80	0.47–1.38	0.66	0.32–1.33
Age	1.01	0.98–1.04	1.09***	1.05–1.15
Education				
Basic (ref.)	–	–	–	–
Secondary	0.70	0.34–1.42	0.32*	0.13–0.77
Tertiary	0.73	0.34–1.55	0.31*	0.12–0.81
Work experience (years)				
1–10 (ref.)	–	–	–	–
11–20	1.50	0.82–2.76	0.75	0.35–1.62
21+	2.82**	1.38–5.74	0.57	0.21–1.58
Number of personal interviews in wave 2016				
1–49	–	–	–	–
50–99	0.43**	0.24–0.77	0.69	0.33–1.46
100+	0.37*	0.17–0.79	0.94	0.36–2.40
Personality-score (Big five)[1]				
Openness	1.04	0.79–1.35	1.12	0.78–1.64
Conscientiousness	1.37	0.96–1.95	1.41	0.89–2.21
Extraversion	0.91	0.65–1.26	1.61*	1.03–2.50
Agreeableness	0.98	0.70–1.39	1.42	0.91–2.21
Neuroticism	1.19	0.96–1.49	1.11	0.84–1.48
Party identification				
No party (ref.)	–	–	–	–
SPD	0.72	0.33–1.57	0.62	0.22–1.75
CDU/CSU	0.59	0.31–1.14	1.47	0.65–3.34
Greens	0.65	0.23–1.80	0.37	0.08–1.67
The Left	0.44	0.15–1.35	1.71	0.44–6.71
Other	0.76	0.24–2.52	0.98	0.21–4.48
Third person present during interview				
No information (ref.)	–	–	–	–
Partner	0.71	0.44–1.15	0.63	0.38–1.04
Other	0.61	0.37–1.02	0.82	0.47–1.45
No third person	0.52**	0.34–0.80	0.38***	0.24–0.61

Table A2. Continued.

Variable (interviewer)	Nonresponse (y/n)		Rep. "other" (y/n)	
	OR	95% CI	OR	95% CI
Time of interview within individual interviewer workload				
First half (ref.)	–	–	–	–
Second half	1.11	0.88–1.39	0.56***	0.43–0.73
Subsample				
Cross-sectional (ref.)	–	–	–	–
Migrants	1.99***	1.36–2.93	1.44	0.88–2.34
Families	0.64*	0.42–0.96	3.79***	2.54–5.65
Other	0.82	0.49–1.37	1.14	0.66–1.97
German region				
North	–	–	–	–
East	1.74	0.74–4.10	0.60	0.21–1.68
South	1.97	0.85–4.59	0.92	0.34–2.54
West	1.98	0.90–4.34	0.80	0.32–2.01
Municipality size				
<2,000 (ref.)	–	–	–	–
2,000–5,000	0.65	0.39–1.09	0.88	0.46–1.70
5,000–20,000	0.76	0.49–1.16	0.97	0.56–1.79
20,000–50,000	0.59*	0.36–0.95	0.81	0.44–1.51
50,000–100,000	0.61	0.34–1.08	0.63	0.31–1.28
100,000–500,000	0.94	0.57–1.54	0.70	0.37–1.34
500,000+	0.58	0.33–1.02	0.56	0.26–1.20
Respondent panel experience				
1–2 years (ref.)	–	–	–	–
3–5 years	0.77	0.52–1.14	0.65	0.41–1.03
6–10 years	0.81	0.50–1.31	1.11	0.65–1.88
11+ years	1.05	0.63–1.73	0.91	0.50–1.64
Respondent sex				
male (ref.)	–	–	–	–
female	1.30**	1.07–1.59	1.21	0.97–1.52
Respondent age				
<25 (ref.)	–	–	–	–
26–35	0.81	0.50–1. 31	0.57*	0.33–0.97
36–50	0.96	0.62–1.47	0.98	0.62–1.54
51–65	1.70*	1.09–2.64	1.36	0.84–2.22
66+	2.58***	1.60–4.16	3.29***	1.92–5.65

Table A2. Continued.

Variable (interviewer)	Nonresponse (y/n)		Rep. "other" (y/n)	
	OR	95% CI	OR	95% CI
Respondent education				
basic (ref.)	–	–	–	–
secondary	1.06	0.84–1.35	1.01	0.77–1.32
tertiary	0.60**	0.43–0.82	0.74	0.53–1.04
In a relationship?				
No (ref.)	–	–	–	–
Yes	0.29***	0.22–0.37	0.40***	0.30–0.55
Household size				
Single (ref.)	–	–	–	–
2	0.91	0.64–1.31	0.81	0.52–1.26
3	0.94	0.62–1.40	0.92	0.57–1.46
4	1.02	0.67–1.55	0.84	0.51–1.39
5	1.31	0.83–2.09	1.10	0.65–1.89
6+	1.14	0.70–1.86	0.87	0.51–1.50
Constant	0.00**	0.00–0.24	0.00***	0.00–0.00
$n_{respondents}$	14,521		13,389	
$n_{interviewers}$	432		401	

*p < .05, **p < .01, ***p < .001.
[1]Range: 1 to 5 with lower values representing less open, agreeable, and so on.
Multilevel logistic regression with respondents nested in interviewers.

7. References

Aicken, C.R.H., C.H. Mercer, and J.A. Cassell. 2013. "Who reports absence of sexual attraction in Britain? Evidence from national probability surveys." *Psychology and Sexuality* 4(2): 121–135. DOI: https://doi.org/10.1080/19419899.2013.774161.

Antecol, H., A. Jong, and M. Steinberger. 2008. "The Sexual Orientation Wage Gap: The Role of Occupational Sorting and Human Capital." *Industrial and Labor Relations Review* 61(4): 518–543. DOI: https://doi.org/10.1177/001979390806100405.

Australian Bureau of Statistics. 2016. Same-Sex Couples in Australia, 2016. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2071.0Main+Features852016?OpenDocument (accessed October 2019).

Bates, N. and T.J. DeMaio. 2013. "Measuring same-sex relationships." *Contexts* 12: 66–69. DOI: https://doi.org/10.1177/1536504213476251.

Black, D.A. 2000. "Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources." *Demography* 37(2): 139–154. DOI: https://doi.org/10.2307/2648117.

Black, D.A., S.G. Sanders, and L.J. Taylor. 2007. "The Economics of Lesbian and Gay Families." *The Journal of Economic Perspectives* 21(2): 53–70. DOI: https://doi.org/10.1257/jep.21.2.53.

Booker, C.L., G. Riger, and J.B. Unger. 2017. "Sexual orientation health inequality: Evidence from Understanding Society, the UK Longitudinal Household Study." *Preventive Medicine* 101: 126–132. DOI: https://doi.org/10.1016/j.ypmed.2017.06.010.

Campanelli, P. and C. O'Muircheartaigh. 1999. "Interviewers, Interviewer Continuity, and Panel Survey Nonresponse." *Quality & Quantity* 33(1): 59–76. DOI: https://doi.org/10.1023/a:1004357711258.

Chandra, A., C.E. Copen, and W.D. Mosher. 2013. "Sexual Behavior, Sexual Attraction, and Sexual Identity in the United States: Data from the 2006–2010 National Survey of Family Growth." In *International Handbook on the Demography of Sexuality*, vol 5, edited by A. Baumle, 45–66. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-007-5512-3_4.

Cortina, C., and P. Festy. 2014. *Identification of same-sex couples and families in censuses, registers and surveys*. Families and Societies Working Paper Series 8. Available at: http://www.familiesandsocieties.eu/wp-content/uploads/2014/12/WP8CortinaFesty2014.pdf (accessed October 2019).

Dahlhamer, J.M., A.M. Galinksy, S.S. Joestl, and B.W. Ward. 2014. "Sexual orientation in the National Health Interview Survey: A quality assessment." *Vital Health Stat* 2(169) Available at: https://www.researchgate.net/publication/269713986_Sexual_Orientation_in_the_2013_National_Health_Interview_Survey_A_Quality_Assessment (accessed November 2019).

De Leeuw, E., J. Hox, and S. Kef. 2003. "Computer-Assisted Self-Interviewing Tailored for Special Populations and Topics." *Field Methods* 15(3): 223–251. DOI: https://doi.org/10.1177/1525822x03254714.

DeMaio, T.J., N. Bates, and M. O'Connell. 2013. "Exploring Measurement Error Issues in Reporting of Same-Sex Couples." *Public Opinion Quarterly* 77(1): 145–158. DOI: https://doi.org/10.1093/poq/nfs066.

Dillman, D.A., G. Phelps, R. Tortora, K. Swift, J. Kohrell, J. Berck, and B.L. Messer. 2009. "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet." *Social Science Research* 38: 1–18. DOI: https://doi.org/10.1016/j.ssresearch.2008.03.007.

Durrant, G.B. and J. D'Arrigo. 2014. "Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal." *Sociological Methods & Research* 43(3): 490–518. DOI: https://doi.org/10.1177/0049124114521148.

Federal Statistical Office of Germany. 2019. Gleichgeschlechtlichen Paare (darunter: gleichgeschlechtliche Ehepaare und eingetragene Lebenspartnerschaften). Available at: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Tabellen/3-4-gleichgeschlechtliche-lebensgemeinschaften.html?nn=209096 (accessed October 2019).

Festy, P. 2007. "Enumerating same-sex couples in censuses and population registers." *Demographic Research* 17: 339–368. DOI: https://doi.org/10.4054/demres.2007.17.12.

Frederiksen-Goldsen, K.I. and H.-J. Kim. 2014. "Count Me In: Response to Sexual Orientation Measures Among Older Adults." *Research on Aging* 37(5): 464–480. DOI: https://doi.org/10.1177/0164027514542109.

Galupo, M.P., E. Lomash, and R.C. Mitchell. 2016. "All of My Lovers Fit Into This Scale: Sexual Minority Individuals' Responses to Two Novel Measures of Sexual Orientation." *Journal of Homosexuality* 64: 145–165. DOI: https://doi.org/10.1080/00918369.2016.1174027.

Gates, G.J. 2011. *How many people are lesbian, gay, bisexual and transgender?* Los Angeles: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf (accessed October 2019).

Geary, R.S., C. Tanton, B. Erens, S. Clifton, P. Prah, K. Wellings, K.R. Mitchell, J. Datta, K. Gravningen, E. Fuller, A.M. Johnson, P. Sonnenberg, and C.H. Mercer. 2018. "Sexual identity, attraction and behaviour in Britain: The implications of using different dimensions of sexual orientation to estimate the size of sexual minority populations and inform public health interventions." *PLoS ONE* 13(1). DOI: https://doi.org/10.1371/journal.pone.0189607.

Goebel, J., M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp. 2018. "The German Socio-Economic Panel (SOEP)." *Journal of Economic and Statistics* 239(2): 345–360. DOI: https://doi.org/10.1515/jbnst-2018-0022.

Groves, R.M. 2004. *Survey Error and Survey Costs*. Hoboken, NJ: John Wiley & Sons.

Hank, K. and M. Wetzel. 2018. "Same-sex relationship experiences and expectations regarding partnership and parenthood." *Demographic Research* 39: 701–718. DOI: https://doi.org/10.4054/demres.2018.39.25.

Haseldon, L. and T. Joloza. 2009. "Measuring Sexual Identity: A Guide for Researchers." Office for National Statistics, Newport UK. Available at: https://osservatoriogender.it/wp-content/uploads/2016/10/sexualidentityuserguidefinal_tcm77-181188.pdf (accessed October 2019).

Heckman, J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement* 5: 475–492. Available at: https://ideas.repec.org/h/nbr/nberch/10491.html (accessed November 2019).

Heckman, J. 1979. "Sample selection bias as a specification error." *Econometrica* 47(3): 153–161. DOI: https://doi.org/10.2307/1912352.

Hilgert, L. M. Kroh, and D. Richter 2016. "The effect of face-to-face interviewing on personality measurement." *Journal of Research in Personality* 63: 133–136. DOI: https://doi.org/10.1016/j.jrp.2016.05.006.

Jans, M., J. Viana, D. Grant, S.D. Cochran, A.C. Lee, and N.A. Ponce. 2015. "Trends in sexual orientation missing data over a decade of the California Health Interview Survey." *American Journal of Public Health* 105(5): e43–e50. DOI: https://doi.org/10.2105/ajph.2014.302514.

Kish, Leslie. 1962. "Studies of Interviewer Variance for Attitudinal Variables." *Journal of the American Statistical Association* 57(297): 92–115. DOI: https://doi.org/10.1080/01621459.1962.10482153.

Klawitter, M. 2011. "Multilevel Analysis of the Effects of Antidiscrimination Policies on Earnings by Sexual Orientation." *Journal of Policy Analysis and Management* 30(2): 334–358. DOI: https://doi.org/10.1002/pam.20563.

Klawitter, M. 2014. "Meta-Analysis of the Effects of Sexual Orientation on Earnings." *Industrial Relations: A Journal of Economy and Society* 54(1): 4–32. DOI: https://doi.org/10.1111/irel.12075.

Kreider, R.M. and D.A. Lofquist. 2015. "Matching Survey Data with Administrative Records to Evaluate Reports of Same-Sex Married Couple Households." *SEHSD Working Paper*, 2014-36. U.S. Census Bureau: Washington, DC.

Kroh, M., S. Kühne, C. Kipp, and D. Richter. 2017. "Income, social support networks, life satisfaction: lesbians, gays, and bisexuals in Germany." *DIW Economic Bulletin* 2017: 335–345. Available at: https://diw.de/documents/publikationen/73/diw_01.c.563931.de/diw_econ_bull_2017-33-2.pdf (accessed October 2019).

Kroh, M., S. Kühne, R. Siegers, and V. Belcheva. 2018. "SOEP-Core: Documentation of Sample Sizes and Panel Attrition (1984 until 2016)." *SOEP Survey Papers* 480. Available at: https://www.diw.de/documents/publikationen/73/diw_01.c.579464.de/diw_ssp0480.pdf (accessed October 2019).

Krosnick, J.A., S. Narayan, and W.R. Smith. 1996. "Satisficing in surveys: Initial evidence." *New Directions for Evaluation* 1996(70): 29–44. DOI: https://doi.org/10.1002/ev.1033.

Kuha, J., C. Skinner, and S. Butt. 2014. *The effect of interviewer probing on item nonresponse and measurement error in cross-national surveys*. Presentation held at the 2014 American Association for Public Opinion Research Conference, May 15–18, 2014, Anaheim, California, U.S.A. Available at: https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/Abstract_Book_FNL.pdf (accessed November 2019).

Kühne, S. 2018. "From Strangers to Acquaintances? Interviewer Continuity and Socially Desirable Responses in Panel Surveys." *Survey Research Methods* 12(2): 121–146. DOI: https://doi.org/10.18148/srm/2018.v12i2.7299.

Leppel, K. 2009. "Labour Force Status and Sexual Orientation." *Economica* 76(301): 197–207. DOI: https://doi.org/10.1111/j.1468-0335.2007.00676.x.

Lewis, J.M., N. Bates, and M. Streeter. 2019. *Measuring Same-Sex Couples: The What And Who of Misreporting on Relationship And Sex*. U.S. Census Bureau. Available at: https://census.gov/content/dam/Census/library/working-papers/2015/demo/SEHSD-WP2015-12.pdf (accessed October 2019).

Lengerer, A. 2017. *Quality of Official Data on Cohabiting Same-Sex Couples in Germany*. Presentation held at the 2017 European Survey Research Association Conference, July 17th–21st 2017, Lisbon, Portugal.

Liu, H., C. Reczek, and D. Brown. 2013. "Same-sex Cohabitors and Health: the role of race-ethnicity, gender, and socioeconomic status." *Journal of Health and Social Behavior* 54(1): 25–45. DOI: https://doi.org/10.1177/0022146512468280.

Lofquist, D., T. Lugaila, M. O'Connell, and S. Feliz. 2012. "Households and Families: 2010." *Census Briefs*. C2010BR-14. Available at: https://www.census.gov/prod/cen2010/briefs/c2010br-14.pdf (accessed November 2019).

McCrae, R.R. and O.P. John. 1992. "An Introduction to the Five-Factor Model and Its Applications." *Journal of Personality* 60(2): 175–215. DOI: https://doi.org/10.1111/j.1467-6494.1992.tb00970.x.

Meyer, I.H. and P.A. Wilson. 2009. "Sampling lesbian, gay, and bisexual populations." *Journal of Counseling Psychology* 56(1): 23–31. DOI: https://doi.org/10.1037/a0014587.

Moleiro, C. and N. Pinto. 2015. "Sexual orientation and gender identity: review of concepts, controversies and their relation to psychopathology classification systems." *Frontiers in psychology* 6: 1511. DOI: https://doi.org/10.3389/fpsyg.2015.01511.

Puhani, P. 2000. "The Heckman correction for sample selection and its critique." *Journal of economic surveys* 14(1): 53–68. DOI: https://doi.org/10.1111/1467-6419.00104.

Rasbash, J. and H. Goldstein. 1994. "Efficient Analysis of Mixed Hierarchical and Cross-Classified Random Structures Using a Multilevel Model." *Journal of Educational and Behavioral Statistics* 19(4): 337–350. DOI: https://doi.org/10.3102/10769986019004337.

Richter, D. and J. Schupp. 2015. "The SOEP Innovation Sample (SOEP IS)." *Journal of Contextual Economics – Schmollers Jahrbuch* 135(3): 389–399. DOI: https://doi.org/10.3790/schm.135.3.389.

Schnell, R. and F. Kreuter. 2005. "Separating Interviewer and Sampling-Point Effects." *Journal of Official Statistics* 21(3): 389–41020 Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/separating-interviewer-and-sampling-point-effects.pdf (accessed November 2019).

Schwarz, N., F. Strack, H.-J. Hippler, and G. Bishop. 1991. "The impact of administration mode on response effects in survey measurement." *Applied Cognitive Psychology* 5(3): 193–212. DOI: https://doi.org/10.1002/acp.2350050304.

Sell, R.L. 2007. "Defining and Measuring Sexual Orientation for Research." In: *The Health of Sexual Minorities*, edited by I.H. Meyer and M.E. Northridge. Boston, MA: Springer. DOI: https://doi.org/10.1007/978-0-387-31334-4_14.

The European Parliament. 2014 "Resolution of 4 February 2014 on the EU roadmap against homophobia and discrimination on grounds of sexual orientation and gender identity". Available at: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0062+0+DOC+XML+V0//EN (accessed October 2019)

Tourangeau, R. and T. Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133(5): 859–883. DOI: https://doi.org/10.1037/0033-2909.

Uhrig, N. 2014. *An Examination of Poverty and Sexual Orientation in the UK*. Essex: Institute for Social and Economic Research. Available at: https://iser.essex.ac.uk/research/publications/working-papers/iser/2014-02.pdf (accessed October 2019).

UK Office for National Statistics. 2019. Sexual Identity lesbian, gay and bisexual population by Occupation by Country. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/sexuality/datasets/sexualidentitylesbiangayandbisexualpopulationbyoccupationbycountry (accessed October 2019).

U.S. Census Bureau. 2019. Same-Sex Couples. Available at: https://www.census.gov/topics/families/same-sex-couples/data/tables.html (accessed October 2019).

West, B.T. and K. Olson. 2010. "How Much Interviewer Variance Is Really Nonresponse Error Variance?" *Public Opinion Quarterly* 74(5): 1004–1026. DOI: https://doi.org/10.1093/poq/nfq061.

Williams Institute. 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys?* Los Angeles: The Williams Institute. Available at: https://www.williamsinstitute.law.ucla.edu/wp-content/uploads/SMART-FINAL-Nov-2009.pdf (accessed October 2019).

Wolff, M., B. Wells, C. Ventura-DiPersia, A. Renson, and C. Grov. 2016. "Measuring Sexual Orientation: A Review and Critique of U.S Data Collection Efforts and Implications for Health Policy." *The Journal of Sex Research* 54: 507–531. DOI: https://doi.org/10.1080/00224499.2016.1255872.

Wooden, M. 2014. *The Measurement of Sexual Identity in Wave 12 of the HILDA Survey (and associations with mental health and earnings)*. Melbourne: Melbourne Institute for Applied Economic and Social Research. Available at: https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-discussion-papers/hdps114.pdf (accessed October 2019).

Received November 2019
Revised June 2019
Accepted October 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 807–833, http://dx.doi.org/10.2478/JOS-2019-0034

Asking about Sexual Identity on the National Health Interview Survey: Does Mode Matter?

James M. Dahlhamer[1], *Adena M. Galinsky*[1], *and Sarah S. Joestl*[2]

Privacy, achieved through self-administered modes of interviewing, has long been assumed to be a necessary prerequisite for obtaining unbiased responses to sexual identity questions due to their potentially sensitive nature. This study uses data collected as part of a split-ballot field test embedded in the National Health Interview Survey (NHIS) to examine the association between survey mode (computer-assisted personal interviewing (CAPI) versus audio computer-assisted self-interviewing (ACASI)) and sexual minority identity reporting. Bivariate and multivariate quantitative analyses tested for differences in sexual minority identity reporting and non-response by survey mode, as well as for moderation of such differences by sociodemographic characteristics and interviewing environment. No significant main effects of interview mode on sexual minority identity reporting or nonresponse were found. Two significant mode effects emerged in subgroup analyses of sexual minority status out of 35 comparisons, and one significant mode effect emerged in subgroup analyses of item nonresponse. We conclude that asking the NHIS sexual identity question using CAPI does not result in estimates that differ systematically and meaningfully from those produced using ACASI.

Key words: Sexual orientation; mode of administration; question sensitivity; item nonresponse; field experiment.

1. Introduction

In recent years, there has been a call for more research on the health of lesbian, gay, bisexual, and transgender (LGBT) persons. In 2011, the Institute of Medicine (IOM 2011) published a seminal report that assessed the overall state of science on sexual minority health and identified a number of gaps in the scientific literature on this population. It noted, for instance, that more research is needed on inequities in health care, and that such research depends on the collection of sexual orientation data in community, state, and national health surveys. In line with the IOM's call for ongoing collection of sexual orientation data in federally funded surveys Healthy People 2020, the federal initiative that sets ten-year national objectives for improving the health of Americans, set an explicit objective of increasing the number of population-based data collection systems that can be used to monitor LGBT health (Healthy people 2020). A number of federal surveys collect sexual orientation data, including the National Health Interview Survey (NHIS), the National Survey of Family Growth, and the National Health and Nutrition Examination

[1] National Center for Health Statistics, Division of Health Interview Statistics, Hyattsville, Mayland, U.S.A. Email: jdahlhamer@cdc.gov, agalinsky@cdc.gov
[2] RAND Corporation, Santa Monica, California, U.S.A. Email: sjoestl@rand.org

Survey. The NHIS, with the largest sample size of those three, began collecting data on the sexual identity of adult respondents in 2013.

An important design feature of the NHIS is face-to-face (rather than telephone or mail-based) interviewing (FTFI). To collect data for the NHIS, an interviewer visits respondents in their homes and administers the survey questions aided by a laptop computer, otherwise referred to as computer-assisted personal interviewing (CAPI). (To convert some reluctant respondents and/or to complete missing portions of the interview, telephone follow-up is permissible.) There are several reasons for administering the NHIS face-to-face. Interviewers' ability to address respondents' questions and concerns about the survey can lead to higher cooperation rates (Groves et al. 2004), potentially reducing nonresponse bias in key survey statistics. Compared to other modes of administration, FTFI tends to produce higher response rates (Hox and De Leeuw 1994; Sykes and Collins 1988), lower item nonresponse rates (Brazier et al. 1992; De Leeuw and Van der Zouwen 1988), and longer verbal responses (De Leeuw and Van der Zouwen 1988; Sykes and Collins 1988). It also allows for longer, more complex interviews (Dialsingh 2008; Fowler 1993), and enables the collection of observational data on the part of the interviewers (Fowler 1993). During FTFI, interviewers can also assist in clarifying terms, probing, and motivating respondents to provide complete and accurate responses.

While FTFI offers many advantages over other modes of survey data collection, there are drawbacks to using it to collect sensitive information. A question is considered sensitive if it "raises concerns about disapproval or other consequences (such as legal sanctions) for reporting truthfully, or if the question itself is seen as an invasion of privacy" (Tourangeau and Smith 1996, 276). Asking such questions face-to-face may lead to greater nonresponse and deliberate misreporting than when such questions are included in self-administered formats.

Questions on sexual behavior, attraction, and identity, the three facets of sexual orientation, are generally considered sensitive. Take, for example, the following excerpt from a "best practices" document on asking sexual orientation survey questions: "Survey administrators need to be aware that (Lesbian, Gay and Bisexual, LGB) individuals are socially stigmatized, and disclosure of a gay, lesbian, or bisexual orientation (or same-sex sexual behavior or attraction) can have meaningful negative consequences for individuals with respect to workplace, family, and social outcomes" (Sexual Minority Assessment Research Team (SMART) 2009, 17). The report goes on to emphasize privacy as a guiding principle for collecting sexual orientation data and recommends self-administered modes of interviewing such as paper-and-pencil (PAPI), audio computer-assisted self-interviewing (ACASI), and telephone-ACASI (T-ACASI, also known as interactive voice response) (SMART 2009). However, there is little research evaluating these recommendations with regard to asking questions on sexual identity.

To address a gap in the scientific community's knowledge about the impact of survey data collection mode on sexual identity reporting, we present results from a field test conducted with the NHIS. The primary goal of this field test was to inform selection of a data collection mode for fielding the sexual identity questions beginning with the 2013 NHIS. To this end, adult respondents participating in the field test were randomly assigned to receive a ten-minute battery of questions on sexual identity and other topics of differing levels of sensitivity (e.g., neighborhood attachment, mental health, financial worries,

sleep, HIV testing) in either ACASI or CAPI (the standard mode of administration for the NHIS). The specific research questions we set out to answer and report on here include:

- Do estimates of the prevalence of sexual minorities (i.e., gay/lesbian and bisexual) differ by whether sexual identity questions are asked via CAPI or ACASI? If so, does ACASI produce a higher estimate of sexual minorities, as suggested by the literature?
- Does the impact of mode of administration on the reporting of sexual identity vary by subgroups defined by respondent sociodemographics and characteristics of the interviewing environment?
- Do item nonresponse rates differ by mode? If so, are the rates lower in ACASI compared to CAPI?

Before addressing these questions, we briefly summarize the literature on the difficulties inherent in obtaining accurate answers in response to sensitive questions, and then detail the findings of research on mode differences in the collection of data on sexual behavior, sexual attraction and sexual identity. We then describe the NHIS field test and the statistical analyses designed to address our research questions. Following the results of our analyses, we conclude by discussing the implications of our findings for the survey collection of sexual identity data.

2. Literature

2.1. Asking Sensitive Questions

As noted above, Tourangeau and Smith (1996) define a sensitive question as one that "raises concerns about disapproval or other consequences (such as legal sanctions) for reporting truthfully or if the question itself is seen as an invasion of privacy" (Tourangeau and Smith 1996, 276). Of particular concern to survey researchers are systematic misreporting and item nonresponse, especially refusal responses (Bradburn 1983; Fowler 1995; Tourangeau et al. 2000) that may occur when respondents are confronted with sensitive questions. Beyond studies documenting high item nonresponse to income questions (Dahlhamer et al. 2003, Dahlhamer et al. 2004; Juster and Smith 1997; Moore et al. 1999), few studies have formally addressed the link between question sensitivity and item nonresponse. Shoemaker et al. (2002) had students rate the sensitivity of survey questions and found that question sensitivity was positively related to item refusals, while Tourangeau and Yan (2007) identified what appeared to be a positive relationship between question sensitivity and item nonresponse, although the authors noted that a formal measure of sensitivity was not used.

With regard to misreporting, it has been demonstrated that asking sensitive questions can and does elicit systematic under- or over-reporting on a range of topics including abortion (Fu et al. 1998), substance use (Aquilino 1994; Gfroerer and Hughes 1992; Turner et al. 1992), and voter turnout (Bernstein et al. 2001; McDonald 2003). Tourangeau and Yan (2007) concluded that misreporting about sensitive topics is fairly common in surveys, that the extent of misreporting is contingent on whether the respondent has anything embarrassing to report, and that the level of misreporting is responsive to certain survey design features. They also conclude that misreporting is a motivated process in which

respondents alter their responses to avoid embarrassing themselves, particularly in the presence of an interviewer or other people. Hence, survey design features found to be effective in reducing motivated misreporting include self-administered data collection modes and providing respondents with a private setting in which to answer (Tourangeau and Smith 1996; Tourangeau and Yan 2007). Few empirical analyses report differences in responses to sensitive questions across different types of self-administered modes (Couper et al. 2003; Tourangeau et al. 2000), but removal of the interviewer from the interview setting consistently reduces misreporting on sensitive questions (Tourangeau and Yan 2007).

Consistent with "best practice" documents and research on the impacts of sensitive questions, questions about sexual identity and the other facets of sexual orientation (sexual attraction and sexual behavior) may be best suited for private survey settings and self-administered modes of data collection, such as computer-assisted self-interviewing (CASI), ACASI, and T-ACASI. In the next section, we review the existing literature on mode effects when asking about sexual orientation.

2.2. *Mode Effects with Questions on Sexual Orientation*

Sexual orientation consists of three distinct constructs: sexual behavior, sexual attraction, and sexual identity. Although heterosexuality is indeed a sexual orientation, in this review we consider only studies that have examined mode differences in reporting of sexual minority identities, attractions, and behaviors. In addition, this review focuses on those whose sexual minority status is defined either by their identity as a gay/lesbian or bisexual person, or by their attraction to or sexual behavior with persons of the same sex. Finally, we focus exclusively on studies comparing self-administered to interviewer-administered modes of data collection, as these are the most pertinent to our research.

A small number of studies have examined mode effects when asking about sexual identity. Midanik and Greenfield (2008) compared responses to questions on sexual identity, sexual behavior, and sexual and physical abuse between T-ACASI and CATI with the 2005 National Alcohol Survey and found that a significantly greater percentage of adults answering with T-ACASI identified as bisexual or homosexual. However, in multivariate analyses, significant differences held only for adults aged 40 or older. Interestingly, no differences in reporting of same-sex sexual behavior by mode were identified. Among patients at a sexually transmitted disease (STD) clinic, a significantly greater percentage identified as gay, lesbian, or bisexual when answering in ACASI compared to FTFI (Ghanem et al. 2005). Finally, field testing of a sexual identity question for inclusion in Office for National Statistics (United Kingdom) surveys revealed CASI to produce higher (but not statistically significantly different) estimates of people with a sexual minority identity (gay/lesbian or bisexual) than CAPI. In three CASI trials, prevalence estimates of sexual minorities ranged from 1.4% to 2.5%, with a combined estimate of 1.9%. (In the third trial, the interviewer could administer the sexual identity question in CAPI if the respondent did not want to use the laptop to enter their answers to the sexual identity question and other sensitive items.) For the fourth and final trial, CAPI produced an estimate of 1.6% (Malagoda and Traynor 2008).

Studies examining reporting of sexual attraction by mode found a similar pattern. For example, Caltabiano and Dalla-Zuanna (2012) found that a higher percentage of

respondents to a self-administered questionnaire (SAQ) reported same-sex attraction compared to those who answered by CATI. In addition, and in one of the few studies examining item nonresponse, the authors found that CATI elicited considerably more refusal responses to the same-sex attraction question than the SAQ. Both reported mode effects held in multivariate analyses. Based on data collected as part of the National STD and Behavior Measurement Experiment (NSBME), Villarroel et al. (2006) found more reporting of same-gender attraction (as well as same-gender sexual experiences and same-gender genital contact) among respondents answering by T-ACASI compared to respondents answering by CATI, effects that held in multivariate analyses. Item nonresponse rates to the gender attraction question, however, did not differ significantly by mode.

A greater number of studies exploring mode effects with sexual orientation reporting have focused on same-sex sexual behaviors. Other than the Villarroel et al. (2006) and Midanik and Greenfield (2008) studies reported earlier, these studies have relied on clinic or community samples. Simoes et al. (2006) explored mode differences in sexual behaviors reporting among a sample of adults seeking treatment for drug and alcohol abuse. Controlling for age, education, race, and marital status, they found that ACASI elicited more reports of men-having-sex-with-men (MSM) than FTFI. Similar results were observed among a sample of syringe-exchange program participants, with ACASI producing higher incidence rates of same-sex sexual behavior than FTFI (Des Jarlais et al. 1999). Likewise, among patients of an STD clinic, Kurth et al. (2004) found that ACASI elicited significantly more reporting of same-sex sexual encounters among both men and women compared to a clinician-administered health interview. ACASI also produced a lower item nonresponse rate to the sexual behavior questions than the clinician interview, although the authors note that this difference may have been due to other factors than item sensitivity (e.g., data entry error). Finally, Potdar and Koenig (2005) explored mode effects among two contrasting samples of urban men aged 18–22 from India: college students and slum residents not attending college. Among the college students, SAQ and ACASI both produced more reports of same-sex oral sex than FTFI. In addition, ACASI produced a significantly higher percentage of 2+ same-sex partner reports than did FTFI. Among the slum residents, ACASI elicited significantly higher reports of same-sex oral sex compared to FTFI, although FTFI elicited a significantly higher percentage of respondents reporting same-sex anal intercourse compared to ACASI.

Other studies using clinic and community samples found no differences in same-sex sexual behavior reporting by mode. A study of patients aged 15–39 at an urban STD clinic found no statistically significant differences in the percentage of respondents reporting same-sex sexual experiences in ACASI versus FTFI (Rogers et al. 2005). Similarly, a study of patients at an Australian sexual health clinic found no difference in the percentage reporting same-sex sexual behavior nor in the mean number of same-sex sexual partners reported across CASI and FTFI (Tideman et al. 2007). A study of perinatally HIV-exposed youth aged 9–16 attending an urban medical clinic identified no mode differences in responses to questions about same-sex sexual behavior (Dolezal et al. 2012). Furthermore, Jaya et al. (2008) found no mode differences in reports of same-sex sexual intercourse among economically disadvantaged youth in urban India.

In sum, research on the impact of survey mode on the reporting of sexual orientation has generally found that self-administered modes such as CASI, ACASI, and T-ACASI elicit more reports of gay/lesbian and bisexual self-identities, same-sex and bisexual sexual attraction, and, to a lesser extent, more reports of same-sex sexual behaviors than interviewer-administered modes (see Table 1). When significant effects have not been identified, the trend is generally toward greater reporting in the self-administered modes. However, many of the studies utilized very small, specialized, or international samples (e.g., clinic patients, youth in India), potentially limiting the generalizability of findings to large-scale, U.S. data collections. In addition, only a handful of these studies have looked at mode effects with regard to item nonresponse rates to questions on sexual orientation, with results being somewhat mixed. In the next section we describe the field test designed to address the question of whether ACASI would yield a greater percentage of adults identifying as a sexual minority than CAPI in the NHIS, a large-scale, general purpose, nationally representative health survey.

3. Data and Methods

3.1. National Health Interview Survey

The National Health Interview Survey (NHIS) is a multi-purpose survey of the health of the civilian, noninstitutionalized household population of the United States. Conducted by the National Center for Health Statistics (NCHS), the survey has been in the field continuously since 1957. Utilizing a multistage, clustered sample design, the NHIS produces nationally representative data on health insurance coverage, health care access and utilization, health status, health behaviors, and other health-related topics. The data are collected by trained interviewers with the U.S. Census Bureau using CAPI. Each year, interviews are conducted in roughly 35,000 households, yielding data on approximately 85,000–100,000 persons. Most interviews are conducted face-to-face in or immediately outside of respondents' homes.

The core survey instrument contains four main components: Household Composition, Family, Sample Child, and Sample Adult. For the household composition module, a household respondent provides basic sociodemographic information on all members of the household. Within each family, the family module is completed by a family respondent who provides health information on each member of the family. Additional health information is subsequently collected from the parent or guardian of one randomly selected child under aged 18 (the "sample child"), and one randomly selected adult (the "sample adult") aged 18 years or older. For the field test behind the analyses presented here, the ACASI module was located toward the end of the sample adult interview.

3.2. Description of the Field Test

Implemented between August 1 and October 15, 2012, the field test had three primary goals: 1) to test the ACASI instrument using normal interviewing protocols with a nationally representative sample, 2) to evaluate response rates for the newly-developed NHIS sexual identity question and the effect of adding the sexual identity question on response to the NHIS, and 3) to compare estimates of sexual identity and sexual minority

Table 1. Summary of studies examining mode effects with questions on sexual orientation.

Study	Measure	Population	Effect (%)	Item nonresponse (%)
Sexual attraction				
Caltabiano and Dalla-Zuanna (2012)	Same sex attraction	Weighted convenience sample (SAQ)/representative national sample (CATI), aged 18–69 in Italy (n = 3058 SAQ, 8285 CATI)	SAQ > CATI (6.9 vs. 3.0)	SAQ < CATI (1.3 vs. 11.6)
Villarroel et al. (2006)	Same sex attraction	U.S. nationally representative sample, and Baltimore representative sample, aged 18–45 (n = 1543 US, 744 Baltimore)	T-ACASI > CATI (17.8 vs. 12.8)	T-ACASI = CATI (1.5 vs. 1.2)
Sexual identity				
Midanik and Greenfield (2008)	Lesbian/gay/bisexual identity	U.S. nationally representative sample aged 18+ (n = 563 T-ACASI, 559 CATI)	T-ACASI > CATI among adults aged 40+ (bisexual: 2.5 vs. 0.6; homosexual: 1.9 vs. 0.9)	
Ghanem et al. (2005)	Lesbian/gay/bisexual identity	Baltimore STD clinic respondents aged 18-65 (n = 671 both modes)	ACASI > FTFI (3.0 vs. 1.0)	
Malagoda and Traynor (2008)	Lesbian/gay/bisexual identity	Nationally representative sample aged 16+ in Britain (n = 6422 CASI, 3429 CAPI)	CASI = CAPI (1.9 vs. 1.6)	
Sexual behavior				
Nationally representative samples				
Villarroel et al. (2006)	Same sex sexual behavior	U.S. nationally representative sample, and Baltimore representative sample, aged 18–45 (n = 1543 US, 744 Baltimore)	T-ACASI > CATI (same-gender sexual experiences: 14.2 vs. 9.1; same-gender genital contact: 10.3 vs. 7.0)	
Midanik and Greenfield (2008)	Same sex sexual behavior	U.S. nationally representative sample aged 18+ (n = 563 T-ACASI, 559 CATI)	T-ACASI = CATI (both genders: 6.9 vs. 4.7; same gender: 1.0 vs. 1.5)	
Clinic and community samples				
Simoes et al. (2006)	Male-male sexual behavior	Men aged 18+ seeking treatment for drug and alcohol abuse in Brazil (n = 367 ACASI and 368 FTFI)	ACASI > FTFI (12.6 vs. 5.7)	
Des Jarlais et al. (1999)	Same sex sexual behavior	Participants in syringe exchange programs in 4 U.S. cities (n = 724 ACASI and 757 FTFI)	ACASI > FTFI (10.0 vs. 5.0)	
Kurth et al. (2004)	Same sex sexual behavior	Patients ages 14+ at an urban, public STD clinic in the U.S. (n = 609 in both modes)	ACASI > clinician interview (men: 36.9 vs. 28.7; women: 19.6 vs. 11.5)	ACASI < clinician interview (men: 2.5 vs. 7.3; women: 0.7 vs. 6.2)

Table 1. Continued.

Study	Measure	Population	Effect (%)	Item nonresponse (%)
Rogers et al. (2005)	Same sex sexual behavior	Patients ages 15–39 at an urban U.S. STD clinic (n = 677 ACASI and 673 FTFI)	ACASI = FTFI (men: 10.1 vs. 8.5; women: 26.6 vs. 21.5)	
Tideman et al. (2007)	Same sex sexual behavior and median number of same sex partners	Patients at a sexual health clinic in Melbourne Australia (n = 255 CASI and 356 FTFI)	CASI = FTFI (men: 37.0 vs. 34.0; women: 11.0 vs. 7.0)	
Dolezal et al. (2012)	Ever any same sex sexual behavior	Urban, ethnic-minority, perinatally HIV-exposed medical clinic patients ages 9–16 in New York City (n = 135 ACASI and 139 FTFI)	ACASI = FTFI (baseline: 4.0 vs. 4.0; follow-up: 5.0 vs 11.0)	
Potdar and Koenig (2005)	Male-male oral sex	Unmarried male college students ages 18–22 in India (n = 300 ACASI, 300 SAQ and 300 FTFI)	ACASI, SAQ > FTFI (5.0 vs. 2.3 vs. 0.7)	
	2+ same sex partners		ACASI > FTFI (8.3 vs. 4.3)	
	Male-male oral sex		ACASI > FTFI (6.0 vs. 2.0)	
	Same-sex anal intercourse	Unmarried slum-residents not attending college ages 18–22 in India (n = 300 ACASI and 300 FTFI)	FTFI > ACASI (7.3 vs. 4.3)	
Jaya et al. 2008	Ever sexual intercourse with someone of the same sex	15–19 year old economically disadvantaged residents of a neighborhood in Delhi India (n = 1058 FTFI and 523 ACASI)	FTFI = ACASI (boys: 6.6 vs. 6.2; girls: 0.4 vs. 1.4)	

SAQ is Self-administered questionnaire.
CATI is Computer-assisted telephone interviewing.
ACASI is Audio computer-assisted self-interviewing.
T-ACASI is Telephone audio computer-assisted self-interviewing.
FTFI is Face-to-face interviewing.
CASI is Computer-assisted self-interviewing.
CAPI is Computer-assisted personal interviewing.

status between ACASI and CAPI. The test was designed to achieve a final minimum sample size of 5,000 completed interviews.

3.2.1. Sample Design

To achieve a nationally representative sample for the test, previously worked and unworked sample addresses from the 2006–2010 NHIS were utilized. To facilitate mode comparisons of sexual identity estimates, a split-ballot experiment was conducted in which 60% of sample adults were randomly assigned to receive the sexual identity questions by ACASI and 40% by CAPI. To ensure that both the ACASI and the CAPI samples were nationally representative, random assignment of mode took place at the time

of sample formation. No identifying information was present on a case that would permit interviewers to determine beforehand the assigned mode.

3.2.2. Field Test Implementation

After receiving a self-study and one-day classroom training on ACASI and the purposes of the test, roughly 475 U.S. Census Bureau interviewers were assigned caseloads to complete over a two-month period. Interviewers were asked to work these cases as they would their regular NHIS cases. Normal NHIS interviewing protocols were in place, including refusal conversion and telephone follow-up to complete missing portions of the interview. Interviewers were given permission to complete the module including the sexual identity question by telephone if it was not possible to obtain the data otherwise. This applied to both the CAPI and ACASI paths. While this placed constraints on our ability to isolate the effects of mode, telephone follow-up is and will continue to be a regular part of NHIS interviewing. Hence, it was decided that the field test should reflect current and future interviewing procedures rather than produce a true test of mode effects. In addition to telephone follow-up, interviewers were allowed to conduct ACASI cases by CAPI if respondents were reluctant to use the computer and would otherwise break off the interview. Roughly 13% of the interviews on each path were completed primarily by telephone. An additional 6% of interviews on the ACASI path were completed in CAPI.

At the conclusion of the test, 5,445 interviews had been completed at least to the beginning of the sexual identity module, exceeding the initial goal of 5,000. Of these, 3,210 sample adults were randomly assigned to ACASI and 2,235 to CAPI. The final family response rate for the field test was 77.3%. The final response rate for sample adults was 64.9% and 64.0% for the CAPI and ACASI paths, respectively. Overall, the field test achieved response rates that were slightly higher than the final response rates for the 2012 NHIS (family: 76.8%; sample adult: 61.2%).

3.2.3. Design of the ACASI Module

Because the NHIS is a general health survey with a diverse array of respondents, respondents with little to no computer experience and/or low levels of literacy would still need to be able to complete the ACASI module. Therefore, a simple three-key interface was developed. When the respondent was presented with a question, he/she would press the Space bar to scroll through the available response options, with a circle appearing around the currently selected response. Once the desired answer was circled, he/she would press the Enter key to select and retain that answer. When a respondent wanted to back up to review or change a previous answer, he/she would use the Tab key. Audio recordings of the question and response options automatically played when each question appeared on the screen. Recordings and question text were available in either English or Spanish.

3.2.4. Survey Instrument

The sexual identity questions were included in the Adult Selected Items (ASI) section of the Sample Adult interview. Both the CAPI and ACASI version of the ASI section was available in English and Spanish. Since the questions were not translated into other languages, interviewers were asked to skip the ASI section (both ACASI and CAPI) if the respondent was not comfortable answering in either English or Spanish. The ASI section

appeared toward the end of the Sample Adult interview and followed questions on access to health care and health care utilization. The section began with questions on computer use, satisfaction with health care, and neighborhood tenure and attachment. The sexual identity questions then followed. The remainder of the module consisted of questions on financial worries, sleep, mental health, and HIV testing.

In CAPI, interviewers proceeded seamlessly from the prior section into the ASI module. For ACASI, interviewers explained to respondents that they would complete the next set of questions on their own. They were asked in which language, English or Spanish, they would like to complete the questions. The interviewer then plugged headphones into the computer, turned the computer to the respondent, and proceeded to give a short tutorial on how to use the keyboard to enter responses and advance to the next question. Respondents were then asked to don the headphones and begin. At the outset of the ACASI module, the interviewer instructions were reinforced with a short set of practice questions. Throughout, respondents could wear the headphones or simply read the questions. Either way, they were asked to leave the headphones plugged in to mute the audio recordings. Once the respondent completed the questions, an exit screen appeared and asked them to return the laptop computer to the interviewer.

3.3. The Sexual Identity Question

The development and testing of a sexual identity question was an extensive effort carried out over an 11-year period. A total of 377 in-depth cognitive interviews were conducted by the NCHS Questionnaire Design Research Laboratory to better understand the interpretive and response process patterns people use to answer questions on sexual identity. A thorough description of the process is beyond the scope of this article, but readers may refer to the cognitive testing report by Miller and Ryan (2011). The resulting question used with the ACASI module read as follows:

Do you think of yourself as:

1. Gay
2. Straight, that is, not gay
3. Bisexual
4. Something else
5. I don't know the answer

Female ACASI respondents received a version where the first response option read "Gay or lesbian" and the second response option read "Straight, that is, not gay or lesbian." Although the goal was to keep question wording consistent across modes, some minor revisions were necessary in CAPI. To provide as much privacy as possible during face-to-face administration, the decision was made to use a flashcard listing the sexual identity response categories. To accommodate the use of a flashcard, the wording of the question stem was slightly different in CAPI compared to ACASI. When the main sexual identity question was reached, the interviewer would hand the flashcard to the respondent and read the following text: "Which of the following best represents how you think of yourself?" Looking at the flashcard, the respondent was asked to report the number associated with the most appropriate answer. The response categories on the flashcard were identical to

those that appeared on the computer screen for the ACASI respondents, with separate flashcards for male and female respondents. Respondents who answered gay, lesbian, or bisexual were considered to be sexual minorities.

3.4. Other Measures

For all mode comparisons of sexual minority estimates and item nonresponse, results are presented overall and for a select set of respondent sociodemographic and interviewing environment characteristics. Sociodemographic measures include age (18–44 versus 45 and older); sex; race and ethnicity (non-Hispanic white versus other); education (less than a high school diploma/General Educational Development high school equivalency diploma (GED) versus high school diploma/GED and higher); employment status (working versus not working); marital status (never married vs. other); reported health status (excellent/very good health versus poor/fair/good health); whether or not the respondent has a functional limitation; total family income from the prior calendar year (less than USD50,000 versus USD50,000 or more); whether or not the residence is owned/being bought or rented/some other arrangement; whether or not the residence is located in the central city of a metropolitan statistical area (MSA); and whether or not the residence is located in the West region. Interviewing environment measures included whether or not other family members aged 17 or older were present during the entire interview (including the ASI section) (yes or unknown); the number of contact attempts required to complete the interview (a commonly used measure to characterize the difficulty of the case; 1–2, 3–4, 5 or more attempts); whether or not the case was re-assigned to a different interviewer; whether or not householders expressed time constraints and/or privacy-related concerns prior to or during the interview; and location of the interview (inside the home, outside the home). Interviews completed by telephone were excluded from the "location of the interview" measure.

3.5. Statistical Procedures

Consistent with an intent-to-treat analysis, we retained sample adult interviews randomly assigned to one mode but conducted in another in the assigned mode group. Analyses performed with these cases removed did not substantially alter the results presented here.

To assess whether randomization had been successfully achieved, we compared the CAPI and ACASI groups on a set of sociodemographic and interviewing environment characteristics. Next, we compared overall estimates of sexual minority status by administration mode, including crude and adjusted odds ratios from logistic regressions. Covariates included in the multiple logistic regressions, listed in the footnote under Table 4, were found to be significantly associated with sexual minority status ($P < 0.15$) in bivariate analyses.

We then compared estimates of sexual minority status by mode within sociodemographic subgroups and types of interviewing environment. For each subgroup, we present the crude odds ratio for mode from a bivariate logistic regression of sexual minority status. If a significant association was identified within a subgroup, we then estimated a logistic regression model with sexual minority status as the dependent variable and mode, the sociodemographic or interviewing environment measure under analysis

(e.g., age), and an interaction term for the two as covariates. This allowed us to further assess whether the impact of mode on sexual minority reporting was homogeneous across subpopulations or whether certain subpopulations were particularly sensitive to mode.

We next compared sexual identity item nonresponse (a category which included not only refused but also "something else", and "I don't know the answer") rates by mode. For ACASI, "refused" included respondents who failed to provide a response to the question. We present crude and adjusted odds ratios from logistic regression models. Covariates included in the multiple logistic regressions (listed in the footnote under Table 7), were found to be significantly associated with sexual identity nonresponse ($P < 0.15$) in bivariate analyses.

Finally, we compared sexual identity item nonresponse rates by mode within sociodemographic subgroups and types of interviewing environment. We again present the crude odds ratio for mode from a bivariate logistic regression of sexual identity nonresponse. If a significant association between mode and item nonresponse was identified within a subgroup, we estimated a logistic regression model with item nonresponse as the dependent variable and mode, the sociodemographic/interviewing environment measure under analysis (e.g., number of contact attempts on the household), and an interaction term for the two as covariates.

We present 95% confidence intervals for all estimates. CAPI is the reference category for all mode odds ratios. All analyses were performed using SAS-callable SUDAAN version 11.0.1 to account for the complex sample design of the NHIS. Finally, to mimic normal NHIS production procedures and to ensure that estimates from each data collection mode were generalizable to the U.S. adult, civilian noninstitutionalized population aged $\geq$ 18 years, all analyses (unless otherwise noted) used final sample adult weights adjusted for nonresponse and calibrated to population control totals.

4. Results

4.1. Sample Equivalency

To determine whether the field test provides a valid assessment of mode differences, if any, in sexual minority reporting and item nonresponse, we compared the two mode groups on 13 respondent sociodemographic and social environmental measures (see Table 2). As shown, the distributions were similar by mode, with no significant differences being identified. The similar sample compositions by mode bolster our confidence in the subsequent results.

4.2. Estimates of Sexual Minority Status by Mode

Table 3 presents response distributions for the sexual orientation question by mode. A higher percentage of adults identified as gay or lesbian (1.4%) in CAPI compared to ACASI (0.9%), while a slightly higher percentage of adults identified as bisexual in ACASI (1.2%) compared to CAPI (1.0%). However, neither difference reached statistical significance. Compared to CAPI, ACASI also yielded a slightly higher percentage of adults answering "something else", "I don't know the answer", and refused. Again, the differences were not statistically significant. Since the overall number of adults identifying as a sexual minority

Table 2. Characteristics of sample adults who reached the sexual identity questions by interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI		CAPI		
	%	95% CI	%	95% CI	χ^2 p-value
Gender					0.87
Male	48.0	45.83, 50.24	48.3	45.98, 50.59	
Female	52.0	49.76, 54.17	51.7	49.41, 54.02	
Age					0.85
18–24	12.9	11.10, 14.73	12.3	10.07, 14.58	
25–44	34.5	31.67, 37.40	35.5	32.51, 38.46	
45–64	35.4	32.93, 37.86	34.6	32.15, 36.99	
65+	17.2	14.96. 19.35	17.6	15.43, 19.80	
Race/ethnicity					0.87
Hispanic	14.9	11.40, 18.48	14.8	11.59, 18.09	
Non-hispanic white	66.6	61.20, 72.05	66.3	60.82, 71.74	
Non-hispanic black	11.6	8.75, 14.41	11.4	8.65, 14.12	
Non-hispanic other	6.9	4.35, 9.35	7.5	5.10, 9.90	
Education					0.77
Less than high school	14.8	12.52, 17.00	15.1	12.80, 17.42	
High school/GED	27.2	24.81, 29.64	26.8	24.06, 29.45	
Some college	30.7	28.24, 33.07	32.1	28.95, 35.22	
Bachelor's+	27.4	23.64, 31.08	26.1	22.15, 29.95	
Employment status					0.07
Working	59.7	56.81, 62.64	62.4	59.43, 65.31	
Not working	40.3	37.36, 43.20	37.6	34.69, 40.57	
Marital status					0.36
Never married	20.8	18.42, 23.24	22.0	19.28, 24.74	
Married or cohabiting	61.7	59.07, 64.32	60.0	57.07, 63.00	
Divorced	10.8	9.52, 12.16	12.0	10.35, 13.56	
Widowed	6.6	5.53, 7.74	6.0	4.98, 7.02	
Reported health status					0.70
Excellent/very good	58.5	55.83, 61.12	59.8	56.80, 62.80	
Good	27.9	25.86, 29.86	27.2	24.90, 29.46	
Poor/fair	13.7	11.89, 15.44	13.0	10.95, 15.10	
Family income					0.81
< USD20,000	15.9	13.88, 17.92	16.8	14.38, 19.23	
USD20,000 – <USD50,000	29.5	26.54, 32.53	28.3	25.81, 30.89	
USD50,000 – <USD100,000	28.3	25.65, 30.98	28.6	25.86, 31.34	
≥ USD100,000	18.4	14.88, 21.99	17.8	14.55, 20.99	
Unknown	7.8	6.45, 9.18	8.5	6.71, 10.25	
Own or rent					0.56
Own or buying	64.6	60.36, 68.82	65.4	60.99, 69.81	
Rent or some other arrangement	35.4	31.18, 39.64	34.6	30.19, 39.01	
MSA status					0.73
MSA, central city	30.3	23.45, 37.15	30.6	23.85, 37.35	
MSA, non-central city	52.8	44.95, 60.75	53.2	45.58, 60.90	
Non-MSA	16.9	10.85, 22.85	16.2	10.27, 22.04	
Region					0.19
Northeast	21.2	12.15, 30.35	20.0	11.56, 28.50	
Midwest	24.0	14.14, 33.89	22.9	13.87, 32.02	
South	32.5	23.96, 41.01	33.4	24.67, 42.10	
West	22.2	13.98, 30.52	23.6	15.06, 32.22	

Note. CI = confidence interval.

Table 3. Responses to main sexual identity question by interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI		
	n	%	95% CI	n	%	95% CI
Sexual identity						
Gay or lesbian	36	0.9	0.43, 1.38	28	1.4	0.73, 1.99
Straight, that is not gay or lesbian	2,952	94.6	93.58, 95.58	2,105	95.2	94.01, 96.42
Bisexual	43	1.2	0.80, 1.63	29	1.0	0.53, 1.38
Something else	14	0.4	0.11, 0.60	8	0.3	0.05, 0.56
Don't know	70	1.9	1.31, 2.54	33	1.3	0.71, 1.92
Refused[a]	38	1.0	0.61, 1.44	21	0.9	0.28, 1.44

Note. CI = confidence interval.
[a]Included in the "refused" for ACASI participants are 33 cases where the respondent skipped the question.

was small (ACASI = 79, CAPI = 57), subsequent analyses focus on a dichotomous measure of sexual minority status (gay/lesbian or bisexual versus straight).

Table 4 presents estimates of sexual minority status by mode. Overall, ACASI produced a slightly lower estimate (2.2%) of sexual minorities than CAPI (2.4%), although the difference was not statistically significant.

Table 5 shows estimates of sexual minority status by mode and respondent sociodemographics. Again, the goal of these analyses is to test whether the effect of interview mode is homogeneous across subpopulations. For example, "Is the impact of survey mode on reporting a sexual minority status equivalent for men and women?" While the CAPI estimates were, on average, slightly higher (in 20 of 26 comparisons), only one statistically significant difference was identified. Adult respondents from families with annual incomes of USD50,000 or more were significantly more likely to identify as a sexual minority in ACASI (2.2%) than in CAPI (0.8%) (unadjusted odds ratio (UOR) = 2.95, 95% confidence interval (CI) = 1.16–7.50). A logistic regression of sexual minority reporting in which interview mode, total family income, and their interaction were included as covariates yielded a

Table 4. Percentage of sample adults identifying as a sexual minority (gay/lesbian or bisexual) by interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI			
	n	%	95% CI	n	%	95% CI	OR	95% CI	AOR	95% CI
Identified as a sexual minority	3,031	2.2	1.56, 2.84	2,162	2.4	1.59, 3.14	0.93	0.59, 1.45	0.92	0.59, 1.44

Note. CI = confidence interval; OR = odds ratio; AOR = adjusted odds ratio.
[a]The following covariates were included in the multivariate logistic regression: age, gender, education, marital status, reported health status, total family income from the prior calendar year, whether the residence is owned/being bought or rented/some other arrangement, whether or not the residence is in the West region, whether or not the residence is in the central city of an MSA, total count of contact attempts on the household, whether or not householder(s) expressed privacy or trust concerns, and whether or not householder(s) expressed time constraints.

Table 5. Percent sexual minority by select sociodemographics and interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI	
	n	%	95% CI	n	%	95% CI	OR	95% CI
Age								
18–44	1,405	2.5	1.53, 3.46	1,011	3.5	2.15, 4.79	0.71	0.39–1.30
45+	1,626	1.9	1.06, 2.78	1,151	1.3	0.41, 2.27	1.44	0.63–3.31
Sex								
Male	1,366	1.8	1.09, 2.48	931	2.1	1.10, 3.09	0.85	0.45–1.62
Female	1,665	2.6	1.61, 3.54	1,231	2.6	1.48, 3.76	0.98	0.56–1.71
Race/ethnicity								
Non–hispanic white	1,810	1.9	1.17, 2.67	1,263	2.3	1.39, 3.13	0.85	0.47–1.51
Other	1,221	2.8	1.39, 4.16	899	2.6	0.93, 4.24	1.08	0.46–2.50
Education								
Less than a high school diploma/GED	1,269	2.5	1.40, 3.59	903	3.3	1.73, 4.88	0.75	0.41–1.35
High school diploma/GED or more	1,756	2.0	1.33, 2.65	1,253	1.7	1.05, 2.38	1.16	0.70–1.94
Employment status								
Working	1,732	2.5	1.63, 3.37	1,282	2.3	1.30, 3.23	1.11	0.65–1.90
Not working	1,267	1.7	0.99, 2.48	879	2.5	1.20, 3.86	0.68	0.33–1.38
Marital status								
Never married	712	4.2	2.32, 6.10	539	4.4	2.44, 6.42	0.95	0.47–1.93
Other	2,312	1.7	1.05, 2.28	1,616	1.8	0.99, 2.60	0.93	0.51–1.69
Reported health status								
Excellent/very good	1,708	1.8	1.05, 2.54	1,236	2.0	1.10, 2.85	0.91	0.49–1.70
Good/fair/poor	1,322	2.8	1.60, 3.92	923	3.0	1.49, 4.42	0.93	0.50–1.73
Family income								
< USD50,000	1,577	2.5	1.64, 3.31	1,151	3.4	2.14, 4.71	0.72	0.41–1.24
≥ USD50,000	1,228	2.3	1.18, 3.33	835	0.8	0.20, 1.36	2.95	1.16–7.50
Own or rent residence								
Own or buying	1,787	1.5	0.82, 2.11	1,238	1.0	0.50, 1.58	1.42	0.70–2.85
Rent or some other arrangement	1,236	3.6	2.02, 5.11	920	4.9	2.91, 6.88	0.72	0.35–1.46
MSA status								
MSA, central city	1,049	3.4	2.08, 4.67	784	3.7	2.00, 5.39	0.91	0.48–1.72
Other	1,982	1.7	1.03, 2.35	1,378	1.8	0.95, 2.61	0.95	0.52–1.73
Region								
West	735	1.4	0.42, 2.47	576	1.8	0.80, 2.79	0.80	0.31–2.05
Other	2,296	2.4	1.66, 3.15	1,586	2.5	1.57, 3.51	0.95	0.57–1.56

Note. CI = confidence interval; OR = odds ratio.

significant interaction term ($p < .001$), lending support to non-equivalence in the impact of survey mode on reporting a sexual minority status across income subgroups.

Next, we turned to measures of the interviewing environment (Table 6) under the hypothesis that ACASI would elicit greater sexual minority (i.e., more accurate) reporting in more challenging (e.g., reluctant respondents), less private scenarios. Only one significant difference emerged by mode. The percentage of adults identifying as a sexual minority was significantly greater in CAPI (3.5%), compared to ACASI (1.4%), when one or more householders expressed privacy concerns (UOR = 0.40, 95% CI = 0.19–0.83). As with total family income, a logistic regression of sexual minority status with interview mode, expression of privacy concerns, and their interaction included as covariates produced a significant interaction term ($p < .05$). Here again, there appears to be non-equivalence in the impact of survey mode on reporting a sexual minority status across subgroups defined by whether or not householders expressed privacy concerns.

4.3. Item Nonresponse

For the next set of analyses we examined item nonresponse as a measure of data quality. Table 7 presents item nonresponse rates (all responses other than those coded as sexual minority or straight) to the main sexual identity question by mode. While the overall item nonresponse rate was slightly higher in ACASI (2.7%) compared to CAPI (2.3%), the difference did not reach statistical significance (UOR = 1.20, 95% CI = 0.77–1.89; AOR = 1.27, 95% CI = 0.83–1.96).

Table 8 presents item nonresponse rates by mode and respondent sociodemographics. Are the effects of mode on item nonresponse to the sexual identity question equivalent across subgroups? Consistent with the overall rate, we observe slightly higher nonresponse to the main sexual identity question in ACASI compared to CAPI across 12 respondent sociodemographics measures. In all, ACASI produced a higher item nonresponse rate for 19 of 26 comparisons. However, none of the observed differences were statistically significant.

Mode comparisons by the interviewing environment measures yielded similar results (see Table 9). Thirteen comparisons across six variables produced one significant difference: interviews conducted inside respondents' homes led to a higher nonresponse rate to the main sexual identity question in ACASI (2.6%) compared to CAPI (1.4%) (UOR = 1.79, 95% CI = 1.07–2.99). However, a logistic regression of item nonresponse to the sexual identity question with interview mode, location of interview, and the interaction of the two as covariates did not yield a significant interaction term ($P = .06$).

5. Discussion

5.1. Sexual Minority Status

While the larger literature on sexual orientation reporting is consistent that self-administered interview modes yield more sensitive and socially undesirable responses, we found no statistically significant differences in the overall percentage of adults identifying as a sexual minority (gay/lesbian and bisexual) in ACASI (2.2%) compared to CAPI (2.4%) in the NHIS field test. In addition, no significant mode differences in sexual minority

Table 6. Percent sexual minority by select interview environment measures and interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI	
	n	%	95% CI	n	%	95% CI	OR	95% CI
Presence of others								
Yes	989	2.5	1.33, 3.69	692	3.0	1.55, 4.43	0.84	0.44–1.59
Unknown	2,042	2.0	1.29, 2.64	1,449	1.9	1.09, 2.69	1.04	0.60–1.82
Location of interview								
Inside respondent's home	1,913	1.9	1.27, 2.62	1,361	2.3	1.28, 3.35	0.84	0.47–1.51
Outside respondent's home	743	3.1	1.50, 4.73	511	2.4	1.00, 3.79	1.31	0.59–2.89
Number of contact attempts								
1–2	1,310	2.0	1.16, 2.84	921	2.9	1.57, 4.24	0.68	0.35–1.34
3–4	746	1.5	0.50, 2.47	578	1.5	0.56, 2.45	0.99	0.43–2.29
5+	971	3.0	1.61, 4.43	656	2.4	0.89, 3.82	1.29	0.56–2.98
Case reassigned to different interviewer								
Yes	717	1.7	0.83, 2.61	522	1.8	0.44, 3.19	0.95	0.40–2.23
No	2,314	2.3	1.54, 3.15	1,640	2.5	1.65, 3.43	0.92	0.54–1.58
Householder(s) expressed privacy-related concerns								
Yes	525	1.4	0.54, 2.27	410	3.5	1.75, 5.19	0.40	0.19–0.83
No	2,492	2.4	1.66, 3.09	1,738	2.1	1.30, 2.99	1.11	0.67–1.84
Householder(s) expressed time-related concerns								
Yes	897	2.9	1.44, 4.27	662	2.9	1.32, 4.51	0.98	0.43–2.23
No	2,120	1.9	1.22, 2.62	1,486	2.2	1.24, 3.06	0.89	0.48–1.65

Note. CI = confidence interval; OR = odds ratio.

Table 7. Item nonresponse rate to sexual identity question by interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI			
	n	%	95% CI	n	%	95% CI	OR	95% CI	AOR[a]	95% CI
Item nonresponse rate to sexual identity question	3,150	2.7	2.00, 3.42	2,223	2.3	1.40, 3.12	1.20	0.77, 1.89	1.31	0.83, 2.07

Note. CI = confidence interval; OR = odds ratio; AOR = adjusted odds ratio.
[a]The following covariates were included in the multivariate logistic regression: age, sex, race/ethnicity, education, employment status, marital status, reported health status, family income, own/rent resident, MSA status, region, presence of others, location of interview, number of contact attempts, case reassigned to different interviewer, householder(s) expressed privacy concerns, and householder(s) expressed time-related concerns.

estimates were observed within 24 of 26 sociodemographic and interview environment subgroups. The two exceptions were the following: 1) adult respondents from higher income families, compared to adults from lower income families, were significantly more likely to identify as a sexual minority in ACASI compared to CAPI; and 2) among adults in households with privacy concerns, a higher percentage identified as a sexual minority in CAPI than they did in ACASI. This second finding is counterintuitive and neither of these exceptions to the pattern of null results can be easily explained. For both CAPI and ACASI the information is being entered into a computer, so concerns about electronic data security cannot explain the difference. Given the number of comparisons performed, these findings may simply reflect type I error. If we had set the p-value cutoff for statistical significance lower, to account for the multiple comparisons, we would have found no significant differences.

Assuming that significant mode differences reveal question sensitivity, these findings suggest that the sexual identity question was not considered sensitive by field test participants. As a further potential indication that respondents did not find the sexual identity question to be sensitive, there were only three breakoffs (i.e., respondent quit the survey) at the sexual identity question across the two modes: two in ACASI, one in CAPI.

There could be a number of explanations for the lack of mode differences found in sexual minority estimates, including features of the survey design and larger societal trends. From a question design perspective, it is important to recall that the use of a flashcard with the CAPI version of the sexual identity question was designed to maximize privacy in a face-to-face setting. By using a flashcard, the two parties can navigate the question without the respondent directly disclosing their sexual identity to the interviewer, or the interviewer reading the response options to the respondent. When the flashcard is handed to the respondent, they are asked to report the number on the card that corresponds to their answer, not their actual sexual identity. In both the ACASI and CAPI versions of the question, the question text does not utilize terminology or allude to the fact that the question is attempting to capture the respondent's sexual identity. To what extent this design minimized differences in estimates between CAPI and ACASI is difficult to measure. It is not always easy to use the flashcard, especially with interviews conducted on doorsteps or other difficult interviewing environments. Thus, it is unknown to what extent interviewers used the flashcard, even when the interview setting was conducive to its use. With that said, another mode comparison study that used a flashcard in its CAPI mode also

Table 8. Item nonresponse rate to sexual identity question by select sociodemographics and interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI	
	n	%	95% CI	n	%	95% CI	OR	95% CI
Age								
18–44	1,454	2.4	1.55, 3.27	1,035	1.8	0.75, 2.84	1.35	0.76, 2.39
45+	1,696	3.0	2.05, 3.91	1,188	2.7	1.56, 3.82	1.11	0.64, 1.92
Sex								
Male	1,422	2.7	1.67, 3.66	957	2.5	1.15, 3.95	1.05	0.52, 2.10
Female	1,728	2.8	1.80, 3.70	1,266	2.0	1.07, 2.92	1.39	0.79, 2.46
Race/ethnicity								
Non-hispanic white	1,847	1.4	0.79, 1.96	1,287	1.5	0.64, 2.37	0.91	0.45, 1.84
Other	1,303	5.4	3.91, 6.84	936	3.7	2.20, 5.30	1.46	0.94, 2.25
Education								
Less than a high school diploma/GED	1,344	4.2	2.89, 5.47	939	3.5	1.97, 5.01	1.21	0.75, 1.94
High school diploma/GED or more	1,799	1.6	0.92, 2.34	1,277	1.3	0.61, 1.97	1.27	0.62, 2.59
Employment status								
Working	1,789	2.1	1.38, 2.87	1,313	1.7	1.01, 2.47	1.23	0.72, 2.08
Not working	1,357	3.6	2.41, 4.76	909	3.1	1.65, 4.62	1.15	0.66, 2.01
Marital status								
Never married	738	2.6	1.42, 3.84	559	2.6	1.18, 4.05	1.00	0.47, 2.13
Other	2,404	2.7	1.92, 3.51	1,655	2.1	1.15, 3.09	1.29	0.76, 2.18
Reported health status								
Excellent/very good	1,768	2.3	1.52, 3.05	1,268	2.1	1.13, 2.98	1.11	0.62, 2.00
Good/fair/poor	1,381	3.3	2.18, 4.45	952	2.6	1.43, 3.73	1.29	0.76, 2.19
Family income								
< USD50,000	1,646	3.3	2.20, 4.48	1,184	2.8	1.30, 4.26	1.21	0.68, 2.14
≥ USD50,000	1,261	1.8	0.95, 2.68	850	1.3	0.50, 2.07	1.42	0.67, 3.03
Missing	243	4.4	2.25, 6.54	189	4.9	2.06, 7.64	0.90	0.40, 2.05
Own or rent residence								
Own or buying	1,839	2.0	1.32, 2.73	1,268	2.0	0.83, 3.25	0.99	0.50, 1.99
Rent or some other arrangement	1,303	4.0	2.68, 5.27	949	2.5	1.31, 3.79	1.58	0.96, 2.61
MSA status								
MSA, central city	1,104	3.8	2.21, 5.48	810	2.1	1.00, 3.21	1.86	0.98, 3.35
Other	2,046	2.2	1.49, 2.94	1,413	2.3	1.18, 3.48	0.95	0.54, 1.66
Region								
West	774	4.5	2.51, 6.39	600	2.9	1.59, 4.20	1.56	0.94, 2.61
Other	2,376	2.2	1.50, 2.92	1,623	2.1	1.01, 3.12	1.07	0.58, 1.96

Note. CI = confidence interval; OR = odds ratio.

Table 9. Item nonresponse rate to the sexual identity question by select interview environment measures and interview mode: NHIS sexual identity field test, 2012 (weighted).

	ACASI			CAPI			ACASI versus CAPI	
	n	%	95% CI	n	%	95% CI	OR	95% CI
Presence of others								
Yes	1,022	2.3	1.37, 3.30	732	1.8	0.75, 2.85	1.31	0.66, 2.61
Unknown	2,128	3.0	2.00, 3.96	1,491	2.6	1.50, 3.74	1.14	0.70, 1.87
Location of interview[a]								
Inside respondent's home	1,982	2.6	1.70, 3.41	1,385	1.4	0.72, 2.18	1.79	1.07, 2.99
Outside respondent's home	768	2.3	0.93, 3.68	528	3.1	1.19, 4.95	0.74	0.31, 1.81
Number of contact attempts								
1–2	1,358	2.6	1.57, 3.72	945	2.0	1.03, 2.98	1.32	0.71, 2.46
3–4	781	3.6	2.05, 5.19	596	3.4	0.94, 5.80	1.08	0.45, 2.60
5+	1,007	2.1	1.09, 3.07	675	1.7	0.77, 2.54	1.26	0.61, 2.59
Case reassigned to different interviewer								
Yes	745	2.7	1.42, 3.89	537	1.6	0.60, 2.57	1.69	0.78, 3.69
No	2,405	2.7	1.92, 3.53	1,686	2.5	1.47, 3.49	1.10	0.68, 1.79
Householder(s) expressed privacy-related concerns								
Yes	567	5.0	3.11, 6.79	428	4.6	1.81, 7.34	1.09	0.56, 2.12
No	2,569	2.2	1.56, 2.92	1,780	1.7	1.02, 2.44	1.30	0.77, 2.20
Householder(s) expressed time-related concerns								
Yes	942	2.9	1.58, 4.29	678	1.7	0.32, 3.01	1.78	0.70, 4.55
No	2,194	2.6	1.85, 3.41	1,530	2.5	1.52, 3.55	1.04	0.63, 1.71

Note. CI = confidence interval; OR = odds ratio.
[a]Telephone interviews are excluded from this measure.

found no significant difference between CAPI and CASI in the percentage of persons reporting a sexual minority identity (Malagoda and Traynor 2008).

The minimal differences by mode could potentially also be explained by decreased stigma associated with a sexual minority identity. If sexual minority respondents have less reason to fear embarrassment or reprisal from interviewers and/or third parties for revealing their sexual identity, they would have less reason to conceal it in face-to-face interviews. Unlike sexual behavior and attraction, which are still considered private matters, sexual identity is increasingly perceived as a standard demographic characteristic that can be shared in social contexts (Fredriksen-Goldsen and Muraco 2010; Rosenfeld 1999). Indeed, the phenomenon of "coming-out" (disclosing one's sexual identity) is a well-known and studied one (e.g., Legate et al. 2012; McGarrity and Huebner 2014). In contrast, there is no analogous widespread phenomenon of revealing one's sexual attractions or activities to colleagues or family members. In addition, although nationally representative trend data is scarce, research using convenience samples suggest that LGB people may be coming out earlier and in more contexts today than they did in the past (Pew Research Center 2013; Floyd and Bakeman 2006). As a result of these changes, the sensitivity of sexual identity questions may have decreased. Alternatively, it may be that the public's declining trust in the ability of the government, or any institution, to keep computerized information secure and confidential means that even an ACASI instrument no longer gives the assurance of privacy that it once did, and both the ACASI and CAPI estimates are underestimates.

Such a discrepancy between the sensitivity of questions on sexual identity versus those on attraction and behavior could explain why studies examining sexual attraction reporting by mode (Caltabiano and Dalla-Zuanna, 2012; Villarroel et al. 2006) and some studies examining sexual behavior by mode (e.g., Villarroel et al. 2006; Potdar and Koenig 2005) found differences by mode. Likewise, in the two studies which found differences in sexual identity reporting by mode (Midanik and Greenfield, 2008; Ghanem et al. 2005), the survey instruments also asked about sexual behavior, and that implicit linking of identity to behavior may explain the mode effects found in those studies and not found here. The temporal differences between those studies and the present one could also explain the difference.

5.2. Item Nonresponse

While few empirical assessments of the sensitivity-item nonresponse link have been performed, it is widely assumed that sensitive items produce more item nonresponse (Tourangeau and Yan 2007), an assumption largely attributable to high refusal rates to income questions (Dahlhamer et al. 2003; Dahlhamer et al. 2004). The very limited literature examining this issue with sexual orientation questions has found either lower or roughly equal item nonresponse in self-administered modes compared with interviewer-administered surveys (Caltabiano and Dalla-Zuanna 2012; Kurth et al. 2004; Villarroel et al. 2006). As we found for sexual minority reporting, there was no overall difference in item nonresponse to the sexual identity question by mode (CAPI = 2.3%, ACASI = 2.7%). In addition, while slightly higher sexual identity item nonresponse rates were observed in ACASI across a number of the sociodemographic and interview environment subgroups, none of the differences reached statistical significance.

Since results indicated a lack of perceived sensitivity, the lack of significant differences in item nonresponse by mode is also consistent with these results. However, we also believe the design of the ACASI screens for this field test contributed to the lack of significant mode differences in item nonresponse rates (Dahlhamer et al. 2013). An "implicit filter" design (Derouvray and Couper 2002) was adopted whereby "don't know" and "refused" options were not presented on the screen. The goal was to mimic a face-to-face interview as closely as possible. For example, an interviewer in the CAPI setting does not provide explicit "don't know" and "refused" options to a respondent. If the respondent attempted to skip the question in ACASI (pressing Enter without selecting a response), he/she was routed to a follow-up question that provided an option to return to the main question. Refused and don't know options were also provided on-screen. Here again, the design attempted to mimic CAPI interviewing. In the face-to-face setting, interviewers are trained to probe respondents in an attempt to convert don't know responses.

5.3. *Limitations*

This study was subject to at least six limitations. First, we lacked the statistical power necessary to detect modest differences in sexual minority estimates by mode. Given effective sample sizes and a sexual minority estimate of 2.4% in CAPI, we had less than 20% power to detect a half percentage point difference in estimates by mode (80% power to detect a difference of 1.8 percentage points). The problem was further compounded for subgroup comparisons. Second, the small number of adults identifying as a sexual minority in this study precluded us from exploring mode differences in gay/lesbian responses separate from bisexual responses. Third, while the CAPI and ACASI instruments were available in Spanish, there was insufficient sample to explore associations between mode of interview and sexual minority status by language of interview. Fourth, we approached the field test and subsequent data analysis from an "intent to treat" perspective. For roughly 17% of cases where the sexual identity module was completed, the questions were asked in a mode different from the one assigned. While separate analyses removing these cases revealed no substantive differences in the conclusions reported here, we cannot say with certainty that our results would be the same if strict adherence to the experimental protocol had been maintained. Fifth, the use of a flashcard with the sexual identity question in the CAPI administration may have afforded the CAPI respondent a level of privacy approaching that of ACASI, contributing to the null findings of this study. However, we do not have information on how often the flashcard was actually used. And sixth, both unit and item nonresponse may have influenced the results reported here. The final sample adult response rates for the two mode paths were 64.9% (CAPI) and 64.0% (ACASI), respectively. It is possible that subgroups who are more sensitive to mode effects with regard to sexual minority reporting may have had lower propensities to respond to the survey as a whole or to the sample adult interview. Nonresponse rates to the sexual identity question were 2.3% for CAPI and 2.7% for ACASI, rates that were similar to or higher than the percentage of adults identifying as a sexual minority. Furthermore, item nonresponse to the sexual identity question was lower among better educated, employed, and more affluent adults,

which may have minimized the effect of survey mode in this test. As Drydakis (2014) notes, sexual minorities with higher socioeconomic status may be more open and forthcoming with their sexual identity.

5.4. Conclusion

In conclusion, the findings from this field test may be of interest for other general health surveys that want to add questions on sexual orientation, but cannot afford the additional costs that may come with more complex mixed-mode designs. Our results suggest that one or more questions on sexual identity can be integrated within the existing design structure of such surveys with little implication for social desirability bias. With that said, this research represents one of a small number of studies that have attempted to isolate the effects of administration mode on responses to a question on sexual identity. Future experimental research could attempt to replicate our findings and/or extend it by exploring other interviewer- and self-administered modes (e.g., CATI, T-ACASI, web surveys) and/or experimentally-manipulating the context in which sexual identity questions are asked.

6. References

Aquilino, W. 1994. "Interview Mode Effects in Surveys of Drug and Alcohol Use." *Public Opinion Quarterly* 58: 210–240. DOI: http://doi.org/10.1086/269419.

Bernstein, R., A. Chadha, and R. Montjoy. 2001. "Over-reporting Voting: Why it Happens and Why it Matters." *Public Opinion Quarterly* 65: 22–44. DOI: http://doi.org/10.1086/320036.

Bradburn, N. 1983. "Response Effects." In *Handbook of Survey Research*, edited by P. Rossi, J. Wright, and A. Anderson, 289–328. New York: Academic Press.

Brazier, J.E., R. Harper, N.M.B. Jones, A. O'Cathain, K.J. Thomas, T. Usherwood, and L. Westlake. 1992. "Validating the SF-36 Health Survey Questionnaire: New Outcome Measure for Primary Care." *British Medical Journal* 305(6846): 160–164. DOI: http://doi.org/10.1136/bmj.305.6846.160.

Caltabiano, M. and G. Dalla-Zuanna. 2012. "A Comparison of Survey Techniques on Sensitive Sexual Behavior in Italy." *Journal of Sex Research* 50(6): 537–547. DOI: http://doi.org/10.1080/00224499.2012.674573.

Couper, M.P., E. Singer, and R. Tourangeau. 2003. "Understanding the Effects of Audio-CASI on Self-Reports of Sensitive Behavior." *Public Opinion Quarterly* 67: 385–395. DOI: http://doi.org/10.1086/376948.

Dahlhamer, J., J. Dixon, P. Doyle, J. Eargle, D.H. Griffin, and P. McGovern. 2003. "Quality at the Item Level: Terms, Methods, and Guidelines for Cross-Survey Comparisons." In *Proceedings of the Federal Committee on Statistical Methodology Research Conference*, Sheraton Crystal City Hotel, VA (November 2003): 17–19. Available at: https://www.researchgate.net/profile/James_Dahlhamer/publication/265540556_Quality_at_the_Item_Level_Terms_Methods_and_Guidelines_for_Cross-Survey_Comparisons_1/links/54b441750cf26833efd0123b.pdf (accessed November 2019).

Dahlhamer, J., J. Dixon, P. Doyle, J. Eargle, and P. McGovern. 2004. "Quality at the Item Level: Decomposing Item and Concept Response Rates." *Proceedings of the European*

Conference on Quality and Methodology in Official Statistics (Q2004), May 24–26, Mainz, Germany.

Dahlhamer, J.M., A. Galinsky, S. Joestl, M. Cynamon, J. Madans, and V. Cain. 2013. "Minor Design Changes with Major Impacts: Testing Explicit Versus Implicit Don't Know and Refused Options in Audio Computer-Assisted Self-Interviewing." Presented at the AAPOR 68th Annual Conference, May 17. Boston, Massachusetts, U.S.A.

De Leeuw, E.D. and J. van der Zouwen. 1988. "Data Quality in Telephone and Face-to-Face Surveys: A Comparative Meta-Analysis." In *Telephone survey methodology*, edited by R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. II, Nicholls, and J. Waksberg, 283–299. New York: Wiley.

Des Jarlais, D.C., D. Paone, J. Milliken, C.F. Turner, H. Miller, J. Gribble, Q. Shi, H. Hagan, and S.R. Friedman. 1999. "Audio-Computer Interviewing to Measure Risk Behaviour for HIV among Injecting Drug Users: A Quasi-Randomised Trial." *The Lancet* 353(9165): 1657–1661. DOI: http://doi.org/10.1016/S0140-6736(98)07026-3.

Derouvray, C. and M.P. Couper. 2002. "Designing a Strategy for Reducing 'No Opinion' Responses in Web Surveys." *Social Science Computer Review* 20(1): 3–9. DOI: http://doi.org/10.1.1.877.199.

Dialsingh, I. 2008. "Face-to-Face Interviewing." In *Encyclopedia of Survey Research Methods*, edited by P. Lavrakas, 260–262. Thousand Oaks, CA: SAGE Publications, Inc. DOI: http://doi.org/10.4135/9781412963947.

Dolezal, C., S.L. Marhefka, E.K. Santamaria, C-S. Leu, E. Brackis-Cott, and C.A. Mellins. 2012. "A Comparison of Audio Computer-Assisted Self-Interviews to Face-to-Face Interviews of Sexual Behavior among Perinatally HIV-Exposed Youth." *Archives of Sexual Behavior* 41: 401–410. DOI: http://doi.org/10.1007/s10508-011-9769-6.

Drydakis, N. 2014. "Sexual Orientation and Labor Market Outcomes." *IZA World of Labor* 111: 1–10. DOI: http://doi.org/10.15185/izawol.111.

Floyd, F.J. and R. Bakeman. 2006. "Coming-Out across the Life Course: Implications of Age and Historical Context." *Archives of Sexual Behavior* 35(3): 287–296. DOI: http://doi.org/10.1007/s10508-006-9022-x.

Fowler, F. 1993. *Survey Research Methods*. Newbury Park, CA: Sage Publications.

Fowler, F. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks, CA: Sage Publications.

Fredriksen-Goldsen, K.I. and A. Muraco. 2010. "Aging and Sexual Orientation: A 25-Year Review of the Literature." *Research on Aging* 32(3): 372–413. DOI: http://doi.org/10.1177/0164027509360355.

Fu, H., J.E. Darroch, S.K. Henshaw, and E. Kolb. 1998. "Measuring the Extent of Abortion Underreporting in the 1995 National Survey of Family Growth." *Family Planning Perspectives* 30: 128–133 & 138. DOI: http://doi.org/10.2307/2991627.

Ghanem, K.G., H.E. Hutton, J.M. Zenilman, and E.J. Erbelding. 2005. "Audio Computer Assisted Self Interview and Face to Face Interview Modes in Assessing Response Bias among STD Clinic Patients." *Sexually Transmitted Infections* 81: 421–425. DOI: http://doi.org/10.1136/sti.2004.013193.

Gfroerer, J. and A. Hughes. 1992. "Collecting Data on Illicit Drug Use by Phone." In *Survey Measurement of Drug Use: Methodological Studies*, edited by C. Turner,

J. Lessler, and J. Gfroerer, 277–295. Washington, DC: U.S. Government Printing Office.

Groves, R.M., F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2004. *Survey Methodology*. Hoboken, NJ: John Wiley & Sons, Inc. DOI: http://doi.org/10.2307/27590808.

Healthy People 2020. (n.d.) "Lesbian, Gay, Bisexual, and Transgender Health." Available at: https://www.healthypeople.gov/2020/topics-objectives/topic/lesbian-gay-bisexual-and-transgender-health (accessed November 2019).

Hox, J.J. and E.D.D. de Leeuw. 1994. "A Comparison of Nonresponse in Mail, Telephone, and Face-to-Face Surveys." *Quality and Quantity* 28(4): 329–344. DOI: http://doi.org/10.1007/BF01097014.

IOM (Institute of Medicine). 2011. *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press. DOI: http://doi.org/0.17226/13128.

Jaya, M.J.H., M.J. Hindin, and S. Ahmed. 2008. "Differences in Young Peoples' Reports of Sexual Behaviors According to Interview Methodology: A Randomized Trial in India." *American Journal of Public Health* 98(1): 169–174. DOI: http://doi.org/10.2105/AJPH.2006.099937.

Juster, F.P. and J.T. Smith. 1997. "Improving the Quality of Economic Data: Lessons from HRS and AHEAD." *Journal of the American Statistical Association* 92: 1268–1278. DOI: http://doi.org/10.1080/01621459.1997.10473648.

Kurth, A.E., D.P. Martin, M.R. Golden, N.S. Weiss, P.J. Heagerty, F. Speilberg, H.H. Handsfield, and K.K. Holmes. 2004. "A Comparison between Audio Computer-Assisted Self-Interviews and Clinician Interviews for Obtaining the Sexual History." *Sexually Transmitted Diseases* 31(12): 719–726. DOI: http://doi.org/10.1097/01.olq.0000145855.36181.13.

Legate, N., R.M. Ryan, and N. Weinstein. 2012. "Is Coming Out Always a "Good Thing"? Exploring the Relations of Autonomy Support, Outness, and Wellness for Lesbian, Gay, and Bisexual Individuals." *Social Psychological and Personality Science* 3(2): 145–152. DOI: http://doi.org/10.1177/1948550611411929.

Malagoda, M. and J. Traynor. 2008. "Developing Survey Questions on Sexual Identity: Report on National Statistics Omnibus Trial 4." London: Social Economic Micro-Analysis and Reporting Division, Office for National Statistics.

McDonald, M.P. 2003. "On the Over-Report Bias of the National Election Study Turnout Rate." *Political Analysis* 11: 180–186. DOI: http://doi.org/10.1093/pan/mpg006.

McGarrity, L.A. and D.M. Huebner. 2014. "Is Being Out about Sexual Orientation Uniformly Healthy? The Moderating Role of Socioeconomic Status in a Prospective Study of Gay and Bisexual Men." *Annals of Behavioral Medicine* 47(1): 28–38. DOI: http://doi.org/10.1007/s12160-013-9575-6.

Midanik, L.T. and T.K. Greenfield. 2008. "Interactive Voice Response Versus Computer-Assisted Telephone Interviewing (CATI) Surveys and Sensitive Questions: The 2005 National Alcohol Survey." *Journal of Studies on Alcohol and Drugs* 69: 580–588. DOI: http://doi.org/10.15288/jsad.2008.69.580.

Miller, K. and J.M. Ryan. 2011. *Design, Development and Testing of the NHIS Sexual Identity Question*. Hyattsville, MD: Cognitive Testing Report, Questionnaire Design Research

Laboratory, Office of Research and Methodology, National Center for Health Statistics. Available at: https://wwwn.cdc.gov/qbank/report/Miller_NCHS_2011_NHIS%20Sexual%20Identity.pdf (accessed June 2013).

Moore, J., L. Stinson, and E. Welniak. 1999. "Income Reporting in Surveys: Cognitive Issues and Measurement Error." Chapter 10 in *Cognition and Survey Research*, edited by M. Sirkin, D. Herrmann, S. Schechter, N. Schwarz, J. Tanur, and R. Tourangeau. New York: John Wiley & Sons.

Pew Research Center. 2013. "A Survey of LBGT Americans: Attitudes, Experiences and Values in Changing Times." Washington, DC: The Pew Research Center. Available at: https://www.pewsocialtrends.org/files/2013/06/SDT_LGBT-Americans_06-2013.pdf (accessed June 2014).

Potdar, R. and M.A. Koenig. 2005. "Does Audio-CASI Improve Reports of Risky behavior? Evidence from a Randomized Field Trial among Young Urban Men in India." *Studies in Family Planning* 36(2): 107–116. DOI: http://doi.org/10.1111/j.1728-4465.2005.00048.x.

Rogers, S.M., G. Willis, A. Al-Tayyib, M.A. Villarroel, C.F. Turner, L. Ganapathi, J. Zenilman, and R. Jadack. 2005. "Audio Computer Assisted Interviewing to Measure HIV Risk Behaviors in a Clinic Population." *Sexually Transmitted Infections* 1: 501–507. DOI: http://doi.org/10.1136/sti.2004.014266.

Rosenfeld, D. 1999. "Identity Work among Lesbian and Gay Elderly." *Journal of Aging Studies* 13(2): 121–144. DOI: http://doi.org/10.1016/s0890-4065(99)80047-4.

Shoemaker, P., M. Eicholz, and E. Skewes. 2002. "Item Nonresponse: Distinguishing between Don't Know and Refuse." *International Journal of Public Opinion Research* 14: 193–201. DOI: http://doi.org/10.1093/ijpor/14.2.193.

Simoes, A.A., F.I. Bastos, R.I. Moreira, K.G. Lynch, and D.S. Metzger. 2006. "A Randomized Trial of Audio Computer and In-Person Interview to Assess HIV Risk among Drug and Alcohol Users in Rio De Janeiro, Brazil." *Journal of Substance Abuse Treatment* 30: 237–243. DOI: http://doi.org/10.1016/j.jsat.2005.12.002.

SMART (Sexual Minority Assessment Research Team). 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys*. Los Angeles: The Williams Institute, University of California School of Law. Available at: https://escholarship.org/uc/item/706057d5 (accessed November 2019).

Sykes, W. and M. Collins. 1988. "Effect of Mode of Interview: Experiments in the UK." In *Telephone Survey Methodology*, edited by R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, and J. Waksberg, 301–320. New York: Wiley.

Tideman, R.L., M.Y. Chen, M.K. Pitts, S. Ginige, M. Slaney, and C.K. Fairley. 2007. "A Randomised Control Trial Comparing Computer-Assisted with Face-to-Face Sexual History Taking in a Clinical Setting." *Sexually Transmitted Infections* 83: 52–56. DOI: http://doi.org/10.1136/sti.2006.020776.

Tourangeau, R., L. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press. DOI: http://doi.org/10.1017/CBO9780511819322.

Tourangeau, R. and T.W. Smith. 1996. "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context." *Public Opinion Quarterly* 60: 275–304. DOI: http://doi.org/10.1086/297751.

Tourangeau, R. and T. Yan. 2007. "Sensitive Questions in Surveys." *Psychological Bulletin* 133(5): 859–883. DOI: http://doi.org/10.1037/0033-2909.133.5.859.

Turner, C.F., J.T. Lessler, and J. Devore. 1992. "Effects of Mode of Administration and Wording on Reporting of Drug Use." In *Survey Measurement of Drug Use: Methodological Studies*, edited by C. Turner, J. Lessler, and J. Gfroerer, 177–220. Rockville, MD: National Institute on Drug Abuse.

Villarroel, M.A., C.F. Turner, E. Eggleston, A. Al-Tayyib, S.M. Rogers, A.M. Roman, P.C. Cooley, and H. Gordek. 2006. "Same-Gender Sex in the United States: Impact of T-ACASI on Prevalence Estimates." *Public Opinion Quarterly* 70(2): 166–196. DOI: http://doi.org/10.1093/poq/nfj023.

Received March 2018
Revised October 2018
Accepted February 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 835–858, http://dx.doi.org/10.2478/JOS-2019-0035

Measuring Sexual Orientation and Gender Identity in the National Crime Victimization Survey

Jennifer L. Truman[1], *Rachel E. Morgan*[1], *Timothy Gilbert*[2], *and Preeti Vaghela*[2]

The National Crime Victimization Survey (NCVS) collects information on nonfatal personal and property crimes both reported and not reported to police. As part of the ongoing redesign efforts for the NCVS, the Bureau of Justice Statistics (BJS) added sexual orientation and gender identity (SOGI) questions to the survey's demographic section in July 2016. The inclusion of these measures will provide important national-level estimates of victimization among lesbian, gay, bisexual, and transgender (LGBT) people and allow researchers to understand victimization risk and access to victim services. This article includes a discussion of the sexual orientation and gender identity measures that were added to the NCVS, and findings from the monitoring activities conducted during the first six months of data collection. In addition, population counts by sexual orientation and gender identity are estimated using July through December 2016 NCVS data.

Key words: National crime victimization survey; sexual orientation; sexual identity; gender identity; victimization.

1. Introduction

There is a growing interest in understanding the national status of the lesbian, gay, bisexual, and transgender (LGBT) population across key indicators of social, health, and economic well-being (IOM 2011; SOGI Federal Working Group 2016a, 2016b, 2016c). Historically, few national surveys have collected data on sexual orientation and gender identity. Research on LGBT persons is developing in the health and social fields, specifically in the US Federal Statistical System, with the addition of sexual orientation measures to the National Health Interview Survey (NHIS), both sexual orientation and gender identity measures on the Survey of Prison Inmates (SPI), Behavioral Risk Factor Surveillance System (BRFSS), and Population Assessment of Tobacco and Health (PATH), and the potential inclusion of sexual orientation and gender identity (SOGI) measures to the Current Population Survey (CPS) (Dahlhamer et al. 2014; Ellis et al. 2017; SOGI Federal Working Group 2016a, 2016b). However, sexual orientation and gender identity have been identified in other research as correlates of victimization, and national-level data are needed on the criminal victimization experiences of LGBT people.

[1] Bureau of Justice Statistics, 810 7th Street, N.W., Washington, D.C. 20531, U.S.A. Emails: Jennifer.Truman@usdoj.gov, and Rachel.Morgan@usdoj.gov
[2] U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746, U.S.A. Emails: Timothy.R.Gilbert@census.gov, and Preeti.Vaghela@census.gov
Acknowledgments: The authors would like to thank Lynn Langton, Meagan Meuchel, and Meg Ruhnke for their thoughtful review of this article. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position and policies of the Bureau of Justice Statistics (BJS) or the U.S. Census Bureau.

As one of two primary sources of information on the nature of criminal victimization incidents in the United States, the National Crime Victimization Survey (NCVS) is a perfect vehicle for collecting information on the victimization experiences of LGBT persons. (The other primary source of information on criminal victimization in the United States is the Federal Bureau of Investigation's Uniform Crime Reporting program.) In 2016, after substantial research and testing, measures of SOGI were added to the NCVS (Martinez et al. 2017). The estimates produced through the survey will provide researchers and policy makers with information on the types of victimization experienced by the LGBT population and their access to victim services. Measuring sexual orientation and gender identity on the NCVS also provides data on other types of victimization experienced by LGBT people, including identity theft and stalking; and their interactions with law enforcement using data from the NCVS supplemental surveys (i.e., NCVS Identity Theft Supplement, Supplemental Victimization Survey, and Police Public Contact Survey). Overall, the inclusion of sexual orientation and gender identity in the NCVS provides more accurate and detailed data that can be used to inform public policy regarding this vulnerable population.

LGBT persons are at risk of experiencing certain types of victimization at a disproportionately higher rate or the same rate as their heterosexual peers. In particular, for both women and men, LGBT persons report intimate partner violence and sexual violence at rates equal to or higher than heterosexual women and men (Krebs et al. 2016; Walters et al. 2013). Transgender persons experience intimate partner and sexual violence at higher rates than those who do not identify as transgender (Krebs et al. 2016; National Coalition of Anti-Violence Programs 2016). Sexual minority youth disproportionately experience health risks, including violence; and are at risk of peer victimization related to their sexual orientation and gender identity or expression (Collier et al. 2013; Kann et al. 2011).

Important changes in federal laws related to protecting LGBT survivors of violence have occurred in recent years, including the Violence Against Women Reauthorization Act (VAWA) of 2013. VAWA sought to improve care and access to victim services for LGBT victims and explicitly prohibited discrimination of victims or survivors of violence based on actual or perceived SOGI status, which works to ensure access to key services (VAWA 2013, Pub. L. No. 113-4, Stat. 47). VAWA also identified LGBT victims as an underserved community, which allowed organizations to receive more funding to focus on LGBT domestic violence, dating violence, sexual assault, and stalking victims. In addition, The Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act of 2009 (HCPA) included language to allow for prosecution of hate crimes committed against persons based on their actual or perceived sexual orientation or gender identity (HCPA, 18 U.S.C. § 249). Measuring SOGI on the NCVS provides important data on victimization, access to victim services, and experiences of hate crimes to support these laws.

As research continues to develop in this area, federal data are needed to expand the knowledge on criminal victimization of the LGBT population. The addition of these measures to the NCVS provides important national-level estimates of victimization among LGBT people and allows researchers to understand victimization risk and access to victim services. This article will address the following research questions: (1) How did interviewers and respondents react to SOGI questions asked in the context of a crime survey; (2) What was item nonresponse for the SOGI questions, and how did this vary by demographic characteristics; and (3) How do SOGI population estimates collected on a

crime survey differ from other types of population surveys, specifically health-related surveys? It includes a discussion of the sexual orientation and gender identity measures that were added to the NCVS, and findings from the monitoring of data collection activities. In addition, population counts by sexual orientation and gender identity are estimated using July through December 2016 NCVS data.

2. Methodology

2.1. Timeline of Pretesting, Implementation, and Monitoring of SOGI Data Collection

In the fall of 2015, the Center for Survey Measurement (CSM) at the U.S. Census Bureau conducted cognitive testing of the proposed sexual orientation and gender identity questions for the NCVS (Figure 1). For more information about the cognitive interviews, please review the report by the U.S. Census Bureau (Martinez et al. 2017). Data collection of these questions began in July 2016. The U.S. Census Bureau conducted a debriefing questionnaire (August 2016 to September 2016), focus groups (September 2016), and targeted interviews (October 2016) with NCVS interviewers who collected SOGI data. Additional monitoring of the SOGI data has been conducted from November 2016 to the present.

2.2. SOGI Data Collection Production Interviews in the 2016 NCVS

The inclusion of the SOGI questions in the NCVS began in July 2016. Before administering these new items, interviewers completed a self-study training to introduce them to the new items in the NCVS instrument and allow them to practice with the items before their first interview. In addition to training, the NCVS computer-assisted personal interviewing (CAPI) instrument includes a Frequently Asked Questions (FAQs) section for interviewers if respondents have questions about the SOGI items, including why the questions are important and relevant to a crime survey and definitions of each of the concepts (see Appendix 1, Section 5). The SOGI questions were placed at the end of the interview in a section with questions on disability, citizenship, veteran status, and household income. Once the questions were in the field, the U.S. Census Bureau and BJS began monitoring responses, refusal rates, and any information reported by interviewers.

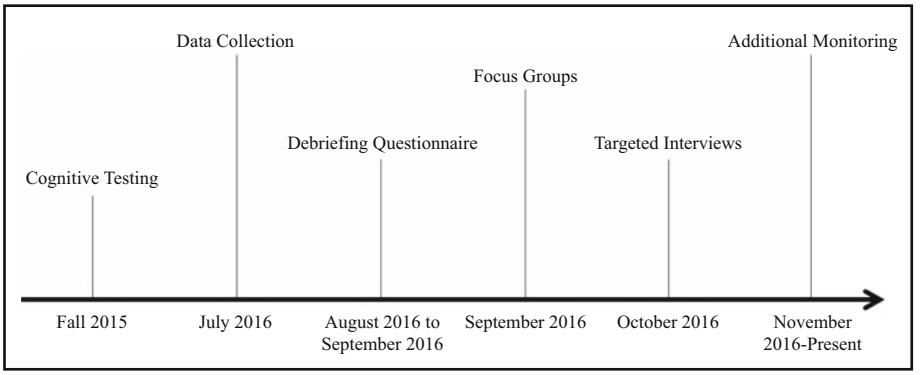


Fig. 1. Timeline of pretesting, implementation, and monitoring of SOGI data collection.

2.3. *Defining Sexual Orientation and Gender Identity*

Sexual orientation is defined by three dimensions: sexual identification (identity), sexual attraction, and sexual behavior. The measure used in the NCVS focuses on sexual identity. *Sexual identification (identity)* refers to the way a person identifies with a given sexual orientation (SMART 2009; SOGI Federal Working Group 2016a). The most commonly used terms to describe sexual orientation include lesbian, gay, bisexual, and heterosexual/straight (SOGI Federal Working Group 2016a). *Sexual attraction* refers to the relationship between a person's gender and the gender of the person(s) that they feel attracted to. *Sexual behavior* refers to the relationship between a person's gender and the gender of the individual(s) with whom they engage in sexual activity. Some measures of sexual orientation include all three dimensions of the concept and some only focus on sexual identity.

Like sexual orientation, gender identity is comprised of several dimensions, specifically differences between the concepts of sex and gender. *Sex* is an individual's biological classification at birth as either male or female (IOM 2011; SOGI Federal Working Group 2016a). *Gender* is socially constructed and based on how the individual presents to society, as either male or female, and encompasses the concepts of gender identity and gender expression (SOGI Working Group 2016a). *Gender identity* refers to a person's internal sense of gender, while *gender expression* is the way one sees themselves or how they present their gender to society (SOGI Working Group 2016b). An individual's sex and gender may be consistent (cisgender) or may be different (transgender).

2.4. *Measures of Sexual Orientation and Gender Identity*

2.4.1. Sexual Orientation

The sexual orientation questions that were administered in the NCVS came from the National Health Interview Survey (NHIS), conducted by the National Center for Health Statistics (NCHS 2018), and measure sexual identity. These questions are also consistent with recommendations on measuring sexual orientation made by the Sexual Minority Assessment Research Team (SMART 2009). The NHIS questions had been previously tested using cognitive interviews and have performed well with persons age 18 or older. The question used in the 2016 NCVS had also been tested using cognitive interviews (Martinez et al. 2017), and is as follows:

Sexual orientation question

1. *Which of the following best represents how you think of yourself?*
   - ○ *[Lesbian or] Gay*
   - ○ *Straight, that is, not [lesbian or] gay*
   - ○ *Bisexual*
   - ○ *Something else*
   - ○ *I don't know the answer*
   - ○ *REFUSED*

The phrase "lesbian or" is only displayed and read if the respondent had been assigned female on the household roster, and answer categories displayed in all capital letters are not read aloud.

2.4.2. Gender Identity

The gender identity questions that were administered in the NCVS come from recommendations from the Gender Identity in U.S. Surveillance (GenIUSS) group and the California Health Interview Survey (CHIS 2018; GenIUSS Group 2014). The current recommendation for measuring gender identity is to use a two-step approach that asks about assigned sex at birth and current gender identity (GenIUSS Group 2014). Two questions are used to classify respondents as transgender or cisgender; and this method has been successful in identifying transgender individuals compared to single questions (Reisner et al. 2014; Tate et al. 2013; Xavier 2000; Xavier et al. 2007). In addition, using the two-step approach allows those transgender individuals who identify their current gender as male or female and not as transgender to identify as such, but still be classified as transgender using the two-step approach. In the NCVS, persons were identified as transgender if their responses to sex at birth and current gender identity were different, but not if they said "none of these" on the gender identity question. These gender identity questions had been previously tested using cognitive interviews and have performed well with persons age 18 or older, and were cognitively tested again for the NCVS (Martinez et al. 2017). The questions used in the 2016 NCVS are as follows:

Gender identity – assigned sex at birth question

1. *What sex were you assigned at birth, on your original birth certificate?*
   - ○ *Male*
   - ○ *Female*
   - ○ REFUSED
   - ○ DON'T KNOW

Gender identity – current gender identity question

2. *Do you currently describe yourself as male, female, or transgender?*
   - ○ *Male*
   - ○ *Female*
   - ○ *Transgender*
   - ○ *None of these*
   - ○ REFUSED
   - ○ DON'T KNOW

The respondent is asked the following confirmation question if they answer "male" and then "female," "transgender," or "none of these" to the gender identity questions or if they answer "female" and then "male," "transgender," or "none of these" to the gender identity questions.

Gender identity – confirmation question

3. *Just to confirm, you were assigned {FILL} at birth and now describe yourself as {FILL}. Is that correct?*
   - ○ *Yes*
   - ○ *No*
   - ○ REFUSED
   - ○ DON'T KNOW

The first fill in the confirmation question is populated with the answer given by the respondent to the first gender identity question (assigned sex at birth), and the fill options are "male" or "female." The second fill is populated with the answer given by the respondent to the second gender identity question (current gender identity). The fill options are "male," "female," or "transgender." Additionally, if the respondent answered "None of these" to the second gender identity question, then the phrase "describe yourself as {FILL}" is replaced with "do not describe yourself as male, female, or transgender."

If the respondent answers "No" to the confirmation question, then the survey instrument forces the interviewer to have the respondent change their answer to either the first or second gender identity question. The confirmation question will be asked again until the answer is "Yes" or the answers to the gender identity questions are not discordant.

2.5. *Analytical Strategy*

We used a mixed methods approach to answer our three research questions. To understand interviewer and respondent reactions to SOGI questions asked in the context of a crime survey *(research question 1)*, we solicited feedback from interviewers in three ways: a debriefing questionnaire, focus groups, and targeted interviews. The online debriefing questionnaire was sent to all Census Bureau interviewers working on the NCVS, and they responded to the questionnaire in August and September 2016. Approximately, 899 interviewers (77.3% response rate) reported completing at least one NCVS interview between July 1, 2016 and the time of the debriefing questionnaire, and completed the full debriefing questionnaire.

The debriefing questionnaire collected quantitative data about interviewer perceptions of instrument problems, and respondents' experience in and reactions to answering SOGI questions. Interviewers were also allowed to elaborate about their experiences with administrating the questions using a write-in response for comments. All write-in comments for each question were analyzed using grounded theory (for more information, see Charmaz 2006). Initial codes were first created from line by line coding by two independent coders. Then the independent coders developed theoretical memos, which are reports written by the researcher that document their thoughts about the individual codes and how codes can be related to each other. From these memos, the coders were able to generate themes. Then the coders independently coded each response for the agreed-upon themes, met to review any inconsistencies, and recoded responses as necessary. For

each response that had a difference in coding, the coders discussed why they coded each response with their specific theme until they both agreed on the same theme for the response. The researchers then conducted theoretical sampling. Theoretical sampling refers to a sample that is selected to gather more information or to develop a theme. This theoretical sampling took the form of focus groups and targeted interviews with interviewers, and was used to collect more data and help refine themes.

Six different focus groups, each one hour in length, with select interviewers from across the country were conducted by conference call in August 2016, after a full month of data collection with the new questions. Each focus group included two interviewers from one of the six U.S. Census Bureau regional offices (ROs). Interviewers from each RO were eligible to participate if they had conducted more than the average number of NCVS interviews within their region (number of interviews ranged from 13 to 22 across region) in July 2016. At least one interviewer selected per RO worked in a state that had legislative actions regarding transgender issues. The topics discussed in the focus groups included the ease of administering the questions, respondent reactions to the questions, and any other comments the interviewers had about the questions. As the researchers facilitated the focus groups, they wrote notes when interviewers mentioned new or relevant themes. After each focus group session, the researchers typed up their notes about the focus groups. These notes were reviewed again by the coders for themes. The coders also developed memos to refine the themes.

Targeted interviews were also conducted with interviewers who had collected specific answers to the SOGI questions. Twenty interviews were conducted in October 2016 with individual interviewers who had collected answers of *lesbian, gay, bisexual, something else,* or *I do not know the answer* to the sexual orientation question; or *transgender, none of these,* or differing *male* and *female* responses to the gender identity questions. The focus of these interviews was understanding how the SOGI questions worked with LGBT respondents. The researchers took notes during the interviews, created a detailed write-up about the interview and the interview responses, and then these interview notes were reviewed by the coders for themes.

We examined item nonresponse by respondent sociodemographic characteristics and breakoffs for each item to understand item nonresponse for the SOGI questions in the NCVS *(research question 2)*. Logistic regression models were used to examine the likelihood of nonresponse to the SOGI questions across various sociodemographic characteristics (i.e., age, sex, race and Hispanic origin, educational attainment, household income, interview language, and mode of interview). These models were also stratified by Hispanic origin to determine if there were subgroup differences in likelihood of nonresponse to the SOGI questions. Many of these demographic characteristics are used to create post-stratification weights of NCVS data. Therefore, it was important to analyze their effect on nonresponse to the SOGI questions, because these are variables known to account for nonresponse in the data. We also stratified our results by race and Hispanic origin because previous research has found that racial minorities have higher rates of nonresponse to SOGI questions compared to non-Hispanic Whites (Jans et al. 2015; SOGI Federal Working Group 2016b). Finally, population totals were estimated to compare sexual orientation and gender identity estimates administered in the setting of a crime survey to estimates from other types of population surveys, specifically health-related surveys *(research question 3)*.

3. Results

3.1. Debriefing Questionnaire, Focus Groups, and Targeted Interviews

During the debriefing questionnaire, interviewers were asked to report any type of issue experienced while administering the SOGI questions to respondents. Interviewers were able to select all types of issues they encountered. The results from this debriefing questionnaire are not meant to make inferences about all interviewers, but merely to report the experiences of interviewers who responded to the debriefing questionnaire. Although rare, some interviewers reported both experiencing no issues, as well as a general option for other types of issues. The results indicate that interviewers felt the gender identity questions were easier to read than the sexual orientation question. About half of interviewers (52%) reported no issues with the gender identity questions, while a smaller percentage (37%) reported no issues when administering the sexual orientation question (Figure 2). About 50% of interviewers reported at least one respondent having a negative reaction to the sexual orientation question itself, while 39% of interviewers reported a negative reaction the gender identity questions. These findings should be interpreted with caution as it is unknown whether the problems were with a single respondent or a larger proportion of respondents interviewed by each interviewer. Interviewers may also misremember how often respondents reported a concern or may only focus on the concerns that were important or stood out to them. Additionally, a smaller percentage of interviewers reported that respondents had difficulty answering the gender identity questions (4%) than the sexual orientation question (9%).

A major theme that emerged from the qualitative analysis was that respondents questioned the relevance of sexual orientation and gender identity to crime. This theme included any mention that respondents questioned the reason for asking the questions,

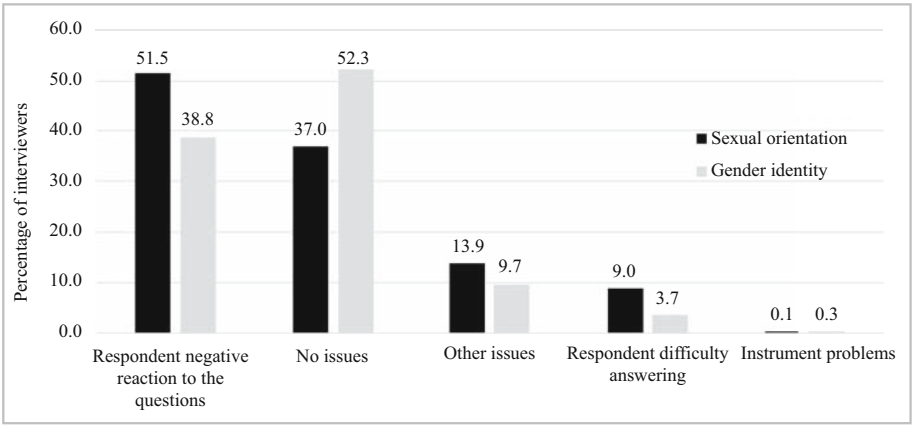


Fig. 2. Issues experienced by interviewers for sexual orientation and gender identity items.
Note: These are response options to the question: Have you experienced any of the following issues? Mark all that apply. Percentages will not add to 100% because respondents were allowed to select more than one answer. Interviewers reported these issues from at least one respondent. An example of other issues reported is respondents questioning the relevancy of the question to crime. Instrument problems refer to issues with the software an interviewer uses to administer the survey. N = 899.
Source: U.S. Census Bureau internal debriefing questionnaire for NCVS interviewers.

including simply questioning the relevance of these questions to experiencing crime and wanting to understand why the government was interested in this information. FAQs were included in the CAPI instrument to aid interviewers in addressing this concern with respondents. The FAQs specify that sexual orientation and gender identity are correlated with victimization, and the questions are included to better understand this relationship. In addition, the FAQs mention that discrimination against persons because of their sexual orientation or gender identity is prohibited by federal hate crime statutes and the 2013 reauthorized VAWA; and the inclusion of these items allows researchers to better address policy-relevant questions about hate crime victimization and victim services. Overall, the inclusion of the FAQs resulted in positive reactions from the respondents and assisted interviewers in being able to address any concerns. Upon hearing these questions, some respondents also tried to answer the relevancy of these questions themselves. They attributed the asking of these questions to current events involving transgender individuals, to politics, or to the change in the cultural discourse around LGBT issues.

A second theme that emerged was some respondents felt the interviewers should have been able to tell their sexual orientation or gender identity by just looking at them, and expressed discomfort at the perceived suggestion of not being straight. These respondents mentioned that the answer to these questions should be obvious without having to ask, and that asking these questions indicated that they might not present as straight or cisgender.

Although the majority of respondents did not have difficulty understanding the question, interviewers remarked that some respondents had negative or emphatic reactions to the content of the question, which indicates that some respondents may have been uncomfortable answering. In particular, some older adults had negative reactions to the gender identity questions, and expressed discomfort at answering the sexual orientation question or did not know how to answer because they were confused by the terms. Additionally, some men (across all age ranges) answered emphatically that they were straight and male. A smaller portion of respondents felt some level of discomfort about answering as *lesbian*, *gay*, or *bisexual*, and hesitated slightly before answering. Some interviewers themselves expressed concerns that the sexual orientation question would impact response rates to future interviews, but these comments were only given by about 1% of all interviewers who answered the full debriefing questionnaire.

Interviewers did state that some respondents had positive reactions to the questions. The sexual orientation question was easy for LGB respondents to understand and answer; some respondents had already divulged their sexual orientation earlier in the interview, and many just answered the question matter-of-factly. Many of these respondents had positive reactions to the question, stating that they were "thrilled" and "appreciated that the question was included" in the survey.

The focus groups also gathered information from interviewers about the *something else* response category for the sexual orientation question. About 0.22% of all respondents age 16 or older selected this response category and the NCVS instrument did not collect any additional information when a respondent chose this answer. In the targeted interviews, interviewers clarified that respondents who chose *something else* tended to move on with the interview without voluntarily providing additional information about why they chose that response. However, some interviewers perceived that English-speaking respondents who identified as something else may not have wanted to disclose their sexual orientation

or felt that their sexual orientation was not captured by the categories presented (queer, pansexual, asexual, etc.). During Spanish-speaking interviews, respondents who selected an answer of *something else* generally needed the interviewer to repeat the question and usually responded that they were "normal." This indicates that respondents who answered *something else* in Spanish might have experienced confusion about the terminology used. Therefore, when some straight respondents were asked about their sexual orientation in a Spanish-speaking interview, those not familiar with the term "straight" selected the *something else* response option.

Interviewers also stated that there were concerns related to the *none of these* response category in the current gender identity question. About 0.17% of all respondents age 16 or older selected this response category and the NCVS instrument did not collect any additional information when a respondent chose this answer. In the targeted interviews, interviewers reported mixed reasons for the use of the *none of these* response category. In some instances, interviewers sensed that this answer was the result of respondents being generally offended and not wanting to answer the gender identity questions, rather than describing themselves as something other than male, female, or transgender. The data from the focus groups and targeted interviews with interviewers suggested that the *none of these* response category may have been marked by interviewers in instances when respondents did not really want to answer the question but did not outright refuse. This issue was discovered early in data collection, and messages were sent to interviewers to reinforce the proper use of the *none of these* response category. However, in other instances interviewers did believe that respondents used this category because their gender identity was not represented in the gender described question (i.e., bigender or genderqueer).

In summation, while the interviewers reported that some respondents had negative reactions or sensitivity to the SOGI questions, overall it appeared that respondents were able to understand and answer the questions. Nonetheless, because many interviewers experienced at least some pushback from respondents, it is important to examine nonresponse patterns for respondents.

3.2. *Item Nonresponse and Breakoffs*

Item nonresponse to the SOGI questions was low compared to other questions in the NCVS. About 2.77% of respondents refused to answer the sexual orientation question. About 0.41% of respondents answered *don't know* to the question. These two nonresponse categories combined with other missing responses (i.e., respondents that were eligible, or in-universe, but had a missing response due to changes to variables used to define the universe of the question, such as age or sex, in postdata collection processing) to the sexual orientation question account for 3.51% of all respondents age 16 or older (Table 1). Less than one percent (0.97%) of respondents refused to answer the gender identity questions. Only 0.01% of respondents answered *don't know* to the questions. These two nonresponse categories combined with other missing responses to the gender identity questions account for 1.33% of all respondents age 16 or older. Comparatively, about 25% of respondents answered *don't know* or refused to answer a question about household income.

These nonresponse rates varied by certain demographic characteristics. A similar percentage of male and female respondents refused to answer the SOGI questions.

Table 1. Unweighted percent of nonresponse to sexual orientation and gender identity (combined responses to sex at birth and current gender identity questions) items.

	Sexual orientation		Gender identity	
	Refusal nonresponse	Total nonresponse[a]	Refusal nonresponse	Total nonresponse[a]
Demographics	Percent	Percent	Percent	Percent
Total	2.77%	3.51%	0.97%	1.33%
Sex[b]				
Male	2.71%	3.48%	0.95%	1.33%
Female	2.82%	3.54%	0.99%	1.33%
Race/Hispanic origin[b]				
Non-Hispanic white	2.77%	3.45%	1.00%	1.32%
Non-Hispanic black	2.92%	3.73%	1.12%	1.59%
Hispanic	2.53%	3.33%	0.71%	1.09%
Age[b]				
16–17	1.97%	3.05%	0.54%	0.88%
18–24	1.97%	2.84%	0.41%	0.73%
25–34	2.63%	3.33%	0.68%	1.04%
35–49	2.49%	3.20%	0.93%	1.35%
50–64	3.22%	3.87%	1.17%	1.49%
65 or older	2.96%	3.80%	1.21%	1.54%
Region				
Northeast	3.66%	4.82%	1.10%	1.76%
Midwest	2.76%	3.39%	0.91%	1.14%
South	2.44%	3.07%	0.95%	1.34%
West	2.78%	3.62%	1.03%	1.29%
Location of residence				
Urban	3.50%	4.38%	0.97%	1.37%
Suburban	2.65%	3.41%	1.07%	1.42%
Rural	1.78%	2.25%	0.68%	0.96%

[a]Total nonresponse includes all nonresponse, including refusals, don't knows, and in-universe missing.
[b]Based on data collected on the household roster for the sampled household, including the name, age, sex, race, Hispanic origin, marital status, and education level of each person living in the household.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July–December 2016.

Nonresponse to the sexual orientation question was similar across the race and Hispanic origin groups, while a slightly lower percentage of Hispanics (0.71%) refused to answer the gender identity questions compared to non-Hispanic Whites (1.00%) and non-Hispanic Blacks (1.12%). Respondents age 16 to 24 generally had lower refusal rates to the sexual orientation question than persons age 25 or older. About one percent of respondents age 50 to 64 and 65 or older refused to answer the gender identity questions, compared to only 0.54% of respondents age 16 to 17 and 0.41% of respondents age 18 to 24. This difference in nonresponse by age supports the qualitative result that found that older respondents had a harder time understanding or answering the sexual orientation and gender identity questions. Differences in nonresponse were also observed by region for the sexual orientation question. More respondents in the Northeast refused to answer the sexual orientation question than respondents in the Midwest, South, and West. For gender

identity, nonresponse was similar across region. A greater percentage of respondents living in urban areas refused to answer the sexual orientation question than respondents in suburban and rural areas. Less than one percent of respondents living in rural areas (0.68%) refused to answer the gender identity questions, compared to 0.97% of respondents in urban areas and 1.07% of respondents in suburban areas.

Breakoffs from the survey at the SOGI questions were also low. Breakoffs include respondents that stopped participating in the survey at this question. Of all breakoffs, 0.24% happened at the sexual orientation question. About 0.10% of all breakoffs happened at the sex at birth question, 0.04% of all breakoffs happened at the current gender identity question, and no respondents broke off from the survey at the gender confirmation question. For comparison, about 13% of all breakoffs occurred when respondents were asked if their house was rented or owned, which is very early on in the instrument, and about 0.29% of all breakoffs occurred at the hearing-based disability question, which is at the end of the instrument with the SOGI items.

The analysis examined various predictors of nonresponse to the sexual orientation question and found that total household income, age, mode of interview, educational attainment, and race predicted nonresponse of sexual orientation (i.e., a refusal). The effect of every one category increase in income was to decrease the odds of nonresponse to the sexual orientation question by a factor of 0.97 holding sex, race and Hispanic origin, age, educational attainment, interview language, and mode of interview constant ($p < 0.001$, Table 2). The effect of every one-year increase in age was to increase the odds of nonresponse by a factor of 1.01 ($p < 0.001$). The effect of interviews conducted over the phone, compared to interviews conducted in person, was to increase the odds of nonresponse by a factor of 1.34 ($p < 0.001$). The effect of every one year increase in educational attainment was to increase the odds of nonresponse by a factor of 1.01 ($p < .05$). The effect of being a race other than Hispanic, non-Hispanic White, or non-Hispanic Black decreased the odds of nonresponse by a factor of 0.76 compared to being non-Hispanic White ($p < .05$).

The analysis also examined various predictors of nonresponse to the combined gender identity questions (i.e., a refusal or I don't know the answer) and found that nonresponse is closely related to age and mode of interview. The effect of every one-year increase in age was to increase the odds of nonresponse to the gender identity question by a factor of 1.01 ($p < .001$), holding sex, race and Hispanic origin, income, educational attainment, interview language, and mode of interview constant (Table 2). The effect of interviews conducted over the phone, compared to interviews conducted in person, was to increase the odds of nonresponse by a factor of 1.77 ($p < .001$). The effect of mode of interview on nonresponse to gender identity differed by Hispanic origin. Among Hispanics, the effect of interviews conducted over the phone was to decrease the odds of nonresponse by a factor of 0.46 compared to interviews conducted in person ($p < .05$, Table 3). Among non-Hispanic respondents, the effect of interviews conducted over the phone, compared to interviews conducted in person, was to increase the odds of nonresponse by a factor of 2.10 ($p < .001$).

3.3. SOGI Population Estimates

Using NCVS data from July through December 2016, population estimates and percent distribution of SOGI among all persons age 16 or older were estimated. The NCVS data

Table 2. Unweighted logistic regression of nonresponse to sexual orientation and gender identity questions by sociodemographics.

	Sexual orientation		Gender identity	
Respondent characteristic	Logged odds	Odds ratios	Logged odds	Odds ratios
Female	−0.03	0.93	0.04	1.08
	(0.03)		(0.05)	
Race (ref = non-Hispanic white)				
Non-Hispanic black	−0.01	0.99	−0.11	0.80
	(0.05)		(0.08)	
Hispanic	−0.09	0.83	−0.01	0.98
	(0.05)		(0.01)	
Non-Hispanic other	−0.14*	0.76*	−0.13	0.77
	(0.06)		(0.10)	
Age	0.01***	1.01***	0.01***	1.01***
	(0.01)		(0.01)	
Household income	−0.03***	0.97***	−0.02	0.98
	(0.01)		(0.02)	
Years of education	0.08*	1.01*	0.01	1.00
	(0.01)		(0.01)	
Language (ref = English)				
Spanish	0.09	1.19	0.29	1.77
	(0.10)		(0.22)	
Other language	0.17	1.40	−0.04	0.92
	(0.26)		(0.36)	
Mode of interview (ref = face to face interviews)				
Telephone	0.14***	1.34***	0.29***	1.77***
	(0.03)		(0.06)	
Intercept	−4.50***		−6.09***	
	(0.30)		(0.47)	

Note: Standard errors in parentheses.
***p<0.001, **p<0.01, *p<0.05.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July-December 2016.

were weighted to the US population age 16 or older in order to produce estimates. Overall, 1.26% of all persons age 16 or older identified as gay or lesbian, 0.60% identified as bisexual, and 0.11% identified as transgender (Figure 3; Appendix Tables 1 and 2). Among transgender respondents, 51.7% identified as transgender on the current gender identity question and 48.3% reported discordant sex at birth and current gender identity. These data indicate that it is important to collect gender identity using the two-step method to provide an accurate measure of the transgender population. It is possible that if only current gender identity was collected, about half of transgender respondents may not be identified as such if they only selected their currently identified gender (i.e., male or female) and not transgender.

Looking at demographic characteristics, the majority of persons age 16 or older regardless of sex, age, race and Hispanic origin, or where they lived identified as straight. The percent distributions of those persons who identified as gay or lesbian were similar for

Table 3. Unweighted logistic regression of nonresponse to gender identity by sociodemographics by Hispanic and non-Hispanic respondents.

	Among Hispanic respondents		Among non-Hispanic respondents	
Respondent characteristic	Logged odds	Odds ratios	Logged odds	Odds ratios
Female	0.05	1.11	0.03	1.07
	(0.16)		(0.05)	
Age	0.01	1.00	0.02***	1.02***
	(0.01)		(0.01)	
Household income	−0.05	0.95	−0.02	0.98
	(0.04)		(0.02)	
Years of education	−0.02	0.98	0.01	1.01
	(0.01)		(0.01)	
Language (ref = English)				
Spanish	0.45	2.47	4.94	>999.9
	(0.23)		(165.0)	
Other language	5.12	>999.8	−0.15	0.75
	(478.80)		(0.36)	
Mode of Interview (ref = face to face interviews)				
Telephone	−0.39*	0.46*	0.37***	2.10***
	(0.18)		(0.06)	
Intercept	−10.09		−11.12	
	(478.80)		(165.0)	

Note: Standard errors in parentheses.
***p<0.001, **p<0.01, *p<0.05.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July-December 2016.

males and females, 1.41% of males age 16 or older identified as gay compared with 1.13% of females who identified as gay or lesbian (Figure 4). A higher percentage of females identified as bisexual (0.85%) than compared to males (0.32%). Looking at distributions by race and Hispanic origin, 1.38% of non-Hispanic Whites age 16 or older, 1.12% of

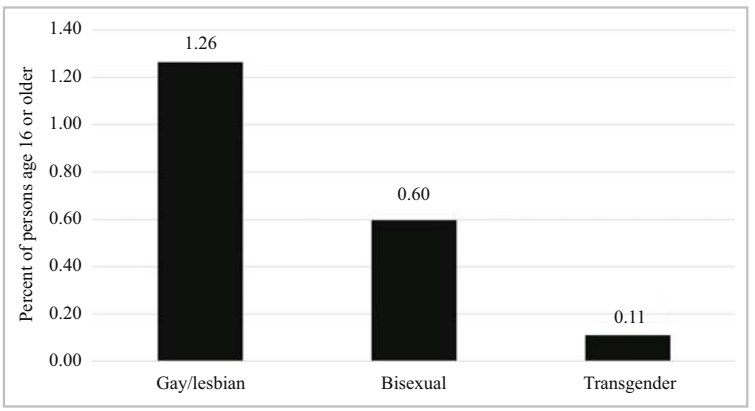


Fig. 3. Percent of persons age 16 or older who identified as LGBT.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July–December 2016.

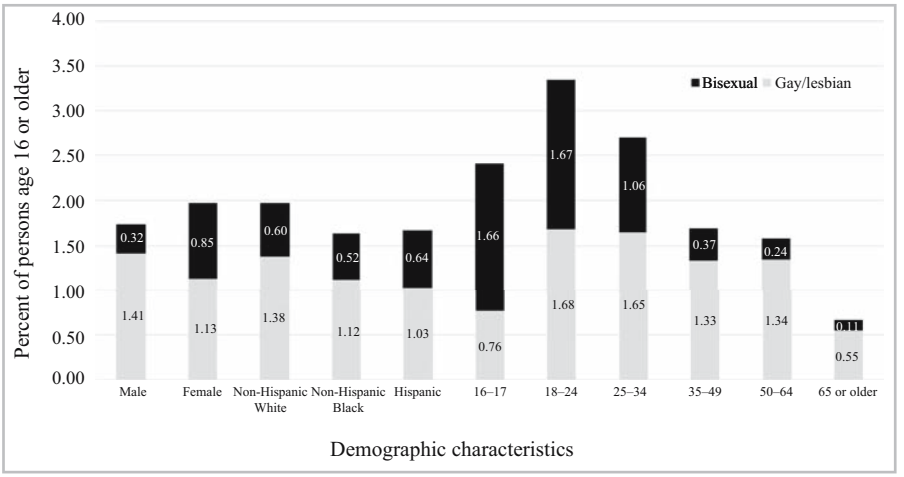


Fig. 4. Percent of persons age 16 or older who identified as LGB, by demographic characteristics.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July–December 2016.

non-Hispanic Blacks, and 1.03% of Hispanics identified as gay or lesbian. About 0.60% of non-Hispanic Whites, 0.52% of non-Hispanic Blacks, and 0.64% of Hispanics identified as bisexual. Percent distributions of those who identified as gay or lesbian varied by age. A higher percentage of persons ages 18 to 24 (1.68%) and ages 25 to 34 (1.65%) identified as gay or lesbian compared to persons ages 16 to 17 (0.76%), 35 to 49 (1.33%), 50 to 64 (1.34%), and 65 or older (0.55%). A larger percentage of younger persons ages 16 to 17 (1.66%) and 18 to 24 (1.67%) identified as bisexual than persons ages 25 to 34 (1.06%), 35 to 49 (0.37%), 50 to 64 (0.24%), and 65 or older (0.11%).

The percent distributions varied among region and location of residence as well. About 1.38% of persons who live in the Northeast and 1.42% of persons who live in the West identified as gay or lesbian, compared to 1.06% of persons who live in the Midwest and 1.23% of persons who live in the South (Figure 5). A higher percentage of persons who

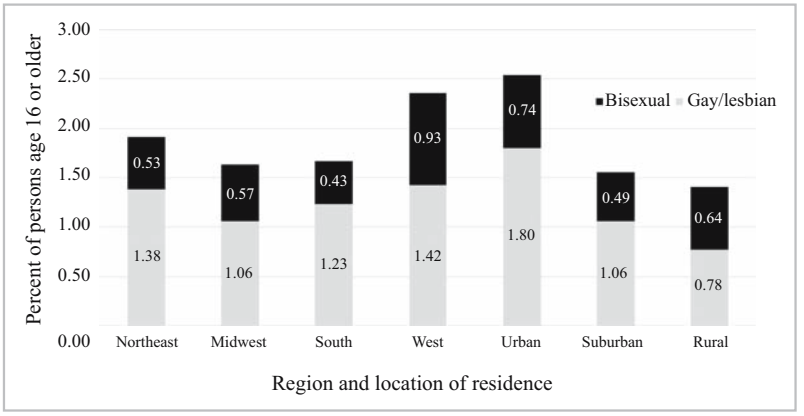


Fig. 5. Percent of persons age 16 or older who identified as LGB, by region and location of residence.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July–December 2016.

live in the West (0.93%) identified as bisexual than persons who live in the Northeast (0.53%), Midwest (0.57%), or South (0.43%). A larger percentage of persons who live in urban areas (1.80%) identified as gay or lesbian compared to persons who live in suburban areas (1.06%) or rural areas (0.78%). Similarly, a higher percentage of persons who live in urban areas (0.74%) identified as bisexual compared to persons who live in suburban areas (0.49%) or rural areas (0.64%).

The majority of persons age 16 or older regardless of sex, age, race and Hispanic origin, or where they lived, identified as either male or female. The distribution of persons age 16 or older who identified as transgender was similar across demographic characteristics. The distribution was similar among race and Hispanic origin, about 0.10% of all non-Hispanic Whites, non-Hispanic Blacks, and Hispanics age 16 or older identified as transgender. Among age groups, 0.22% of persons ages 16 to 17 and 0.21% of persons ages 18 to 24 identified as transgender compared to about 0.10% of persons age 25 or older. Distributions were similar among regions and location of residence. About 0.18% of persons who live in the West identified as transgender, compared to about 0.10% of persons who live in the Northeast (0.08%), Midwest (0.07%), or South (0.11%). About 0.10% each of persons who live in urban (0.13%), suburban (0.11%), and rural (0.06%) areas identified as transgender.

Based on 2015 NHIS data, 97.6% of persons identified as straight, 1.6% identified as gay or lesbian, and 0.8% identified as bisexual (Figure 6) (NCHS 2015b). This compares to 2016 NCVS data where 97.5% of all persons age 18 or older identified as straight, 1.3% identified as gay or lesbian, and 0.6% identified as bisexual. In general, given the differences in methodology the two surveys found reasonably comparable population estimates. It should also be noted that while the National Survey on Drug Use and Health (NSDUH) and National Survey of Family Growth (NSFG) use slightly different question wording, the NCVS estimate of all persons age 18 or older that identified as gay or lesbian was also comparable to those findings (Medley et al. 2016; NCHS 2015a). However, the percentages of adults who identified as bisexual in the NSDUH and NSFG appeared to be higher than the estimate from the NCVS (Medley et al. 2016; NCHS 2015a).

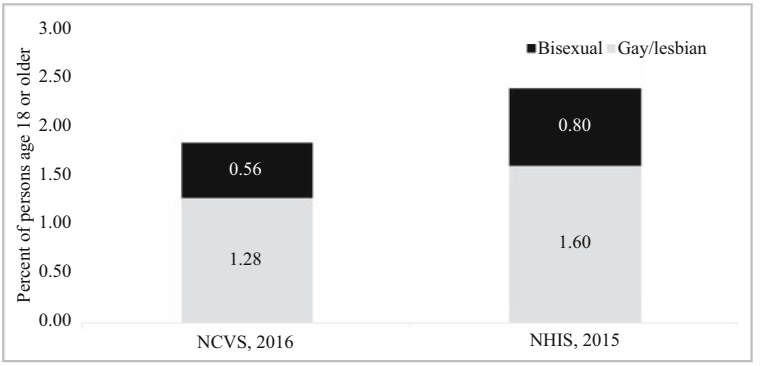


Fig. 6. Percent of adults who identified as LGB.
Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July–December 2016; and National Center for Health Statistics, National Health Interview Survey, 2015.

The NCVS was the first national household-based survey to include a gender identity measure for all respondents ages 16 or older. Currently there are limited national, population-based data collections that the NCVS can be compared to; however, some researchers have estimated the transgender population using the CDC's Behavioral Risk Factor Surveillance System (BRFSS) (Flores et al. 2016; Herman et al. 2017). Although the data collection, methodology, question wording, and sampled population vary by each of these surveys, it is useful to compare estimates to assess data quality. The BRFSS uses a one-step measure to identify transgender persons (*Do you consider yourself to be transgender?*), while the NCVS uses the two-step measure as previously discussed. The research using BRFSS found that 0.6% of U.S. adults identified as transgender. The NCVS estimate for adults who identified as transgender was 0.1%, which is lower than the estimate from the research using the BRFSS data.

4. Discussion

This article addressed three research questions about sexual orientation and gender identity measurement in a large-scale population-based federal survey. A multi-method approach was taken to address these questions including conducting a debriefing questionnaire, focus groups, and targeted interviewers with interviewers, analyzing nonresponse, breakoffs, and estimating the populations. Additionally, the sexual orientation estimates generated from the NCVS data were compared to existing estimates from another household-based survey. Overall, interviewers indicated that both they and the respondents did not have difficulty understanding or comprehending the SOGI items. Nonresponse and breakoffs were low for both sexual orientation and gender identity. Similar LGB population totals using NCVS data were estimated compared to another population-based federal survey.

Interviewers and respondents generally reacted positively to the addition of the SOGI items. Including SOGI items in a crime survey did not present any major problems related to the collection of these data or other demographic data or victimization data. Based on the debriefing interviews, focus groups, and targeted interviews, interviewers indicated that the most common issue respondents had was related to the relevance of these items on a crime survey. However, interviewers were able to address these concerns by using the FAQs that were provided in the CAPI instrument and explain their importance in better understanding the relationship between these characteristics and experiences with criminal victimization. Respondents were understanding and more positive about the questions once this information from the FAQs was provided and explained. During the focus groups, interviewers indicated that English-speaking respondents may have chosen the *something else* response category to sexual orientation because they may not have wanted to disclose their sexual orientation or felt that their sexual orientation was not captured by the categories presented. Whereas, Spanish-speaking respondents may have selected *something else* when there was confusion about the terminology used. This finding is consistent with other research in the field (Stern et al. 2016). Given this, BJS and the Census Bureau continue to monitor the *something else* response category and consider approaches to improving the measurement of sexual orientation among non-English speakers, in particular. In general, interviewers reported fewer issues with the gender

identity questions than the sexual orientation question. Interviewers reported mixed reasons for the use of the *none of these* response category in the second step question for gender identity. In some instances, interviewers felt that respondents chose that response category when they were offended and did not want to answer the question; whereas, in other cases interviewers indicated that respondents used this response category because their gender identity was not represented in the existing response categories. While the interviewers reported that some respondents exhibited sensitivity and other issues toward the questions, overall respondents were able to understand and answer the questions.

Item nonresponse and breakoffs were low for both the sexual orientation and gender identity items. Older and more educated persons were more likely to refuse to answer the sexual orientation question. Older persons were also more likely to refuse to answer the gender identity questions. Respondents who were interviewed over the phone were more likely to refuse both the sexual orientation and gender identity questions compared to those who were interviewed in person.

A direct comparison could be made since the NCVS used the same measure as the NHIS. In general, the NCVS found reasonably comparable population estimates given the methodological differences between the two surveys. These are both household-based surveys using the same question wording, and may be why the results are similar. The NCVS estimates were also compared to other research estimating the transgender population using the BRFSS. The NCVS estimates of the transgender population were less than the population estimated using BRFSS data. These differences are likely a result of differences in the two surveys, including question design and wording, survey context (crime vs. health), and data collection methods.

As with any research, there were some limitations to the collection of SOGI items in the NCVS. The NCVS did not collect data on the *something else* response category for sexual orientation; therefore, it is difficult to fully assess whether persons identified as something else or whether they did not want to respond to the item. In addition, the NCVS did not collect information on the *none of these* response category for the second step question for gender identity. Again, this makes it difficult to determine whether these respondents may be identifying as gender non-conforming, or if they were choosing the response category for another reason. The findings on respondent reactions and any negative reactions came from the interviewers and not directly from the respondents. This is a limitation as the data received was from the perspective of the interviewers whose experience or interpretation of the situation could have been different than the respondents. Future research should seek to address these issues.

SOGI questions can be successfully administered on a large-scale population-based survey. Respondents are able to answer the questions with general ease and have minimally negative reactions to the items. It is recommended that these items be added to other surveys that may be considering including them. However, there are a few considerations to keep in mind. Placement of the questions should be considered; in the NCVS they were placed at the end of the instrument so that they would not have any effect on crime rates. One should also consider why the questions are being added to the survey. It was clear from this research that respondents questioned the relevance of these items on a crime survey, but were willing to answer them once they understood their purpose. Providing FAQs for interviewers to be able to reference when providing additional

clarification to respondents was effective, and therefore is a recommended approach for other surveys. Finally, if gender identity is used, using the two-step method is recommended. Based on data from the NCVS, about half of transgender persons age 16 or older did not identify as transgender in the current gender identity question. Instead, they reported different responses to their sex at birth and current gender identity.

The BJS and the Census Bureau continue to monitor performance, data quality, and field interviewers' concerns as the data are collected. Based on experiences from the NCVS data collection of sexual orientation and gender identity, some key issues emerged as areas to focus future research. These areas are also consistent with those identified in the SOGI Federal Working Group's *Toward a Research Agenda for Measuring Sexual Orientation and Gender Identity in Federal Surveys* report (2016c). These areas include monitoring response categories, specifically *something else* for sexual orientation and *none of these* for gender identity. Current terminology is continually evolving. In addition, more research is needed administering these items in languages other than English. Prior research has indicated that terms related to sexual orientation and gender identity may have different meaning or lack a direct translation in other languages. Finally, as data collection in the NCVS continues, analyses on victimization experiences, including types of victimization and incident characteristics, can be examined. The field of research around sexual orientation and gender identity continues to grow. Despite concerns that may exist related to adding these questions to surveys, the NCVS has shown that they can be successfully administered in a household-based crime survey.

5. Appendix

Table A1. Weighted estimates of persons age 16 or older by sexual orientation.

	Number of persons age 16 or older					Percent of persons age 16 or older				
Demographics	Gay/lesbian	Straight	Bisexual	Something else	Don't know	Gay/lesbian	Straight	Bisexual	Something else	Don't know
Total	2,937,090	226,389,420	1,384,380	545,910	1,021,175	1.26%	97.46%	0.60%	0.24%	0.44%
Sex[a]										
Male	1,580,010	109,099,895	357,990	206,580	534,920	1.41%	97.60%	0.32%	0.18%	0.48%
Female	1,357,080	117,289,525	1,026,390	339,330	486,255	1.13%	97.34%	0.85%	0.28%	0.40%
Race/Hispanic origin[a]										
Non-Hispanic White	2,060,130	145,625,345	902,920	387,775	581,885	1.38%	97.37%	0.60%	0.26%	0.39%
Non-Hispanic Black	315,070	27,541,315	146,235	47,995	99,515	1.12%	97.84%	0.52%	0.17%	0.35%
Hispanic/	380,315	36,040,060	237,225	50,610	197,175	1.03%	97.66%	0.64%	0.14%	0.53%
Age[a]										
16–17	54,385	6,879,450	117,780	34,960	29,455 !	0.76%	96.68%	1.66%	0.49%	0.41%
18–24	446,655	25,406,200	444,175	115,320	151,420	1.68%	95.64%	1.67%	0.43%	0.57%
25–34	666,605	39,119,015	428,790	115,645	150,150	1.65%	96.64%	1.06%	0.29%	0.37%
35–49	756,280	55,478,330	207,295	111,045	220,765	1.33%	97.72%	0.37%	0.20%	0.39%
50–64	773,525	56,561,360	139,435	111,960	214,400	1.34%	97.86%	0.24%	0.19%	0.37%
65 or older	239,635	42,945,065	46,905	56,985	254,990	0.55%	98.63%	0.11%	0.13%	0.59%
Region										
Northeast	547,815	38,541,700	211,480	100,715	264,070	1.38%	97.17%	0.53%	0.25%	0.67%
Midwest	539,575	49,692,250	291,160	129,155	208,905	1.06%	97.70%	0.57%	0.25%	0.41%
South	1,083,410	85,985,440	378,465	125,960	241,835	1.23%	97.92%	0.43%	0.14%	0.28%
West	766,290	52,170,030	503,275	190,085	306,365	1.42%	96.73%	0.93%	0.35%	0.57%
Location of residence										
Urban	1,387,665	74,434,300	573,265	263,430	408,800	1.80%	96.58%	0.74%	0.34%	0.53%
Suburban	1,284,360	118,401,820	593,855	218,425	531,135	1.06%	97.83%	0.49%	0.18%	0.44%
Rural	265,065	33,553,300	217,260	64,055	81,245	0.78%	98.16%	0.64%	0.19%	0.24%

Note: Numbers rounded to nearest 5.

! Interpret with caution. Estimate based on 10 or fewer sample cases, or coefficient of variation is greater than 50%.

[a]Based on data collected on the household roster for the sampled household, including the name, age, sex, race, Hispanic origin, marital status, and education level of each person living in the household.

Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July-December 2016.

Table A2. Weighted estimates of 16 or orlder by gender identity.

	Number of persons age 16 or older					Percent of persons age 16 or older				
Demographics	Male	Female	Transgender[a]	None of these	Don't know	Male	Female	Transgender[a]	None of these	Don't know
Total	113,568,135	122,754,055	265,290	438,135	210,430	47.87%	51.74%	0.11%	0.18%	0.09
Race/Hispanic origin[b]										
Non-Hispanic White	74,150,450	78,093,575	160,180	276,075	90,355	48.54%	51.12%	0.10%	0.18%	0.06
Non-HispanicBlack	12,741,565	15,913,260	35,610	39,415	39,640	44.29%	55.31%	0.12%	0.14%	0.14
Hispanic	17,826,680	19,634,245	40,440	77,155	44,395	47.38%	52.19%	0.11%	0.21%	0.12
Age[b]										
16–17	3,660,205	3,565,610	16,185 !	18,555 !	2,095 !	50.40%	49.10%	0.22% !	0.26% !	0.03
18–24	13,366,840	13,567,930	56,545	69,160	24,130 !	49.35%	50.09%	0.21%	0.26%	0.09
25–34	20,417,715	20,937,435	52,680	83,480	15,255 !	49.19%	50.44%	0.13%	0.20%	0.04
35–49	28,266,940	29,319,320	71,785	79,630	60,850	48.91%	50.73%	0.12%	0.14%	0.11
50–64	28,371,520	30,612,310	37,565	107,570	54,590	47.94%	51.72%	0.06%	0.18%	0.09
65 or older	19,484,920	24,751,455	30,530	79,740	53,505	43.88%	55.75%	0.07%	0.18%	0.12
Region										
Northeast	19,328,775	21,304,905	31,405	93,370	86,080	47.32%	52.16%	0.08%	0.23%	0.21
Midwest	25,098,465	26,851,795	38,715	87,620	36,660	48.16%	51.53%	0.07%	0.17%	0.07
South	42,505,110	46,537,240	96,730	129,615	61,815	47.58%	52.10%	0.11%	0.15%	0.07
West	26,635,780	28,060,115	98,445	127,535	25,875	48.47%	51.07%	0.18%	0.23%	0.05
Location of residence										
Urban	37,866,975	41,179,230	106,905	156,960	61,220	47.71%	51.88%	0.13%	0.20%	0.08
Suburban	59,198,250	63,627,320	138,385	249,660	111,660	48.00%	51.59%	0.11%	0.20%	0.09
Rural	16,502,910	17,947,505	20,000 !	31,515	37,550	47.78%	51.96%	0.06% !	0.09%	0.11

Note: Numbers rounded to nearest 5.

! Interpret with caution. Estimate based on 10 or fewer sample cases, or coefficient of variation is greater than 50%.

[a]Includes respondents with discordant responses to the sex at birth and currently described gender questions. That is, respondents may have identified using the transgender response category or identified with a different gender than their assigned sex at birth.

[b]Based on data collected on the household roster for the sampled household, including the name, age, race, Hispanic origin, marital status, and education level of each person living in the household.

[c]White and black race categories exclude persons of Hispanic or Latino origin.

Source: Bureau of Justice Statistics, based on internal data from the U.S. Census Bureau, National Crime Victimization Survey, July-December 2016.

6. References

Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Los Angeles: Sage. Available at: http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Charmaz_2006.pdf (accessed July 2019).

CHIS (California Health Interview Survey). 2018. *CHIS Questionnaires*. Los Angeles: UCLA Center for Health Policy Research. Available at: http://healthpolicy.ucla.edu/chis/design/Pages/questionnairesEnglish.aspx (accessed November 2018).

Collier, K.L., G. van Beusekom, H.M.W. Bos, and T.G.M. Sandfort. 2013. "Sexual orientation and gender identity/expression related peer victimization in adolescence: A systematic review of associated psychosocial and health outcomes." *Journal of Sex Research* 50: 299–317. DOI: https://doi.org/10.1080/00224499.2012.750639.

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 Health Statistics* 2(169): 1–24. Available at: https://www.cdc.gov/nchs/data/series/sr_02/sr02_169.pdf (accessed July 2019).

Ellis, R., M. Virgile, J. Holzberg, D.V. Nelson, J. Edgar, P. Phipps, and R. Kaplan. 2017. *Assessing the Feasibility of Asking Sexual Orientation and Gender Identity in the Current Population Survey: Results from Cognitive Interviews*. Washington, DC: Center for Survey Measurement, U.S. Census Bureau and Office of Survey Methods Research, Bureau of Labor Statistics. Available at: https://www.bls.gov/osmr/research-papers/2017/html/st170210.htm (accessed July 2019).

Flores, A.R., J.L. Herman, G.J. Gates, and T.N. Brown. 2016. *How Many Adults Identify as Transgender in the United States?* Los Angeles, CA: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/How-Many-Adults-Identify-as-Transgender-in-the-United-States.pdf (accessed July 2019).

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. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/geniuss-report-sep-2014.pdf (accessed July 2019).

Herman, J.L., A.R. Flores, T.N.T. Brown, B.D.M. Wilson, and K.J. Conron. 2017. *Age of Individuals who Identify as Transgender in the United States*. Los Angeles, CA: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/TransAgeReport.pdf (accessed July 2019).

IOM (Institute of Medicine). 2011. *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press. Available at: http://www.nationalacademies.org/hmd/Reports/2011/The-Health-of-Lesbian-Gay-Bisexual-and-Transgender-People.aspx (accessed March 2018).

Jans, M., J. Viana, D. Grant, S.D. Cochran, A.C. Lee, and N.A. Ponce. 2015. "Trends in sexual orientation missing data over a decade of the California Health Interview Survey." *American Journal of Public Health* 105(5): e43–e50. DOI: https://doi.org/10.2105/AJPH.2014.302514.

Kann, L., E.O. Olsen, T. McManus, S. Kinchen, D. Chyen, W.A. Harris, and H. Wechsler. 2011. *Sexual Identity, Sex of Sexual Contacts, and Health-Risk Behaviors Among*

Students in Grades 9–12 — Youth Risk Behavior Surveillance, Selected Sites, United States, 2001–2009. Washington, DC: Centers for Disease Control and Prevention. (MMWR 60(SS07):1–133). Available at: https://www.cdc.gov/mmwr/preview/mmwrhtml/ss6007a1.htm (accessed July 2019).

Krebs, C., C. Lindquist, M. Berzofsky, B. Shook-Sa, K. Peterson, M. Planty, L. Langton, and J. Stroop. 2016. *Campus Climate Survey Validation Study Final Technical Report*. Washington, DC: Bureau of Justice Statistics. (NCJ 249545). Available at: https://www.bjs.gov/content/pub/pdf/ccsvsftr.pdf (accessed July 2019).

Martinez, M., A. Henderson, J. Luck, and M.C. Davis. 2017. *Cognitive Pretesting of the National Crime Victimization Survey Supplemental Victimization Survey*. Suitland, MD: U.S. Census Bureau, Center for Survey Measurement. Available at: https://www.census.gov/library/working-papers/2017/adrm/rsm2017-03.html (accessed November 2018).

Medley, G., R.N. Lipari, J. Bose, D.S. Cribb, L.A. Kroutil, and G. McHenry. 2016. *Sexual Orientation and Estimates of Adult Substance Use and Mental Health: Results from the 2015 National Survey on Drug Use and Health*. Washington, DC: Substance Abuse and Mental Health Services Administration. Available at: https://www.samhsa.gov/data/sites/default/files/NSDUH-SexualOrientation-2015/NSDUH-SexualOrientation-2015/NSDUH-SexualOrientation-2015.htm (accessed July 2019).

National Coalition of Anti-Violence Programs. 2016. *Lesbian, Gay, Bisexual, Transgender, Queer, and HIV-Affected Intimate Partner Violence in 2015*. New York, NY: Emily Waters. Available at: https://avp.org/wp-content/uploads/2017/04/2015_ncavp_lgbtqipvreport.pdf (accessed March 2018).

NCHS (National Center for Health Statistics). 2015a. *Sexual identity for men and women 18–44 years of age*. Washington DC: NCHS. Available at: https://www.cdc.gov/nchs/nsfg/key_statistics/s.htm#sexualorientation (accessed March 2018).

NCHS (National Center for Health Statistics). 2015b. *Sexual Orientation and Health among U.S. Adults: National Health Interview Survey, 2015*. Washington DC: NCHS. Available at: https://www.cdc.gov/nchs/data/nhis/sexual_orientation/asi_2015_stweb-site_tables.pdf (accessed March 2018).

NCHS (National Center for Health Statistics). 2018. *NHIS Data, Questionnaires and Related Documentation*. Washington DC: NCHS. Available at: https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm (accessed November 2018).

Reisner, S.L., K.J. Conron, L.A. Tardiff, S. Jarvi, A.R. Gordon, and S.B. Austin. 2014. "Monitoring the health of transgender and other gender minority populations: validity of natal sex and gender identity survey items in a U.S. national cohort of young adults." *BMC Public Health* 14(1): 1224. Available at: https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-14-1224 (accessed July 2019).

SMART, The Sexual Minority Assessment Research Team. 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys*. Los Angeles, CA: The Williams Institute. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/SMART-FINAL-Nov-2009.pdf (accessed July 2019).

Stern, M.J., S. Michaels, C. Milesi, M.H. Viox, and H. Morrison. 2016. *Task 1.32.a: Develop and Test Sexual and Gender Minority Status (LGBT) Items Final Summary Report*. Washington, DC: Centers for Medicare and Medicaid Services, Office of Minority Health. (HHSM-500-2014-00035I). Available at: https://www.cms.gov/

About-CMS/Agency-Information/OMH/Downloads/Research-Report-sexual-gender-minority-status-items.pdf (accessed July 2019).

SOGI Federal Working Group (Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys). 2016a. *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 March 2018).

SOGI Federal Working Group (Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys). 2016b. *Evaluations of Sexual Orientation and Gender Identity Survey Measures: What Have We Learned?* Available at: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/Evaluations_of_SOGI_Questions_20160923.pdf (accessed March 2018).

SOGI Federal Working Group (Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys). 2016c. *Toward a Research Agenda for Measuring Sexual Orientation and Gender Identity in Federal Surveys: Findings, Recommendations, and Next Steps*. Available at: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/SOGI_Research_Agenda_Final_Report_20161020.pdf (accessed March 2018).

Tate, C.C., J.N. Ledbetter, and C.P. Youssef. 2013. "A two-question method for assessing gender categories in the social and medical sciences." *Journal of Sex Research* 50(8): 767–776. DOI: https://doi.org/10.1080/00224499.2012.690110.

The Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act of 2009 (HCPA), 18 U.S.C. § 249 (2009). Available at: https://www.congress.gov/111/bills/s909/BILLS-111s909is.pdf (accessed July 2019).

VAWA, Violence Against Women Reauthorization Act of 2013, Pub. L. No. 113-4, Stat. 47 (2013). Available at: https://www.congress.gov/113/plaws/publ4/PLAW-113publ4.pdf (accessed July 2019).

Walters, M.L., J. Chen, and M.J. Breiding. 2013. *The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Findings on Victimization by Sexual Orientation*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/violenceprevention/pdf/nisvs_sofindings.pdf (accessed July 2019).

Xavier, J.M. 2000. *The Washington Transgender Needs Assessment Survey: Final report for phase two*. Washington, DC: Administration for HIV/AIDS of the District of Columbia. DOI: https://doi.org/10.13140/RG.2.2.23988.65924.

Xavier, J., J.A. Honnold, and J. Bradford. 2007. *The Health, Health-related Needs, and Lifecourse Experiences of Transgender Virginians*. Richmond, VA: Virginia Department of Health, Division of Disease Prevention. Available at: http://www.vdh.virginia.gov/content/uploads/sites/10/2016/01/THISFINALREPORTVol1.pdf (accessed July 2019).

Received March 2018
Revised April 2019
Accepted August 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 859–884, http://dx.doi.org/10.2478/JOS-2019-0036

Intersections between Sexual Identity, Sexual Attraction, and Sexual Behavior among a Nationally Representative Sample of American Men and Women

Emma Mishel[1]

Social scientists struggle on how to best operationalize and measure sexual orientation. Depending on the survey, researchers can use self-reports of lesbian, gay, or bisexual (LGB) identity, same-sex partner cohabitation, same-sex sexual attraction, or same-sex sexual behavior. All measures come with their own limitations. To illuminate differences in these measures, this study examines the intersections between self-reported sexual identity, attraction, and behavior among a nationally representative sample of US men and women aged 15–45. I explore how and when the three measures align, examine the determinants of self-identifying as gay or bisexual based on sexual behavior and attraction, and assess gender differences in the patterns. I find that about 20% of women and 10% of men aged 15–45 would comprise the LGB community if it were defined to include those who report at least one of the following: gay or bisexual identity, any same-sex attraction, or same-sex sex in the last year. This is much higher than the 6.4% of women and 3.6% of men aged 15–45 who self-identify as LGB. I conclude with recommendations that can aid in measurement of the LGB population, and discuss implications for using certain measures over others when conducting research on the LGB community.

Key words: Sexual orientation; sexuality; gay; population; quantitative.

1. Introduction

Much research on sexual minorities in the United States is qualitative due to challenges of collecting quantitative data on the lesbian, gay, and bisexual (LGB) population. One challenge is that questions about sexual orientation are not always asked in large, nationally representative surveys. Another is that, of the national surveys that do ask about sexual orientation, measures are not always consistent, and some surveys include more than one measure. For example, the concept of sexual orientation involves three main elements: the label we attach to our sexuality (sexual identity), the gender of those whom we engage in sexual activity (sexual behavior), and the gender of those whom we experience attraction to (sexual attraction). Researchers must thus select which measure(s) to use as their indicator of the LGB community – sometimes a crucial decision, as research results may differ depending on the measure used. An additional challenge is that responses to questions of this nature can be especially vulnerable to response bias.

[1] New York University, 295 Lafayette St 4th Floor, New York, New York, 10012, U.S.A. Email: emmamishel@nyu.edu

Acknowledgments: I gratefully thank Paula England, Mike Hout, and David Greenberg for their helpful feedback, as well as the editors of Journal of Official Statistics for their helpful comments.

Since operationalizing and defining sexual orientation is very complex, how to best estimate the *size* of the LGB population is also difficult. However, accurate measurement of the LGB population is vital, as researchers use these estimates to study inequalities between heterosexuals and LGB individuals across dimensions such as urban amenities, health, fertility and morbidity, educational attainment and investment, economic earnings, and household divisions of labor (e.g., Baumle et al. 2009; Berg and Lien 2002; Black et al. 2002; Black et al. 2003; Carpenter 2009; Harper 2007). Psychologists, evolutionary biologists, and geneticists also use these estimates in their respective fields (e.g., Cochran et al. 2003; Gavrilets and Rice 2006), and policy-makers take these estimates into account when assessing civil rights arguments and non-discrimination bills.

Using a self-identification measure to estimate the size of the lesbian, gay, and bisexual community, researchers from the Williams Institute have suggested that LGB individuals make up about 3.5% of the U.S. population (Gates 2011). Yet, self-identifying as lesbian, gay, or bisexual may put one at risk for employment discrimination (Tilcsik 2011; Mishel 2016), housing discrimination (Friedman et al. 2013), as well as at risk for being subjected to negative social stigmas by others, specifically surrounding one's own morality and competency (Mize and Manago 2018a; Webster et al. 1998). Thus, it is likely that not all Americans who feel some same-sex attraction, or who engage in same-sex sexual activity, also *identify* as lesbian, gay or bisexual, and this may translate into not marking it on a survey.

In this vein, this study analyzes the links between the three commonly used measures of sexual orientation: sexual identity, sexual attraction, and sexual behavior, using a nationally representative sample of the United States. Specifically, I pool four waves of a nationally representative dataset of U.S. men and women aged 15–45, and explore the a) patterns and links between self-reported sexual identity, attraction, and behavior, b) determinants of self-identifying as lesbian, gay, or bisexual based on self-reported sexual behavior and sexual attraction, and c) gender differences in these patterns. Based on the findings, I put forth arguments for using certain measures over others when conducting research on the LGB community, and I provide recommendations that can aid in measurement and analysis of the LGB population. In all, this research seeks to shed light on the complexity of defining, measuring, and operationalizing sexual orientation, while providing insights that can help researchers in conducting quantitative data analysis on the LGB population.

2. Past Research

To conduct research on the LGB population, researchers must make an important decision about the specific measure they choose to use as their indicator of LGB identity. Some studies use self-identification as lesbian, gay, or bisexual as their measure (e.g., Carpenter and Eppink 2017; Cochran et al. 2003); others use reports of same-sex sexual behavior, either recently or ever (e.g., Badgett 1995; Berg and Lien 2002; Black et al. 2003); some use reports of same-sex sexual attraction (e.g., Busseri et al. 2006); others use reports of same-sex partner cohabitation (e.g., Baumle and Poston 2011; Fischer 2016); and yet others use a combination of one or more of these measures (e.g., Bostwick et al. 2010; England et al. 2016; Mize 2016). The decision of which measure to use is critical, as past

studies show disjunctures between a person's self-described sexual identity and with whom the individual engages in sexual behavior, or to whom the individual is sexually attracted (Epstein et al. 2012; Igartua et al. 2009; Kinsey et al. 1953; Laumann et al. 1994), and these disjunctures have been shown to change research results, depending on the measure of sexual orientation used (Bostwick et al. 2010; Compton et al. 2015; Korchmaros et al. 2013; Mize 2016).

To understand the correlations between sexual identity, sexual desire, and sexual behavior, previous studies conducted in the United Kingdom and Australia have used nationally representative, country-specific data to analyze the patterns between the three measures among its citizens (e.g., Geary et al. 2018; Richters et al. 2014; Smith et al. 2003; Wellings et al. 1994). For Australian men and women, same-sex attraction and same-sex sexual behavior were more common than gay or bisexual identity would suggest (Richters et al. 2014; Smith et al. 2003), and similar results were found in the United Kingdom (Geary et al. 2018; Wellings et al. 1994).

Past studies based in the United States have examined patterns between sexual identity, attraction, and behavior as well (e.g., Laumann et al. 1994). The more recent studies to do so have either used smaller and unrepresentative samples, such as adolescents/young adults from one Southwestern city (Korchmaros et al. 2013), young women (Diamond 2008), men from New York City (Pathela et al. 2006), or have focused on bisexuality (Compton et al. 2015), asexuality (Poston and Baumle 2010), or on health risks for the LGB community (Bostwick et al. 2010). Two reports from the Center for Disease Control's National Health Statistics Department use nationally representative data of the United States to give a cursory glance at the links between sexual behavior, attraction and identity (Chandra et al. 2011; Copen et al. 2016). However, these reports do not go beyond mere descriptive tables, and with their focus on heterosexual behavior, they are not able to list values in many figures pertaining to non-heterosexual identity, same-sex sexual behavior, or different levels of same-sex attraction due to small sample size.

This research builds upon previous studies by pooling four waves of a large representative sample of U.S. men and women aged 15–45 to examine the links between sexual identity, sexual attraction, and sexual behavior. I go beyond descriptive statistics to estimate relationships using regression analysis, which allows me to control for important demographic characteristics when assessing these patterns. I also examine gender differences in patterns, and discuss implications for using different measures of sexual orientation. I conclude by discussing which measure(s) of sexual orientation may be most useful depending on the research objectives, and by discussing how this analysis has policy implications, as estimates for and research on the LGB community are used to inform a host of different research fields.

3. Data and Measures

I use the National Survey of Family Growth (NSFG) to conduct this research. The NSFG is a nationally representative survey of the U.S. household population, with in-person interviews conducted with men and women between 15 and 45. The survey collects information on family life, marriage and divorce, pregnancy, infertility, use of contraception,

health, and sexuality. The NSFG uses three measures of sexual orientation: sexual identity, sexual attraction, and sexual behavior, which allows me to compare responses from all three variables. To obtain the largest sample size of sexual minorities, I pool data from four NSFG survey waves: 2002, 2006–2010, 2011–2013, and 2013–2015. In wave 2002, 7,643 women and 4,928 men (12,571 in total) were sampled; in wave 2006–2010, 12,279 women and 10,403 men (22,682 in total) were sampled; in wave 2011–2013, 5,601 women and 4,815 men (10,416 in total) were sampled; and in wave 2013–2105, 5,699 women and 4,506 men (10,205 in total) were sampled, making my total sample size 55,874. All analyses were conducted using the designated weights from NSFG. As such, results in all tables and figures reflect the general U.S. population among those aged 15 to 45. Participants were included in analysis if they answered specific items that assessed sexual identity, sexual attraction, and sexual behavior measures, as well as relevant demographic measures, for each table and regression in question; if not, they were excluded from the analysis. Below, I describe how I operationalize each sexual orientation measure, and provide details on NSFG's question ordering and relevant pathways.

3.1. Sexual Orientation Measures

3.1.1. Sexual Identity

To assess sexual identity, I use a question that asked respondents whether they see themselves as "heterosexual or straight," "homosexual, gay, or lesbian," or "bisexual." Note that for brevity, and to facilitate comparing analyses across genders, I use the term "gay" when referring to the men and women who marked "homosexual, gay, or lesbian." I use the term "straight" when referring to the men and women who marked "heterosexual or straight." In the 2002 wave and part of the 2006–2010 wave (through June 2008), respondents were also given the option of choosing "something else" to describe their sexual identity; and, in every wave except 2002, respondents were also given the option of choosing "don't know" to describe their sexual identity. Thus, I create a fourth category, "something else/ don't know," which combines these two additional sexual identity options across survey waves. I include and combine "something else" and "don't know" sexual identities, as I believe it is important to get a sense of the proportions of individuals who do not select one of the standard sexual identities when other options are given.

3.1.2. Sexual Attraction

To assess attraction I use the following question in the NSFG: "People are different in their sexual attraction to other people. Which best describes your feelings? Are you..." Male respondents were then shown categories for: only attracted to females, mostly attracted to females, equally attracted to males and females, mostly attracted to males, and only attracted to males. Response options for women were the same, but the gender order was reversed. Thus, I coded attraction on a five-point scale, as follows: only attracted to the other sex, mostly attracted to the other sex, equal attraction to both sexes, mostly attracted to the same sex, and only attracted to the same sex.

Please not that throughout this article, I use the term "other sex" rather than "opposite sex" when referring to women in relation to men and vice versa, because men and women

are not the opposite of each other. When referring to women reporting sex with women, and men reporting sex with men, I use the term "same-sex" to describe the sexual behavior. I prefer language that refers to an other sex rather than the other sex so as not to imply that there are only two sexes, given that some individuals identify as nonbinary. However, the NSFG provided only "male" or "female" as categories for respondents and their sex partners, so in these data male is the only other sex choice for women and vice versa. Thus, I use language consistent with the limitations of the data.

3.1.3. Sexual Behavior

I include measures for lifetime and recent sexual behavior using the following questions: "Thinking about your entire life, how many [male/female] sex partners have you had?" and "Thinking about the last 12 months, how many [male/female] sex partners have you had?" I operationalize these variables to assess whether someone had *ever* had a same-sex sex partner (0 for no, 1 for yes), and whether they had one *in the last year* (0 for no, 1 for yes). I also use these measures to assess *how many* same-sex sex partners they have ever had, and had in the last year, coded as follows: 0, 1, 2–4, or 5 or more. Lastly, I include a measure of whether the respondent has had sex partners of *both* sexes, both ever and in the last year (0 for no, 1 for yes).

3.2. NSFG's Survey Design and Question Pathways

In this section, I describe the question pathways for the relevant sexual orientation measures. Women were first asked if they have ever performed oral sex on, or received oral sex from, another female. If they said no, they were then asked if they had "ever had any type of sexual experience of any kind with another female." If they said yes to any of these questions, they were then asked about their number of female sex partners ever, and in the last year. (If they said no to any of the questions above, they were assumed to have no same-sex sex partners.) Then, women were asked about their sexual attraction towards men and women, and about their sexual identity. Men were first asked if they have ever had oral or anal sex with another male, either given or received. If they answered yes to any of these questions, they were then asked about their number of male sex partners ever, and in the last year. (If they said no to any of the questions above, they were assumed to have no same-sex sex partners.) Then, men were asked about their sexual attraction towards women and men, and about their sexual identity.

Note that, as indicated above, the pathways used to screen respondents for questions about same-sex sex partners were slightly different for men and women. At first glance, NSFG seems to have created a higher bar for men to report a same-sex sex partner compared to women. That is, men weren't asked how many male sex partners they had had (and thus were assumed to have had none) unless they said they had oral or anal sex with a man, whereas women could be classified as having had sex with a woman if they said they had had "any" sexual experience with a woman, even if they did not report having had oral sex with a woman. Recent attention to the prevalence of women kissing women on dance floors and at parties (Rupp et al. 2014; Hamilton 2007) raises the question of whether women reporting sexual experience with women are referring to experiences such as these, or to more private sexual contact involving genitals. In analyses not shown, I ascertained

that 88% of the women who said they had sex with a woman in the last year (regardless of whether they also said they had sex with a man) also reported that they had ever had oral sex with a woman, as did 86% of (the overlapping group of) women who reported having sex with both men and women in the last year. This suggests that the vast majority of those who say they have had a female sex partner have had private sexual experiences with women beyond kissing. Thus, I am relatively unconcerned that the measures artifactually create a higher bar for men than for women reporting on sexual behavior.

For questions deemed particularly sensitive by NSFG, Audio Computer-Assisted Self-Interview (ACASI) was used, which enables respondents to answer questions privately – that is, without the interviewer seeing their response. Questions about sexual identity, behavior, and attraction were asked in the ACASI portion of the survey, which came at the very end of the NSFG survey. Other questions included in the ACASI part of the survey were about issues such as household income, unemployment, getting expelled from school, drug use, HIV/STDs, and experiences of sexual assault. Use of ACASI may reduce reporting bias on sensitive questions compared to other surveys that do not use this approach (Betts 2009; O'Reilly et al. 1994; Villarroel et al. 2006; Tourangeau and Smith 1996). Overall, NSFG reports that response rates for their recent data releases are around 69% (Centers for Disease Control 2018).

4. Analytical Approach

4.1. Descriptive Statistics

I perform several analyses to examine the links between sexual identity, sexual attraction, and sexual behavior. First, I provide descriptive statistics of men's and women's self-reported sexual identity by self-reported sexual attraction and various sexual behaviors. Then, I assess whether there are statistically significant gender differences in these estimates by performing standard two-tailed t-tests. Following this, I examine the self-reported sexual attraction levels and sexual identities of men and women who report having had same-sex sex *in the past year*. These analyses all help to illuminate the patterns and links between sexual identity, attraction, and behavior among men and women aged 15–45 in the United States, and to assess whether there are gender differences in these patterns.

4.2. Regression Analysis

After analysis of the descriptive statistics, I move to regression analysis. I estimate eight logistic regression models that predict bisexual identity (four for women and four for men) and eight logistic regression models that predict gay identity (four for women and four for men). The dependent variable is either bisexual identity (compared to not bisexual identity) or gay identity (compared to not gay identity). The four main predictors are: (1) self-reported sexual attraction on a five-point scale, (2) whether the individual has ever had same-sex sex, (3) whether the individual has had partners of both sexes in the past year, and (4) whether the individual has had only same-sex sex partners in the past year. Each regression also controls for race, age, age-squared, mother's education, immigrant status, and metropolitan area, as well as survey wave. These regression models seek to assess how certain sexual behaviors and sexual attraction levels predict whether a man or woman age

15–45 takes on a gay or bisexual identity. Such analysis helps to illuminate the links between sexual identity, attraction, and behavior; and, it goes beyond descriptive statistics in a useful way, as it predicts when someone may take on a gay or bisexual identity – arguably the most common measure for counting the LGB population – after controlling for important demographic characteristics.

4.3. Predicted Probabilities

Calculated from the regression results, I plot the predicted probabilities for identifying as gay or bisexual, separately for men and women. I first plot the predicted probabilities corresponding to the four logistic regression models predicting bisexual identity, and then plot the predicted probabilities corresponding to the four logistic regressions predicting gay identity, separately for men and women. Predicted probabilities were obtained using an average marginal effects (AME) approach, and they are adjusted for race, mother's education, immigrant status, metropolitan city, and survey wave. Each plotted predicted probability includes bars to indicate confidence intervals at the 95% level.

5. Results

Table 1 displays the proportion of men and women who identify as straight, bisexual, gay, and something else/don't know. Table 1 also displays the proportions of men and women who have engaged in same-sex sex (such as whether they have ever had a same-sex sex partner, whether they have had partners of both sexes within the last year or ever, whether they have had only same-sex partners in the past year, and the number of same-sex partners they have ever had and had in the past year), as well as their self-reported sexual attraction level, by sexual identity. When examining women's self-reported sexual identities, Table 1 shows that overall, 92.2% of the 30,861 women identify as straight, 4.9% as bisexual, 1.5% as gay, and 1.4% said either something else or don't know. When examining the self-reported sexual identities of men, Table 1 shows that 95.1% of the 24,357 men identify as straight, 1.7% as bisexual, 1.9% as gay, and 1.3% said something else or don't know.

The estimates for those who identify as gay or bisexual are slightly larger here compared to the proportion of people who identify as gay or bisexual in the United Kingdom and Australia (see Geary et al. 2018; Richters et al. 2014); however, this is to be expected with a younger sample, as NSFG only sampled Americans aged 15–45 while the United Kingdom sampled adults aged 16–74 and Australia sampled adults aged 16–69. Indeed, many studies show that younger individuals are more likely to identify as non-heterosexual compared to older individuals (see Bridges and Moore 2018; England et al. 2016; Mishel et al. forthcoming).

Table 1 also shows that there are many gender differences when assessing how men and women age 15–45 label their sexual orientation: significantly more women than men identify as bisexual, and significantly more men than women identify as gay. This pattern is also seen in data from Australia (Richters et al. 2014) and the United Kingdom (Geary et al. 2018), and is consistent with past research in the United States (England et al. 2016; Gates 2011; Korchmaros et al. 2013).

Table 1. Sexual identity by sexual attraction and sexual behavior, U.S. men and women age 15–45.

	Straight			Bisexual			Gay			Something else/ Don't know			Total		
	Women		Men	Women		Men	Women		Men	Women		Men	Women		Men
Total	**92.2%**	***	**95.1%**	**4.9%**	***	**1.7%**	**1.5%**	*	**1.9%**	**1.4%**		**1.3%**	**100.0%**		**100.0%**
Sexual attraction															
Only OS	88.4%	***	96.3%	3.9%	***	14.3%	9.1%		8.4%	74.5%		81.0%	83.1%	***	93.1%
Mostly OS	10.8%	***	3.4%	35.9%		32.8%	1.6%		0.9%	16.2%		11.2%	11.9%	***	3.9%
Equal	0.7%	***	0.2%	52.6%	**	41.0%	4.1%	*	1.3%	4.5%		3.1%	3.2%	***	0.9%
Mostly SS	0.0%	*	0.0%	7.4%		11.7%	28.3%		22.4%	3.5%		1.6%	0.8%		0.7%
Only SS	0.1%		0.2%	0.2%		0.2%	56.9%	*	67.1%	1.4%		3.2%	0.9%	***	1.4%
Sexual behavior															
Total percent who have:															
Ever had SS Sex	10.2%	***	2.6%	76.4%	***	63.3%	83.9%		84.0%	19.7%	***	8.9%	14.7%	***	5.2%
Ever had OS & SS Sex	10.0%	***	2.5%	72.0%	***	55.4%	53.2%	*	41.6%	16.5%	***	7.7%	13.7%	***	4.1%
Had only SS Sex LY	0.1%	*	0.0%	11.9%	*	17.2%	71.1%		74.5%	3.9%		3.0%	1.8%		1.8%
Had SS & OS Sex LY	1.7%	***	0.4%	25.0%		22.3%	9.0%	***	2.1%	5.4%	*	2.2%	3.0%	***	0.8%
# of SS Partners Ever															
0	89.8%	***	97.5%	23.7%	***	36.7%	16.1%		16.0%	80.3%	***	91.1%	85.4%	***	94.9%
1	7.0%	***	1.7%	28.7%	*	20.9%	12.3%	**	6.0%	10.4%	**	4.1%	8.2%	***	2.2%
2 to 4	2.9%	***	0.6%	35.8%	***	22.4%	29.0%	***	15.4%	7.9%	**	2.3%	5.0%	***	1.3%
5 or more	0.3%		0.2%	11.8%	**	20.0%	42.6%	***	62.6%	1.4%		2.5%	1.5%		1.7%
# of SS partners LY															
0	98.3%	***	99.6%	63.1%		60.5%	19.9%		23.4%	90.7%		94.8%	95.3%	***	97.5%
1	1.5%	***	0.3%	26.8%		21.7%	57.1%	***	30.2%	7.4%	**	2.0%	3.6%	***	1.2%
2 to 4	0.2%	**	0.1%	8.6%	*	13.7%	20.3%	*	28.7%	1.8%		1.4%	0.9%		0.9%
5 or more	0.1%		0.1%	1.6%		4.2%	2.7%	***	17.8%	0.1%	*	1.8%	0.2%	***	0.5%
N	**28,217**		**23,017**	**1,648**		**471**	**508**		**573**	**488**		**296**	**30,861**		**24,357**

Notes: Stars indicate significant differences between men and women at alpha levels ***p < .001, **p < .01, and *p < .05. OS is Other Sex; SS is Same Sex; LY is Last Year. In 2002 only, "Don't know" was not an option; in 2009-2015 only, "Something else" was not an option.

5.1. Sexual Identity by Sexual Attraction

How does *sexual identity* relate to *sexual attraction* among U.S. men and women aged 15–45? Table 1 shows some results that may be different from what one might expect. For example, Table 1 shows that about 9.1% of gay women, 3.9% of bisexual women, 8.4% of gay men, and 14.3% of bisexual men report that they are only attracted to the *other-sex*. In addition, over 11% of self-identified straight women report some level of *same-sex* attraction, while a smaller percentage of straight men (about 3.7%) report the same. Furthermore, only 56.9% of self-identified gay women report *only same-sex* attraction, while 67.1% of gay men report the same.

Some of these disjunctures may be surprising. For example, if an individual identifies as gay or bisexual, surely they have some level of same-sex attraction. Indeed, this disjuncture is also found among Australian men and women (about 6% of self-identifying Australian gay women reported only sexual attraction to *men*, and about 2% of self-identifying Australian gay men reported only sexual attraction to *women*) (Richters et al. 2014). Why someone who identifies as gay or bisexual does not report any same-sex attraction can only be speculated. One possible reason for this disjuncture is simple measurement error. Indeed, DeMaio et al. (2013) show that measurement error can be a big issue when conducting research on the LGB population; they find that about 28% of reported same-sex couple households in the 2010 Census are likely to be other-sex couple households, where either the man or woman was wrongly coded as the other gender (and as a result, the 2010 Census has now issued a "preferred" set of state-level estimates of same-sex couples, which are much lower). Other researchers also conclude that the number of same-sex married couples reported in the 2010 Census is likely inflated because gender was mismarked (e.g., Black et al. 2007; O'Connell and Gooding 2006). Coding errors, while not typically a big problem in probability-based analyses, can be a serious issue when analyzing a numerically small group of people, such as those who identify as LGB. While miscoding of gender is not very likely in analysis of NSFG data, since men and women were given different ACASI surveys that asked gender-specific questions about pregnancy and genitalia (see Section 3.2), it is nonetheless possible that responses to other questions were miscoded.

Besides attributing this disjuncture to coding errors, it may also be the case that the respondent did not understand either the sexual identity question or the attraction question, or both. After taking a closer look at the individuals who report gay identity and only *other-sex* attraction in the NSFG, these respondents are disproportionately immigrants (37% immigrant, compared to 15.6% of the entire sample); have less education compared to the entire sample (35% did not graduate high school compared to 21% in the entire sample), and are less likely to be white (37% white, compared to 60% white in the entire sample) (analysis not shown). This may suggest that some of these respondents did not understand one or both of the questions due to a language or cultural barrier. However, when analyzing the *sexual behavior* of these same respondents, 31% indicate that they have had same-sex sex, compared to just 10% of the entire sample who reported same-sex sex. This suggests that there may be more going on than respondents simply misunderstanding the question or marking the wrong answer. In other words, perhaps some individuals legitimately do not feel attracted to members of the same sex, but still identify with a gay or bisexual label for

whatever reason – one possible reason being that they engage in same-sex sex. In any case, whether due to measurement error or due to unique and complex cases of how people align their sexual identity with their sexual attraction and sexual behavior, this finding is indeed puzzling, and it motivates further research perhaps using qualitative methods.

Other disjunctures in Table 1 may be less surprising. For example, Table 1 shows that 11.6% of self-identified straight women age 15–45 report some level of same-sex attraction, as do about 3.7% of self-identified straight men age 15–45. The fact that there are some Americans who feel same-sex attraction but do not identify as gay or bisexual is not surprising, as perhaps they do not feel strongly enough about it to identify as anything other than straight, or are dissuaded from doing so due to the potential for experiencing stigma or discrimination (Friedman et al. 2013; Herek 2009; Mishel 2016; Tilcsik 2011).

Table 1 also shows that about 43% of self-identified gay women report some level of *other-sex* attraction, as do about 33% of self-identified gay men. If we think of sexuality as being on a spectrum, one might assume that gay men and women are only attracted to the same sex, while straight men and women are only attracted to the other sex. This seems to be the case for most straight men and women, as an overwhelming majority of straight women and men age 15–45 report only *other*-sex attraction (88% of women and 96% of men). However, sexual attraction levels for gay women and men aged 15–45 seem to be much more flexible: only 57% of gay women and 67% of gay men report only *same*-sex attraction. Results in Table 1 also reveal that self-identified straight women are significantly more likely to report same-sex attraction compared to self-identified straight men.

5.2. *Sexual Identity by Sexual Behavior*

How does *sexual identity* relate to *sexual behavior* for US men and women aged 15–45? Again, we see some disjunctures, and some significant gender differences. First, Table 1 shows that over 10% of self-identified straight women and 2.6% of self-identified straight men report having had same-sex sex sometime in their life. Furthermore, about 9% of self-identified gay women have had both male and female sex partners in the past year, while about 2.1% of gay men report the same. Of course, it is important to note that these reported sexual experiences could simply be a result of previous sexual exploration, since the survey is measuring *current* sexual identity and *current* sexual attraction, but asking about *past* sexual behaviors. In addition, just because someone engages in sexual behavior with a certain gender does not mean that they identify with a sexual identity that is consistent with that behavior, nor does it automatically mean that they feel attracted to that particular gender; sexual preferences and/or identities may change over time (Diamond 2008).

The gender differences on these estimates are statistically significant, such that significantly more self-identified straight women than straight men report having had same-sex sex sometime in their life, and significantly more self-identified gay women than gay men report having both male and female sex partners in the last year. Among those who identify as bisexual, 76.4% of women report having had same-sex sex sometime in their life, compared to 63.3% of men – again a statistically significant difference. Among *all* men and women age 15–45 in the United States, about 15% of women report having had same-sex sex sometime in their life, as do about 5% of men – also a significant difference by gender.

Interestingly, Table 1 shows that a large proportion of self-identified bisexual and gay individuals aged 15–45 have had 0 same-sex sex partners: about 16% of self-identified gay men and 16% of self-identified gay women report 0 same-sex sex partners, as do about 23.7% of self-identified bisexual women and about 36.7% of self-identified bisexual men. (In comparison, only about 12% of straight men and about 12% of straight women report 0 lifetime other-sex sex partners.) When assessing the *number* of same-sex sex partners for those who identify as gay or bisexual, Table 1 shows that gay and bisexual men report significantly more lifetime same-sex sex partners compared to gay and bisexual women. For example, 62.6% of gay men report five or more lifetime same-sex sex partners compared to just 42.6% of gay women, as do 20% of bisexual men compared to just 11.8% of bisexual women – both significant differences. Along the same lines, significantly more self-identified gay and bisexual women report just one lifetime same-sex sex partner compared to gay and bisexual men.

5.3. Other Sexual Minority Identities

Table 1 also shows that, relative to the number of persons who identified as gay or bisexual, a sizeable portion of both men and women age 15–45 identified as something else or don't know. About 1.3% of men marked something else or don't know, compared to 1.7% men who marked bisexual and 1.9% who marked gay. About the *same* proportion of women marked something else or don't know (1.4%) as those who marked gay (1.5%). Selecting a non-normative identity could be a sign of resistance against traditional sexual identification categories, as more and more individuals who report same-sex attraction or same-sex sexual behavior are taking on other identities, such as "queer." Rejection of normative sexual identity categories (i.e., gay, lesbian, bisexual) and adoption of other identity categories (i.e., queer, pansexual) has been shown to occur more among younger individuals (see Stein 2010; Risman 2018), which is of note since the upper bound of the NSFG sample is age 45. In this case, we might thus expect that individuals who select a non-normative sexual orientation category are, on average, younger, and will have similar self-reported sexual attractions and behaviors to individuals who identify as gay or bisexual. However, Table 1 shows that these individuals seem to be more similar to straight respondents in terms of sexual attraction and behavior. We do see some higher percentages of same-sex sexual behavior among women who reported something else/don't know identities compared to straight women: 19.7% of women who selected something else or don't know reported having had same-sex sex, compared to just 10.2% of straight-identified women, for example. These percentages were slightly higher for men as well, in that 8.9% of men who selected something else or don't know identities reported having had same-sex sex, compared to just 2.6% of straight-identified men.

In results not shown, I analyzed the age of respondents who selected something else or don't know as a sexual identity, and found 23% of those who selected something else or don't know were between the ages of 15 to 19, a higher proportion than the proportion of 15 to 19 year olds in the entire sample (16%). Thus, in addition to the possibility that these individuals would prefer to identify as queer or another sexual identity that is not listed as an option (or do not wish to identify with a label at all), it could also be that they are just

not sure of their sexual orientation and are engaging in sexual exploration to find out, or it could be that they did not understand the question – or a combination of all three.

When assessing whether there are gender differences in sexual behavior among those who identified as something else or don't know, patterns generally follow those found for straight men and women – that significantly more women than men who marked something else/don't know reported having same-sex sex sometime in their life (19.7% versus 8.9%, respectively). There were no significant differences by gender when analyzing sexual attraction levels among those who marked something else/don't know.

5.4. Sexual Attraction and Identity among Those Reporting Same-Sex Sex in Past Year

As several studies use recent sexual behavior as their measure of the LGB community (e.g., Badgett 1995; Berg and Lien 2002; Black et al. 2003), I further examine the sexual attraction and behavior of the sub-sample of individuals who reported having same-sex sex in the last year. Figures 1 and 2 display the self-reported sexual identities (Figure 1) and self-reported sexual attraction levels (Figure 2) among the men and women who reported having had same-sex sex in the past year.

Figure 1 shows that, among women aged 15–45 who have had same-sex sex in the past year, their sexual identities are more or less evenly distributed between straight (34%), bisexual (39%), and gay (25%). In contrast, over half of men aged 15–45 who report having had same-sex sex in the past year identify as gay (56%), while 27% identify as bisexual and 15% identify as straight. Similarly, when looking at sexual attraction among men aged 15–45 who report having had same-sex sex in the last year (Figure 2), about 44% report only being attracted to the same sex, compared to just 17% of women who report the same. These gender differences are quite compelling – among those aged

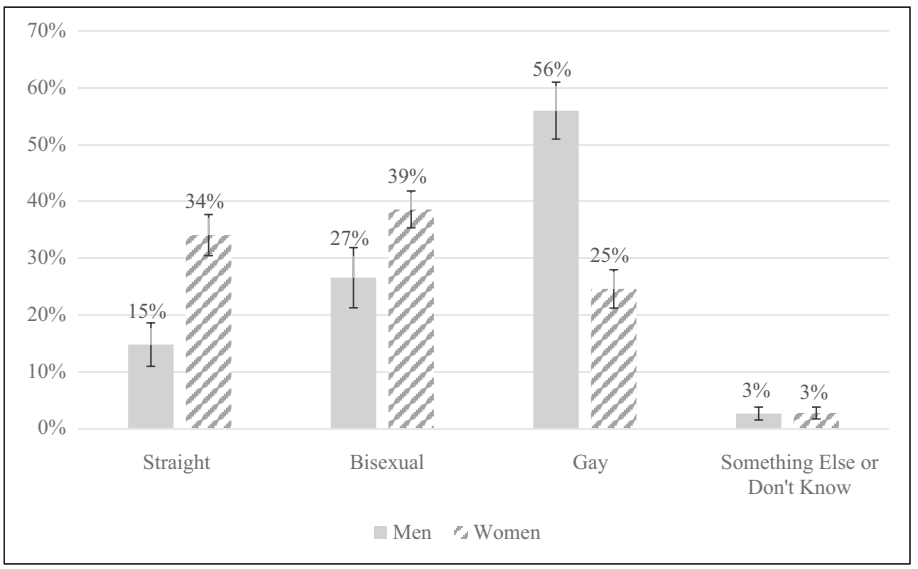


Fig. 1. Self-reported sexual identity among U.S. men and women aged 15–45 who reported having same-sex sex in the last year (n=2,392).
Note: Bars indicate 95% confidence intervals.

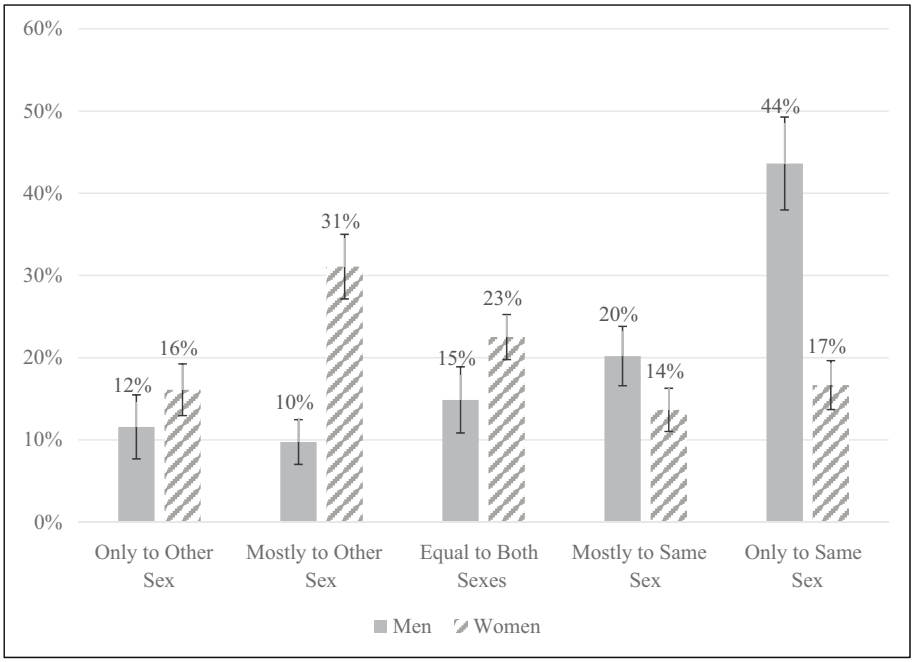


Fig. 2. Self-reported sexual attraction among U.S. men and women aged 15–45 who reported having same-sex sex in the last year (n=2,337).
Note: Bars indicate 95% confidence intervals.

15–45 who report having had same-sex sex in the past year, women's sexual identity and attraction levels are much more evenly distributed among the straight, bisexual, and gay identity categories, as well as all attraction categories. In contrast, the majority of men who have had sex with men in the past year identify as gay and report only same-sex attraction. Generally speaking, results displayed in Figures 1 and 2 are striking: among U.S. men and women age 15–45 who have recently had same-sex sex, their self-reported sexual identity and attraction levels vary greatly.

5.5. Combining Measures: Identity, Attraction, Behavior

My analysis on how sexual identity, attraction, and behavior relate to each other allows me to assess what estimates for the LGB population in the United States may look like (for those age 15–45); if not solely based on sexual identity questions, but rather, if they were based on individuals who had checked at least one of the following: gay or bisexual identity, same-sex sex in the last year, or any same-sex attraction. Using this combination of measurement, 6,277 women in the sample (about 20%) reported at least one of the following: gay or bisexual identity, same-sex sex in the last year, or some level of same-sex attraction, and 2,470 men in the sample (about 10%) also reported at least one of the three measures. Thus, of those aged 15–45, about 20% of women and about 10% of men would comprise the LGB community if it were defined to include all individuals who report a non-heterosexual identity, any same-sex attraction, and/or sex with a same-sex partner in the last year. This is much higher than the observed 6.4% of women and 3.6% of men aged 15–45 who self-identify as gay or bisexual.

5.6. *Predicting Gay or Bisexual Identity from Regressions with Controls*

Next, I plot the predicted probabilities for identifying as bisexual (Figure 3) or gay (Figure 4) from self-reported sexual attraction and various sexual behaviors. Predicted probabilities are calculated from logistic regression models that predict either gay or bisexual identity from sexual attraction levels and various sexual behaviors. All models include sociodemographic controls (see Supplemental material Table A-1 for a description of all control variables used, and please see Tables A-2 and A-3 for logistic regression results).

Figures 3 and 4 show that predicted probabilities for identifying as gay or bisexual based on *sexual attraction* are quite similar for men and women. Specifically, men and women age 15–45 who report equal attraction to both sexes have about a 75% chance of identifying as bisexual (Figure 3), while men and women age 15–45 who report only same-sex attraction have just under a 90% chance of identifying as gay (Figure 4). The distributions of the predicted values are also about what we would expect to see in terms of how self-reported sexual attraction predicts gay and bisexual identity: the highest probability for bisexual identity is attraction to both sexes, while the highest probability for gay identity is only same-sex attraction, for both men and women.

While results are similar for men and women when examining how sexual attraction predicts sexual identity, there are clear gender differences when examining how *sexual behavior* predicts sexual identity for men and women age 15–45. For example, Figure 3 shows that women age 15–45 who report only same-sex sex in the last year have about a

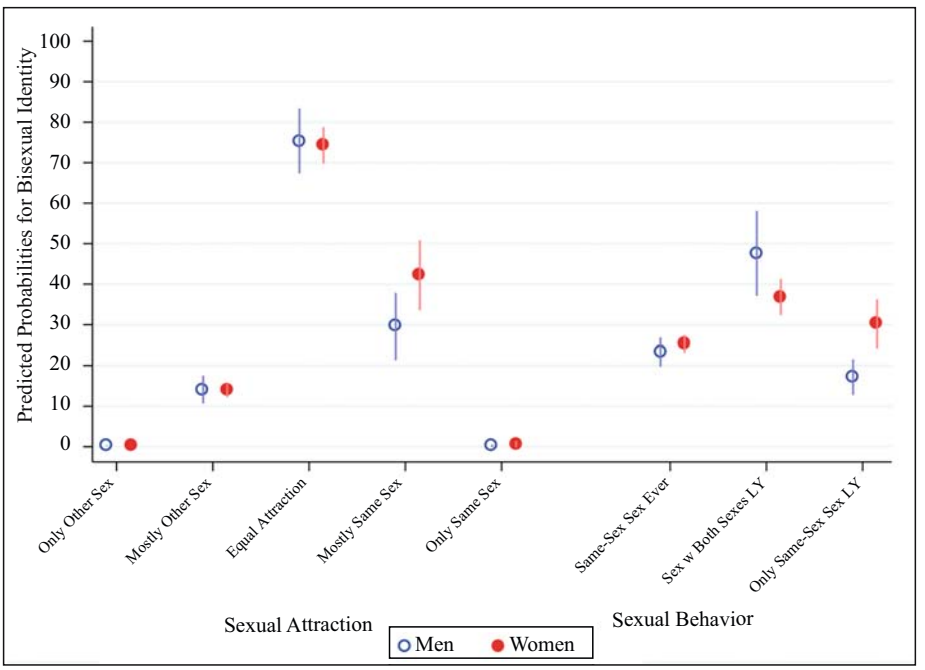


Fig. 3. Predicted probabilities for bisexual identity among U.S. men and women age 15–45.
Note: Predicted probabilities correspond to logistic regressions and have been adjusted for race, age, mother's education, immigrant status, MSA, and survey wave. Bars indicate 95% confidence intervals. LY is Last Year.

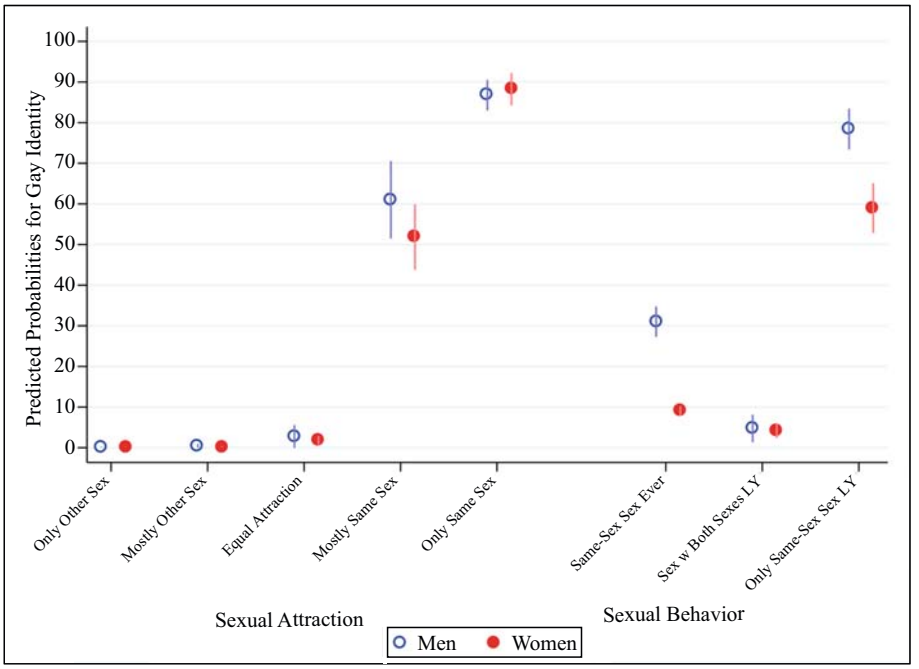


Fig. 4. Predicted probabilities for gay identity among U.S. men and women age 15–45.
Note: Predicted probabilities correspond to logistic regressions and have been adjusted for race, age, mother's education, immigrant status, MSA, and survey wave. Bars indicate 95% confidence intervals. LY is Last Year.

30% chance of identifying as bisexual, while men age 15–45 who report only same-sex sex in the last year have only about a 17% chance of identifying as bisexual. We also see a large difference between men and women in Figure 4 predicting gay identity, in whether they have ever had same-sex sex, and whether they have had only same-sex sex in the last year; probabilities for men are much higher than for women in both cases. Specifically, women age 15–45 who report only same-sex sex in the last year have about a 60% chance of identifying as gay, while men age 15–45 who report only same-sex sex in the last year have about an 80% chance of identifying as gay. Similarly, women age 15–45 who report ever having had same-sex sex have about a 10% chance of identifying as gay, but men age 15–45 who report ever having had same-sex sex have about a 30% chance of identifying as gay. In other words, while having had same-sex sex sometime in one's life and having had only same-sex partners in the last year both have positive effects on identifying as gay for men and women, the effect is much larger for men than for women in both cases.

6. Discussion

6.1. Intersections between Sexual Identity, Sexual Attraction, and Sexual Behavior

This article pools four waves of a nationally representative dataset to examine the relationship between self-reported sexual behavior, sexual identity, and sexual attraction among U.S. Americans aged 15 to 45. Results reveal some disjunctures between the three measures. First, I find that more Americans age 15–45 report same-sex sexual experience

and/or same-sex attraction than identify as gay or bisexual. About 15% of women and 5% of men age 15–45 report having same-sex sex sometime in their life, and about 17% of women and 7% of men age 15–45 report some level of current same-sex attraction. This is compared to just 6.4% of women and 3.6% of men age 15–45 who self-identify as either gay or bisexual. The finding that more men and women report same-sex sexual experience or same-sex attraction than report gay or bisexual identity is consistent with past research (e.g., Compton et al. 2015; Gates 2011; Korchmaros et al. 2013), and there may be many reasons for this. One possible reason is fear of discrimination, as much research provides causal evidence of discrimination against out LGB individuals when they seek employment (Mishel 2016; Tilcsik 2011) and apply for housing (Friedman et al. 2013). Another reason could be negative stereotypes and social stigma often associated with identifying as gay or bisexual. Research by Ward (2008) and Silva (2016) show that some men who sleep with men still choose to identify as straight, and purposely distance themselves from any sort of queer identity to construct (and reinforce) ideals of normative masculinity. Studies have also shown that out LGB individuals are sometimes viewed as less competent, less moral, less trustworthy, and less warm than straight individuals (Mize and Manago 2018a; Webster et al. 1998). In addition, a whopping 23% of the American public still believe that gay and lesbian relations between consenting adults should be illegal (Gallup 2018). Results from these past studies suggest that fear of stigma and social penalties may be one reason that many individuals who have sex with or are attracted to those of the same sex do not take on a non-heterosexual identity.

A further reason that more Americans report same-sex sexual behavior and same-sex attraction than report gay or bisexual identity could be rejection of such labels all together. Research shows that some individuals who report engaging in same-sex sex or who report same-sex attraction also express discomfort in labeling their sexual orientation, saying that current definitions of sexual identity categories do not accurately describe them, that they find them to be restrictive, or that they prefer other identities over non-heterosexual identity labels (Budnick 2016; Diamond 2008; Savin-Williams and Vrangalova 2013). Previous research has also documented the strong hetero-normative expectations and proscriptions – such as promoting gender conventionality and heterosexuality as the norm and "correct" way to be – that exist as background context in contemporary U.S. culture (e.g., Neilsen et al. 2000; Pascoe 2012), which some people may be responding to when asked to label their sexual identity, regardless of their sexual desires and history.

Findings from this study also illustrate that the extent to which sexual identity, attraction, and behavior aligns differs by gender. For example, I find that straight women age 15–45 are significantly more likely to report same-sex attraction and same-sex sexual experience than straight men of the same age. As another example, I find that gay women age 15–45 are significantly more likely to report other-sex attraction and sex with both sexes, both ever and in the last year, compared to gay men of the same age. This suggests that sexual identity may be more aligned with attraction and behavior for men than for women. It also suggests that women are more sexually fluid in terms of same-sex sex and same-sex attraction compared to men.

England et al. (2016) and Mishel et al. (forthcoming) argue that there is more same-sex sex among women than men, due to the asymmetry of the gender revolution. Specifically,

they argue that the gender revolution sent the message that gender nonconformity was more acceptable than before; and, given that deviations from exclusive heterosexuality are seen as gender nonconforming, part of the implicit message was permission to have same-sex sex. However, this message was received much more strongly by women than by men, because society continued to devalue anything seen as feminine. This meant that the gender revolution was largely a one-way street – that is, seen as less applicable to men. Thus, the social costs of any kind of gender nonconformity, including deviating from exclusive heterosexuality, decreased much more for women than for men (Mishel et al. forthcoming). What's more, many researchers have documented the strong ties between cultural notions of masculinity and of heterosexuality (e.g., Bridges 2014, Connell 1995; Connell and Messerschmidt 2005; Plummer 1999; Silva 2016; Ward 2008), which may also explain why women may be more sexually fluid in terms of same-sex sexuality compared to men. Indeed, Pascoe (2012) finds that boys in high school engage in what she terms "fag discourse," a form of gender policing where boys routinely call each other 'fags" to socialize one another into performing normative masculine behaviors and to enforce their heterosexuality, while such policing is not found among girls.

Figures 3 and 4 also revealed that men's predicted probabilities of identifying as gay are significantly higher than women's if they ever have engaged in same-sex sex. Recent experimental research by Mize and Manago (2018b) may shed some light as to why this may be. Mize and Manago (2018b) show that people allow women more freedom to experiment with same-sex sex without assigning them a lesbian sexual identity, while men are not allowed this same freedom to experiment with same-sex sexuality without people perceiving them to be gay. The argument put forth is that men's – but not women's – heterosexuality is an especially privileged identity that is easily lost if engaging in same-sex sexual behavior. Along the same lines, a study by Mishel et al. (2018) analyzes Google Trends data, and finds that people turn to Google to ask whether their sons, dads, boyfriends, husbands, brothers, and other boys/men are "gay" at consistently higher rates than analogous searches regarding their daughters, moms, girlfriends, wives, sisters, or other girls/women. Mishel et al. (2018) argue that because men's heterosexuality is seen as more precarious and more difficult to uphold than women's heterosexuality, it follows that people would question the heterosexuality of men and boys in their life more frequently than the women and girls they know. My findings seem to be consistent with this theory, in that, my results illustrate that men and women *self-identify* in alignment with the notion that men's heterosexuality is more precarious than women's heterosexuality. I posit that, due to the growing acceptance of same-sex sexuality among women but not men, and due to the more severe consequences for men when they break gender roles or identify as anything but heterosexual (Page and Yee 1985; Pascoe 2012; Pelligrini 1992), men are less likely than women to explore engaging in same-sex sex unless they are comfortable enough to also take on a gay or bisexual identity. Additional, qualitative research would be beneficial to further explore this notion.

Next, my analysis also reveals that a sizeable portion of Americans age 15–45 select "something else" or "don't know" when asked to label their sexual identity. Most research on the LGB community will typically exclude those who do not identify with a normative sexual identity such as lesbian, gay, or bisexual – so that individuals choosing something

else or don't know identities would be dropped from analysis. As a fraction of those who do not identify as straight, the proportion of men and women who would be dropped from analysis is quite substantial. As such, excluding those who do not choose a normative sexual identity category when conducting analysis on the LGB community or to explore measurement of the LGB population becomes problematic.

As a supplemental analysis, in results not shown, I ran the same logistic regressions with the same controls as in Supplemental material Tables A-2 and A-3 but used something else/don't know identity as the outcome variable, and then calculated predicted probabilities from these regressions. Predicted probabilities were very low for every attraction level and every sexual behavior measure (most under 3%, for both men and women). Because it is difficult to interpret non-normative sexual identity categories, some social scientists argue for utilizing open-ended responses when asking respondents to label their sexual identity (e.g., Better and Simula 2015), as it allows respondents to use their own words when defining their sexual identities. Others have recommended simply not including "other" or "something else" as a response option for sexual identity, and that if including "don't know" as responses, to make sure to specify *two* options: "I don't know yet," *and* "I don't know what this question means" (Badgett et al. 2009; Saewyc et al. 2004). I agree that using an open-ended response may be most accurate, but note that doing so may not be the most useful to quantitative researchers, as responses may vary greatly and grouping categories may become theoretically difficult. Yet, this may also illuminate ordinarily puzzling findings in the data, such as reports of same-sex sex but no same-sex attraction, which may be explained by someone's sexual identity as asexual, for example (see Poston and Baumle 2010). I also agree with Badgett et al. (2009) and Saewyc et al. (2004) that if national surveys include "I don't know," as a response option for sexual identity, then it is imperative to clarify whether the respondent does not understand the question, or just does not know how to label their sexual identity yet.

6.2. *Measure Limitations, Recommendations, and Implications*

This research sheds light on the complexity of defining sexual orientation using one of the three commonly used measures of identity, attraction, or behavior. This section will summarize some limitations of using each measure to define the LGB population. First, using only a sexual identity measure to define the LGB community can be limiting as more Americans report same-sex sexual behavior and same-sex attraction than report gay or bisexual identity. As discussed, there may be many reasons for this, including fear of societal costs and negative social stigma, or rejection of normative sexual identity categories altogether. In this sense, using only sexual identity to define the LGB community may lead to underestimation of the LGB population.

Second, using only sexual attraction to define the LGB population can be limiting, as more people report some level of same-sex attraction than report same-sex sexual experience or non-heterosexual identity. Yet, stigma and strong heteronormative scripts in American society may discourage these individuals from ever acting on their same-sex desire. Moreover, these individuals may not feel strongly enough about their same-sex attraction to ever act on it. In this sense, using only sexual attraction to define the LGB population may lead to overestimation of this population.

Finally, there are limitations to defining sexual orientation using only sexual behavior measures. For example, using only same-sex sexual behavior as the LGB measure omits large proportions of individuals who self-identify as gay or bisexual but who have no same-sex sexual experience. (Also, how it affects proportions depends on whether straight individuals are defined using sexual experience with other-sex partners – in this scenario, virgins would not be counted as heterosexual.) Recall that about 24% of bisexual women and about 16% of gay women aged 15–45 report 0 lifetime same-sex sex partners, while about 37% of bisexual men and about 16% of gay men aged 15–45 report 0 lifetime same-sex sex partners. Indeed, this phenomenon is also seen when analyzing sexual behaviors of Australian men and women aged 16–69, though percentages are slightly lower: 17% of Australian bisexual men and 9% of Australian bisexual women report 0 same-sex sex partners, as do 3% of Australian gay men and 4% of Australian gay women (Richters et al. 2014).

One reason that large proportions of self-identified gay and bisexual respondents have never had same-sex sex could be due to stigma around having same-sex sexual relations, as different moral and social implications come with homosexuality (Herek 2009); another reason could simply be due to a lack of access to other non-heterosexuals to date (Ghaziani 2014; Rosenfeld and Thomas 2012). In any case, the lack of same-sex sexual experience among some self-identified bisexual and gay individuals age 15–45 in the United States is important to note, as it suggests that measuring the LGB community using only sexual behavior may lead to underestimation of the LGB population. On the other hand, using this measure alone may incorrectly count self-reported straight individuals who have had same-sex sexual experiences for experimentation purposes in the past but no longer do so, or have no intention of doing so in the future – which may lead to *overestimating* the size of the LGB community. In this sense, it is difficult to say whether using only measures of sexual behavior to define the LGB community would lead to over- or underestimation of this population. Another consideration is whether to use recent sexual behavior or lifetime sexual behavior; lifetime behavior typically produces larger estimates, but recent behavior may be more accurate if seeking to assess the LGB community's current reach.

Because of these limitations, deciding which measure to use to define the LGB population in quantitative analyses can be a very difficult decision. Some researchers argue that recent sexual behavior is the best measure (Black et al. 2003); others argue that sexual attraction is the best measure, if limited to a single measure and analysis is on adolescents (Saewyc et al. 2004); others argue for a combination of the three measures (e.g., Baumle et al. 2009; Laumann et al. 1994); and yet other researchers argue that the specific research question should drive the measure (Compton et al. 2015). I agree with Compton et al. (2015): the specific measure of sexuality used in research on the LGB community should depend on the study's research question. After all, using a sexual behavior measure may be most appropriate for studies interested in examining sexually transmitted infections among sexual minorities, while it may be more appropriate to use a combination of sexual orientation measures for researchers interested in political and social change (Compton et al. 2015). Thus, social scientists interested in this topic should let their research question drive their measure of sexual orientation, while being transparent and clear about the measure they use to indicate LGB identity, and address any potential limitations of their measure of choice. If the research question does not strongly

justify using one measure over another, researchers should consider conducting the same analysis using more than one measure or using multiple measures combined (if the data allow them to), and address if and how results vary based on how they define the LGB community. After all, previous studies show differences in findings based on the measure of sexual orientation used. For example, Bostwick et al. (2010) find that identifying as LGB is associated with higher odds of anxiety disorders compared to identifying as straight, for both men and women; however, they also find that women who report only same-sex sex partners in their lifetime had the lowest rates of most disorders. As another example, Compton et al. (2015) find that, if using a behavior dimension of sexual orientation, bisexual men self-report as significantly less healthy compared to straight men; however, they find no significant differences in self-reported health between bisexual men and straight men when using an attraction or identity measure of sexual orientation.

As for what measure(s) to use to best assess the *size* of the LGB community, it ultimately depends on how one defines the "LGB community." As discussed, using sexual identity as the barometer of interest for the LGB population can be limiting, as fear of discrimination, stigma around non-heterosexual identity, prominence of heteronormative social scripts, and/or rejection of sexual identity labels may influence individuals who engage in same-sex sex or who feel same-sex attraction to not explicitly take on an LGB identity, which may translate to not marking it on a survey. In this sense, using multiple measures to assess the LGB population may be the most useful in calculating its reach. As such, this research allowed me to assess what the estimate for the LGB population may look like for those aged 15–45 if using a combination of sexual identity, behavior, and attraction. Specifically, I assessed what it would look like if estimates were based on individuals who checked at least one of the following: gay or bisexual identity, same-sex sexual behavior in the past year, or any same-sex attraction. Using this combination of measurement, I find that about 20% of women and 10% of men would comprise the LGB community (among those aged 15–45). This is much higher than the 6.4% of women and 3.6% of men aged 15–45 who self-identify as gay or bisexual.

In general, understanding how sexual attraction, sexual behavior, and sexual identity are linked is an important aspect of sexuality research and research on the LGB community. The disjunctures between self-reports of non-heterosexual identity, same-sex sexual behavior, and same-sex attraction among men and women age 15–45 in the United States that this study finds have implications for the current methods of measurement for the size of the LGB community – measurement that is critical as these figures inform a host of public policies, human rights issues, and anti-discrimination laws.

One limitation of this analysis is that there are no individuals above the age of 45, which motivates further research on the links between sexual identity, attraction, and behavior among U.S. adults older than 45. In any case, results from this analysis suggest that researchers should be open to more complex definitions of sexuality. In this regard, there has been increasing interest among social scientists to explore and analyze sexual identities and measures that exist on the borders of heterosexuality and homosexuality, such as queer, asexual, pansexual, fluid, or heteroflexible (Bridges 2014; Callis 2014; Compton et al. 2015; Savin-Williams and Vrangalova 2013; Silva 2016; Ward 2008). In general, more research is needed to fully understand whether estimates of health and well-being disparities, inequalities, and victimization of LGB individuals differ according to

the measure of sexual orientation that is used, as oftentimes, researchers do not have a choice in this decision when using existing secondary data. Because of the disjunctures between the three measures, and the fact that more men and women age 15–45 report same-sex attraction and same-sex sexual behavior than identify as gay or bisexual, it is important for surveys to ask other measures of sexual orientation rather than just identity (such as attraction or behavior), and for researchers to utilize multiple measures whenever appropriate, in order to obtain more accurate and all-encompassing analysis on the LGB community.

7. References

Baumle, A.K., D'Lane Compton, and D.L. Poston. 2009. *Same Sex Partners: The Demography of Sexual Orientation*. Albany, NY: SUNY Press.

Baumle, A.K. and D.L. Poston. 2011. "The Economic Cost of Homosexuality: Multilevel Analyses." *Social Forces* 89(3): 1005–1031. DOI: https://doi.org/10.1093/sf/89.3.1005.

Betts, P. 2009. "Developing Survey Questions on Sexual Identity: Cognitive/In-Depth Interviews." *Office for National Statistics*. Available at: https://www.ons.gov.uk/file?uri=/methodology/classificationsandstandards/sexualidentityguidanceandprojectdocumentation/sexidcogreportfinal_tcm77-181190.pdf (accessed October 2019).

Badgett, M.V.L. 1995. "The Wage Effects of Sexual Orientation." *ILR Review* 48(4): 726–739. DOI: https://doi.org/10.1177/001979399504800408.

Badgett, M.V.L. and Sexual Minority Assessment Research Team (SMART). 2009. "Best Practices for Asking Questions about Sexual Orientation on Surveys." *UCLA*: *The Williams Institute*. Available at: https://escholarship.org/uc/item/706057d5 (accessed October 2019).

Berg, N. and D. Lien. 2002. "Measuring the Effect of Sexual Orientation on Income: Evidence of Discrimination?" *Contemporary Economic Policy* 20(4): 394–414. DOI: https://doi.org/10.1093/cep/20.4.394.

Better, A. and B.L. Simula. 2015. "How and for Whom Does Gender Matter? Rethinking the Concept of Sexual Orientation." *Sexualities* 18(56): 665–680. DOI: https://doi.org/10.1177/1363460714561716.

Black, D., G. Gates, S. Sanders, and L. Taylor. 2002. "Why Do Gay Men Live in San Francisco?" *Journal of Urban Economics*: 54–76. DOI: https://doi.org/10.1006/juec.2001.2237.

Black, D.A., H.R. Makar, S.G. Sanders, and L.J. Taylor. 2003. "The Earnings Effects of Sexual Orientation." *Industrial and Labor Relations Review* 56(3): 449–469. DOI: https://doi.org/10.1177/001979390305600305.

Black, D., G.J. Gates, S.G. Sanders, and L. Taylor. 2007. "The Measurement Error of Same-Sex Unmarried Partner Couples in the U.S. Census." California Center for Population Research.

Bostwick, W.B., C.J. Boyd, T.L. Hughes, and S. Esteban McCabe. 2010. "Dimensions of Sexual Orientation and the Prevalence of Mood and Anxiety Disorders in the United States." *American Journal of Public Health* 100(3): 468–475. DOI: https://doi.org/10.2105/AJPH.2008.152942.

Bridges, T. 2014. "A Very 'Gay' Straight? Hybrid Masculinities, Sexual Aesthetics, and the Changing Relationship between Masculinity and Homophobia." *Gender and Society* 28(1): 58–82. DOI: https://doi.org/10.1177/0891243213503901.

Bridges, T. and M. Moore. 2018. "Young Women of Color and Shifting Sexual Identities." *Contexts* 18: 86–88. DOI: https://doi.org/10.1177/1536504218767125.

Budnick, J. 2016. "'Straight Girls Kissing'?: Understanding Same-Gender Sexuality beyond the Elite College Campus." *Gender & Society* 30(5): 1–24. DOI: https://doi.org/10.1177/0891243216657511.

Busseri, M., T. Willoughby, H. Chalmers, and A. Bogaert. 2006. "Same-sex attraction and successful adolescent development." *Journal of Youth and Adolescence* 35: 563–575. DOI:https://doi.org/10.1007/s10964-006-9071-4.

Callis, A.S. 2014. "Bisexual, pansexual, queer: Non-binary identities and the sexual borderlands." *Sexualities* 17(1/2): 63–80. DOI: https://doi.org/10.1177/1363460713511094.

Carpenter, C. 2009. "Sexual Orientation and Outcomes in College." *Economics of Education Review* 28: 693–703. DOI: https://doi.org/10.1016/j.econedurev.2007.10.010.

Carpenter, C.S. and S.T. Eppink. 2017. "Does It Get Better? Recent Estimates of Sexual Orientation and Earnings in the United States." *Southern Economic Journal* 84(2): 426–441. DOI: http://doi.wiley.com/10.1002/soej.12233.

Centers for Disease Control. 2018. "About the National Survey of Family Growth." Available at: https://www.cdc.gov/nchs/nsfg/about_nsfg.htm (accessed November 2018).

Chandra, A., W.D. Mosher, C. Copen, and C. Sionean. 2011. "Sexual behavior, sexual attraction, and sexual identity in the United States: Data from the 2006–2008 National Survey of Family Growth." National health statistics reports; no 36. Hyattsville, MD: National Center for Health Statistics. DOI: https://doi.org/10.1007/978-94-007-5512-3_4.

Cochran, S.D., J.G. Sullivan, and V.M. Mays. 2003. "Prevalence of mental disorders, psychological distress, and mental health services use among lesbian, gay, and bisexual adults in the United States." *Journal of consulting and clinical psychology* 71(1): 53–61. DOI: http://dx.doi.org/10.1037/0022-006X.71.1.53.

Compton, D.R., D.N. Farris, and Y-T. Chang. 2015. "Patterns of Bisexuality in America." *Journal of Bisexuality* 15: 481–497. DOI: https://doi.org/10.1080/15299716.2015.1048919.

Connell, R.W. 1995. *Masculinities*. Cambridge, UK: Polity Press.

Connell, R.W. and J.W. Messerschmidt. 2005. Hegemonic masculinity: Rethinking the concept. *Gender & Society* 19: 829–859. DOI: https://doi.org/10.1177/0891243205278639.

Copen, C., A. Chandra, and I. Febo-Vazquez. 2016. "Sexual behavior, sexual attraction, and sexual orientation among adults aged 18–44 in the United States: Data from the 2011–2013 National Survey of Family Growth." National health statistics reports; no 88. Hyattsville, MD: National Center for Health Statistics Available at: https://www.ncbi.nlm.nih.gov/pubmed/26766410 (accessed October 2019).

DeMaio, T.J., N. Bates, and M. O'Connell. 2013. "Exploring Measurement Error Issues in Reporting of Same-Sex Couples." *Public Opinion Quarterly* 77(S1): 145–158. DOI: https://doi.org/10.1093/poq/nfs066.

Diamond, L. 2008. "Female Bisexuality from Adolescence to Adulthood: Results from a 10-Year Longitudinal Study." *Developmental Psychology* 44(1): 5–14. DOI: http://dx.doi.org/10.1037/0012-1649.44.1.5.

England, P., E. Mishel, and M.L. Caudillo. 2016. "Increases in Sex with Same-Sex Partners and Bisexual Identity Across Cohorts of Women (but Not Men)." *Sociological Science* 3: 951–970. DOI: https://doi.org/10.15195/v3.a42.

Epstein, R., P. MckInney, S. Fox, and C. Garcia. 2012. "Support for a fluid-continuum model of sexual orientation: A large-scale internet study." *Journal of Homosexuality* 59(10): 1356–1381. DOI: https://doi.org/10.1080/00918369.2012.724634.

Fischer, M.M. 2016. "Identifying same-sex couples in cross-national survey data: a comparison of same-sex couples' demographic and socio-economic traits in six European countries." In *Nederland in context: verschillen en overeenkomsten*, edited by R. Meuleman, G. Kraaykamp, and M. Wittenberg, 50–79. Den Haag: DANS.

Friedman, S., A. Reynolds, S. Scovill, F. Brassier, R. Campbell, and M. Ballou. 2013. "An Estimate of Housing Discrimination against Same-Sex Couples." Prepared for U.S. Department of Housing and Urban Development. DOI: http://dx.doi.org/10.2139/ssrn.2284243.

Gallup. 2018. "Gay and Lesbian Rights." Retrieved November 2018. Available at: http://www.gallup.com/poll/1651/gay-lesbian-rights.aspx (accessed October 2019).

Gates, G. 2011. "How Many People are Lesbian, Gay, Bisexual, and Transgender?" *Williams Institute Report*. Available at: https://escholarship.org/uc/item/09h684X2 (accessed October 2019).

Gavrilets, S. and W.R. Rice. 2006. "Genetic models of homosexuality: generating testable predictions." *Proceedings of the Royal Society B: Biological Sciences* 273(1605): 3031–3038. DOI: https://doi.org/10.1098/rspb.2006.3684.

Geary, R.S., C. Tanton, B. Erens, S. Clifton, P. Prah, K. Wellings, K.R. Mitchell, J. Datta, K. Gravningen, E. Fuller, A.M. Johnson, P. Sonnenberg, and C.H. Mercer. 2018. "Sexual identity, attraction and behaviour in Britain: The implications of using different dimensions of sexual orientation to estimate the size of sexual minority populations and inform public health interventions." *PLOS One.* 13(1): e0189607: 3, 7, 9. DOI: https://doi.org/10.1371/journal.pone.0189607.

Ghaziani, A. 2014. *There Goes the Gayborhood?* Princeton University Press.

Hamilton, L. 2007. "Trading on Heterosexuality: College Women's Gender Strategies and Homophobia." *Gender & Society* 21: 145–172. DOI: https://doi.org/10.1177/0891243206297604.

Harper, G. 2007. "Sex isn't that simple: Culture and context in HIV prevention interventions for gay and bisexual male adolescents." *American Psychologist* 62(8): 806–819. DOI: http://dx.doi.org/10.1037/0003-066X.62.8.806.

Herek, G. 2009. "Sexual Stigma and Sexual Bias in the US: A Conceptual Framework." *Contemporary Perspectives on Lesbian, Gay, and Bisexual Identities* 54: 183–195. DOI: https://doi.org/10.1007/978-0-387-09556-1_4.

Igartua, K., B. Thombs, G. Burgos, and R. Montoro. 2009. "Concordance and discrepancy in sexual identity, attraction, and behavior among adolescents." *Journal of Adolescent Health* 45(6): 602–608. DOI: https://doi.org/10.1016/j.jadohealth.2009.03.019.

Kinsey, A., W. Pomeroy, and C. Martin. 1953. *Sexual Behavior in the Human Female*. Philadelphia, PA: WB Saunders.

Korchmaros, J., C. Powell, and S. Stevens. 2013. "Chasing Sexual Orientation: A Comparison of Commonly Used Single-Indicator Measures of Sexual Orientation." *Journal of Homosexuality* 60: 596–614. DOI: https://doi.org/10.1080/00918369.2013.760324.

Laumann, E.O., J. Gagnon, R. Michael, and S. Michaels. 1994. *The social organization of sexuality: Sexual practices in the United States*. Chicago, IL: University of Chicago Press.

Mishel, E. 2016. "Discrimination against Queer Women in the U.S. Workforce: A Résumé Audit Study." *Socius: Sociological Research for a Dynamic World* 2: 1–13. DOI: https://doi.org/10.1177/2378023115621316.

Mishel, E., P. England, J. Ford, and M.L. Caudillo. Forthcoming. "Cohort Increases in Sex with Same-Sex Partners: Do Trends Vary by Gender, Race, and Class?" *Gender & Society*.

Mishel, E., T. Bridges, and M.L. Caudillo. 2018. "Google, Tell Me. Is He Gay?: Masculinity, Homophobia, and Gendered Anxieties in Google Search Queries about Sexuality." *SocArXiv*. November 7. DOI: https://doi.org/10.31235/osf.io/4se75.

Mize, T.D. 2016. "Sexual Orientation in the Labor Market." *American Sociological Review* 81(6): 1132–1160. DOI: https://doi.org/10.1177/0003122416674025.

Mize, T.D. and B. Manago. 2018a. "The Stereotype Content of Sexual Orientation." *Social Currents*: 1–21. DOI: https://doi.org/10.1177/2329496518761999.

Mize, T.D. and B. Manago. 2018b. "Precarious Sexuality: How Men and Women Are Differentially Categorized for Similar Sexual Behavior." *American Sociological Review*. DOI: https://doi.org/10.1177/0003122418759544.

Neilsen, J.M., G. Walden, and C.A. Kunkel. 2000. "Gendered Heteronormativity: Empirical Illusions in Everyday Life." *Sociological Quarterly* 41(2): 283–296. DOI: https://doi.org/10.1111/j.1533-8525.2000.tb00096.x.

O'Connell, M. and G. Gooding. 2006. "The Use of First Names to Evaluate Reports of Gender and Its Effect on the Distribution of Married and Unmarried Couple Households." Paper presented at the Population Association of America Meeting, Date: 3/30/06, Place: Los Angeles, U.S.A. Avaliable at: https://paa2006.princeton.edu/papers/60208 (accessed October 2019).

O'Reilly, J.M., M.L. Hubbard, J. Lessler, P. Biemer, and C. Turner. 1994. "Audio and Video Computer Assisted Self-Interviewing: Preliminary Tests of New Technologies for Data Collection." *Journal of Official Statistics* 10(2): 197–214. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279921/ (accessed October 2019).

Page, S. and M. Yee. 1985. "Conception of Male and Female Homosexual Stereotypes among University Undergraduates." *Journal of Homosexuality* 12: 109–118. DOI: https://doi.org/10.1300/J082v12n01_06.

Pascoe, C.J. 2012. *Dude You're a Fag: Masculinity and Sexuality in High School*. Los Angeles, CA: University of California Press.

Pathela, P., A. Hajat, J. Schillinger, S. Blank, R. Sell, and F. Mostashari. 2006. "Discordance between Sexual Behavior and Self-Reported Sexual Identity: A Population-Based Survey of New York City Men." *Annals of Internal Medicine* 145: 416–425. DOI: https://doi.org/10.7326/0003-4819-145-6-200609190-00005.

Pelligrini, A. 1992. S(h)ifting the terms of Hetero/sexism: Gender, power, homophobia." In *Homophobia: How We All Pay the Price*, edited by W.J. Blumenfeld, 39–56. Boston: Beacon Press.

Plummer, D. 1999. *One of the boys: Masculinity, homophobia, and modern manhood*. New York: Harrington Park Press.

Poston, D.L. and A.K. Baumle. 2010. "Patterns of Asexuality in the United States." *Demographic Research* 23(18): 509–530. DOI: https://doi.org/10.4054/DemRes.2010.23.18.

Richters, J., D. Altman, P.B. Badcock, A.M.A. Smith, R.O. de Visser, A.E. Grulich, C. Rissel, and J.M. Simpson. 2014. "Sexual identity, sexual attraction and sexual experience: the Second Australian Study of Health and Relationships." *Sexual Health* 11: 451–460. DOI: https://doi.org/10.1071/SH14117.

Risman, B.J. 2018. *Where the Millennials Will Take Us*. Oxford University Press.

Rosenfeld, M.J. and R.J. Thomas. 2012. "Searching for a Mate: The Rise of the Internet as a Social Intermediary." *American Sociological Review* 77(4): 523–547. DOI: https://doi.org/10.1177/0003122412448050.

Rupp, L., V. Taylor, S. Regev-Messalem, Alison Fogarty, and Paula England. 2014. "Queer Women in the Hookup Scene: Beyond the Closet?" *Gender & Society* 28: 212–235. DOI: https://doi.org/10.1177/0891243213510782.

Saewyc, E.M., G. Bauer, C. Skay, L. Bearinger, M. Resnick, E. Reis, and A. Murphy. 2004. "Measuring sexual orientation in adolescent health surveys: Evaluation of eight school-based surveys." *Journal of Adolescent Health* 35(345): 1–15. DOI: https://doi.org/10.1016/j.jadohealth.2004.06.002.

Savin-Williams, R. and Z. Vrangalova. 2013. "Mostly Heterosexual as a Distinct Sexual Orientation Group: A Systematic Review of the Empirical Evidence." *Developmental Review* 33(1): 58–88. DOI: https://doi.org/10.1016/j.dr.2013.01.001.

Silva, T. 2016. "Bud-Sex: Constructing Normative Masculinity among Rural Straight Men That Have Sex With Men." *Gender & Society* XX(X): 1–23. DOI: https://doi.org/10.1177/0891243216679934.

Smith, A., C. Rissel, J. Richters, A. Grulich, and R. de Visser. 2003. "Sex in Australia: Sexual identity, sexual attraction and sexual experience among a representative sample of adults." *Australian and New Zealand Journal of Public Health* 27(2): 138–145. DOI: https://doi.org/10.1111/j.1467-842X.2003.tb00801.x.

Stein, A. 2010. "The Incredible Shrinking Lesbian World and Other Queer Conundra." *Sexualities* 13(1): 21–32. DOI: https://doi.org/10.1177/1363460709352724.

Tilcsik, A. 2011. "Pride and Prejudice: Employment Discrimination against Openly Gay Men in the United States." *American Journal of Sociology* 117(2): 586–626. DOI: https://doi.org/10.1086/661653.

Tourangeau, R. and T.W. Smith. 1996. "Asking Sensitive Questions: The Impact of Data Collection Mode." Question Format, and Question Context. *Public Opinion Quarterly* 60: 275–304. DOI: https://doi.org/10.1086/297751.

Villarroel, M.A., C.F. Turner, E. Eggleston, A. Al-Tayyib, S.M. Rogers, A.M. Roman, P.C. Cooley, and H. Gordek. 2006. "Same-Gender Sex in the United States: Impact of T-Acasi on Prevalence Estimates." *Public Opinion Quarterly* 70(2): 166–196. DOI: https://doi.org/10.1093/poq/nfj023.

Ward, J. 2008. "Dude-sex: White masculinities and 'authentic' heterosexuality among dudes who have sex with dudes." *Sexualities* 11(4): 414–434. DOI: https://doi.org/10.1177/1363460708091742.

Webster, M., S.J. Hysom, and E.M. Fullmer. 1998. "Sexual Orientation and Occupation as Status." *Advances in Group Processes* 15: 1–21.

Wellings, K., J. Field, A. Johnson, and J. Wadsworth. 1994. *Sexual Behavior in Britain: The National Survey of Sexual Attitudes and Lifestyles*. Penguin Books.

Received March 2018
Revised November 2018
Accepted May 2019

sciendo

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 885–911, http://dx.doi.org/10.2478/JOS-2019-0037

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

Jessica Holzberg[1], *Renee Ellis*[1], *Robin Kaplan*[2], *Matt Virgile*[1], *and Jennifer Edgar*[2]

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 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

Sexual orientation
[Self-response]: Which of the following best represents how you think of yourself ?
[Proxy response]: To the best of your knowledge, which of the following best represents how [NAME] thinks of themselves?

- Gay or Lesbian
- Straight, that is not gay, lesbian, or bisexual
- Bisexual
- Something else

Gender identity
Question 1: Sex at birth
[Self-response]: Was your sex recorded as male or female at birth?
[Proxy response]: To the best of your knowledge, was [NAME's] sex recorded as male or female at birth?

- Male
- Female

Question 2: Current gender identity
[Self-response]: Do you describe yourself as male, female, or transgender?
[Proxy response]: To the best of your knowledge, does [NAME] describe themselves as male, female, or transgender?

- 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 Surve." 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. Grynkiewicz, 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://ww2.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

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 913–918, http://dx.doi.org/10.2478/JOS-2019-0038

Editorial Collaborators

The editors wish to thank the following referees who have generously given their time and skills to the Journal of Official Statistics during the period October 1, 2018–September 30, 2019. An asterisk indicates that the referee served more than once during the period.

Abraham, Katharine, Joint Program in Survey Methodology, College Park, Maryland, U.S.A.*
Abramowitz, Joelle, University of Michigan, Ann Arbor, Michigan, U.S.A.*
Adeleke, Ismaila, University of Lagos, Lagos, Nigeria*
Adiguzel, Feray, Erasmus University Rotterdam, Rotterdam, the Netherlands*
Akun, Ozgur, University of St. Andrews, St. Andrews, Fife, UK*
Alaimo Di Loro, Pierfrancesco, Sapienza University of Rome, Rome, Italy
Antoni, Manfred, Institute for Employment Research, Nuremberg, Germany
Ashmead, Robert, Ohio Colleges of Medicine Government Resource Center, Columbus, Ohio, U.S.A.*
Ayhan, Oztas, Middle East Technical University, Ankara, Turkey
Aysoy, Cevriye, Central Bank of Turkey, Ankara, Turkey
Baffour, Bernard, University of Queensland, Brisbane, Australia*
Baker, Kellan, Johns Hopkins University, Baltimore, Maryland, U.S.A.
Bates, Nancy, U.S. Census Bureau, Washington, D.C., U.S.A.*
Benassi, Federico, Italian National Institute of Statistics, Rome, Italy
Beresovsky, Vladislav, National Center for Health Statistics, Hyattsville, Maryland, U.S.A.*
Bermudez Parrado, Silvia, Institute of Statistics and Cartography of Andalusia, Seville, Spain*
Bethlehem, Jelke, Leiden University, Leiden, the Netherlands
Białek, Jacek, University of Lodz, Lodz, Poland*
Bianchi, Annamaria, University of Bergamo, Bergamo, Italy
Biemer, Paul, RTI International, Durham, North Carolina, U.S.A.
Biffignandi, Silvia, University of Bergamo, Bergamo, Italy
Bijlsma, Ineke, Maastricht University, Maastricht, the Netherlands
Blumberg, Stephen, Center for Disease Control and Prevention, Atlanta, Georgia, U.S.A.
Boeschoten, Laura, Tilburg University, Tilburg, the Netherlands
Boonstra, Harm Jan, Statistics Netherlands, Heerlen, the Netherlands*
Bottone, Marco, Bank of Italy, Rome, Italy
Braaksma, Barteld, Statistics Netherlands, The Hague, the Netherlands
Brenner, Philip, University of Massachusetts Boston, Boston, Massachusetts, U.S.A.
Broome, Jessica, University of Michigan, Ann Arbor, Michigan, U.S.A.
Brunåker, Fabian, Valueguard Index Sweden, Uppsala, Sweden

Gemenis, Kostas, Max Planck Institute for the Study of Societies, Cologne, Germany
Geisen, Emily, RTI International, Durham, North Carolina, U.S.A.
Golinelli, Daniela, RAND Corporation, Santa Monica, California, U.S.A.*
Graf, Monique, University of Neuchâtel, Neuchâtel, Switzerland
Graham, Patrick, Statistics New Zealand, Christchurch, New Zealand
Grant, David, RAND Corporation, Santa Monica, California, U.S.A.
Grazzini, Jacopo, Eurostat, Luxembourg, Luxembourg
Groenitz, Heiko, Philipps-University Marburg, Marburg, Germany
Gubman, Yury, VATBox – Automated VAT Recovery and Governance, Tel Aviv, Israel
Guidi, Caterina, European University Institute, Florence, Italy*
Guy, Paula, Office for National Statistics, Titchfield, UK
Haan, Marieke, University of Groningen, Groningen, the Netherlands*
Haraldsen, Gustav, Statistics Norway, Kongsvinger, Norway
Haslett, Stephen, Massey University, Palmerston North, Manawatu, New Zealand*
Haque, Shovanur, Queensland University of Technology, Brisbane, Australia
Heldal, Johan, Statistics Norway, Oslo, Norway
Hill, Robert, University of Graz, Graz, Austria
Himelein, Kirsten, World Bank, Washington, D.C., U.S.A.
Hitchcock, David, University of South Carolina, Columbia, South Carolina, U.S.A.
Holbrook, Allyson, University of Chicago, Chicago, Illinois, U.S.A.
Horn, Stephen, Melba, Australia
Hox, Joop, Utrecht University, Utrecht, the Netherlands*
Hu, Jingchen, Vassar College, Poughkeepsie, New York, U.S.A.
Hua, Jianjun, Dartmouth College, Hanover, New Hampshire, U.S.A.
Huang, Ning, Statistics Canada, Ottawa, Ontario, Canada*
Iannaccone, Roberto, Italian National Institute of Statistics, Rome, Italy
Jang, Don, NORC at the University of Chicago, Bethesda, Maryland, U.S.A.*
Jans, Matt, University of California, Los Angeles, California, U.S.A.*
Jaspers, Eva, Utrecht University, Utrecht, the Netherlands
Jonsson, Fredrik, Statistics Sweden, Stockholm, Sweden
Joye, Dominique, University of Lausanne, Lausanne, Switzerland*
Junker, Christoph, Federal Statistical Office Health, Neuchâtel, Switzerland
Kaminska, Olena, University of Essex, Colchester Essex, UK
Karlsson, Sune, Örebro University, Örebro, Sweden
Karanka, Joni, Office for National Statistics, Cardiff, UK
Keeter, Scott, Pew Research Center for the People and the Press, Washington, D.C.,U.S.A.
Kennedy, Courtney, EW Research Center, Washington, D.C., U.S.A.
Ketzaki, Eleni, Aristotle University of Thessaloniki, Thessaloniki, Greece
Keusch, Florian, Lancaster University, Lancaster, UK
Kieffer, Annick, ENS, Paris, France
Kim, Jong-Min, University of Minnesota, Morris, Minnesota, U.S.A.
Kirby, Graham, University of St. Andrews, St. Andrews, Fife, UK
Koerner, Thomas, German Federal Statistical Office, Wiesbaden, Germany
Komaki, Yasuyuki, Nihon University, Tokyo, Japan
Kontula, Osmo, Family Federation of Finland, Helsinki, Finland

Pedlow, Steven, NORC/University of Chicago, Chicago, Illinois, U.S.A.
Peeters, Laurie, Catholic University of Leuven, Leuven, Belgium
Persson, Andreas, Statistics Sweden, Örebro, Sweden
Perry, Pier Francesco, University of Calabria, Cosenza, Italy
Pettersson, Esbjörn, Statistics Sweden, Stockholm, Sweden
Phipps, Polly, Bureau of Labor Statistics, Washington, D.C., U.S.A.
Piccolo, Domenico, University of Naples Federico II, Naples, Italy
Planas, Christophe, Joint Research Centre of EC, Ispra, Varese, Italy
Plewis, Ian, University of Manchester, Manchester, UK
Polettini, Silvia, Sapienza University of Rome, Rome, Italy*
Polidoro, Federico, Italian National Institute of Statistics, Rome, Italy
Powell, Ben, University of York, York, UK
Prasada Rao, D.S., University of Queensland, Brisbane, Australia
Presser, Stanley, University of Maryland, College Park, Maryland, U.S.A.
Proietti, Tommaso, University of Rome, Rome, Italy
Psarakis, Stelios, University of Athens, Athens, Greece
Quick, Harrison, Drexel University, Philadelphia, Pennsylvania, U.S.A.
Ralphs, Martin, Office for National Statistics, London, UK
Rambaldi, Alicia, University of Queensland, Brisbane, Australia
Ranalli, Maria Giovanna, University of Perugia, Perugia, Italy
Rancourt, Eric, Statistics Canada, Ottawa, Ontario, Canada
Rees, Phil, University of Leeds, Leeds, West Yorkshire, UK*
Reist, Benjamin, U.S. Department of Agriculture, Washington, D.C., U.S.A.*
Robbins, Michael, University of Missouri, Columbia, Missouri, U.S.A.
Rocchetti, Irene, High Council of the Judiciary, Rome, Italy
Rueda Garcia, Maria del Mar, University of Granada, Granada, Spain
Sacco, Sam, National Statistics Office, Valletta, Malta
Sala, Emanuela, University of Milan-Bicocca, Milan, Italy
Salgado, David, National Statistical Institute, Madrid, Spain
Salvati, Nicola, University of Pisa, Pisa, Italy
Scannapieco, Monica, Sapienza University of Rome, Rome, Italy
Schaeffer, Nora Cate, University of Wisconsin, Madison, Wisconsin, U.S.A.*
Schmid, Timo, Free University of Berlin, Berlin, Germany*
Schouten, Barry, Statistics Netherlands, The Hague, the Netherlands
Sengupta, Manisha, Center for Disease Control and Prevention, Hyattsville, Maryland, U.S.A.
Serfioti, Maria. European Commission (DG ESTAT), Luxembourg, Luxembourg
Shin, Hee-Choon, National Center for Health Statistics, Hyattsville, Maryland, U.S.A.*
Shlomo, Natalie, University of Manchester, Manchester, UK
Silvestre, Cláudia, ESCS-School of Communication and Media Studies, Lisbon, Portugal
Smith, Duncan, University of Manchester, Manchester, UK
Stagnitti, Marie, Agency for Healthcare Research and Quality, Rockville, Maryland, U.S.A.
Ståhl, Olivia, Statistics Sweden, Stockholm, Sweden*
Sverchkov, Michail, Bureau of Labor Statistics, Washington, D.C., U.S.A.

Journal of Official Statistics, Vol. 35, No. 4, 2019, pp. 919–921, http://dx.doi.org/10.2478/JOS-2019-0039

Index to Volume 35, 2019

Contents of Volume 35, Numbers 1–4

Author Index