CrossMark

Estimates of Non-Heterosexual Prevalence: The Roles of Anonymity and Privacy in Survey Methodology

Ronald E. Robertson[1] · Felix W. Tran[1] · Lauren N. Lewark[1] · Robert Epstein[1]

Abstract When do people feel comfortable enough to provide honest answers to sensitive questions? Focusing specifically on sexual orientation prevalence—a measure that is sensitive to the pressures of heteronormativity—the present study was conducted to examine the variability in U.S. estimates of non-heterosexual identity prevalence and to determine how comfortable people are with answering questions about their sexual orientation when asked through commonly used survey modes. We found that estimates of non-heterosexual prevalence in the U.S. increased as the privacy and anonymity of the survey increased. Utilizing an online questionnaire, we rank-ordered 16 survey modes by asking people to rate their level of comfort with each mode in the context of being asked questions about their sexual orientation. A demographically diverse sample of 652 individuals in the U.S. rated each mode on a scale from −5 (very uncomfortable) to +5 (very comfortable). Modes included anonymous (name not required) and non-anonymous (name required) versions of questions, as well as self-administered and interviewer-administered versions. Subjects reported significantly higher mean comfort levels with anonymous modes than with non-anonymous modes and significantly higher mean comfort levels with self-administered modes than with interviewer-administered modes. Subjects reported the highest mean comfort level with anonymous online surveys and the lowest with non-anonymous personal interviews that included a video recording. Compared with the estimate produced by an online survey with a nationally representative sample, surveys utilizing more intrusive methodologies may have underestimated non-heterosexual prevalence in the U.S. by between 50 and 414%. Implications for public policy are discussed.

Keywords Sexual orientation · Prevalence of non-heterosexuality · Survey methodology · Heteronormativity

Introduction

A long-standing question in survey research has been: When do people feel comfortable enough to provide honest answers to sensitive questions (Hyman, 1944; Parry & Crossley, 1950)? The results of surveys covering topics considered to be sensitive or stigmatized, such as a person's abortion history (Fu, Darroch, Henshaw, & Kolb, 1998; Jones & Forrest, 1992; Jones & Kost, 2007), criminal record (Kleck & Roberts, 2012; Preisendörfer & Wolter, 2013), drug use (Aquilino, 1994; Johnson & Fendrich, 2005; Morral, McCaffrey, & Iguchi, 2000; Tourangeau & Smith, 1996), or sexual inclinations (Coffman, Coffman, & Ericson, 2016; Gates, 2013a; Tourangeau & Smith, 1996; Villarroel et al., 2006), have been found to be especially prone to systematic distortions. For example, less than half of induced abortions in the U.S. are reported (Jones & Kost, 2007), over one-third of subjects from a German sample provided inaccurate reports about their criminal history (Preisendörfer & Wolter, 2013), and over 25% of students sampled from a US university did not report the failing or near failing grades that were, in fact, part of their academic history (Kreuter, Presser, & Tourangeau, 2008). One meta-analysis of self-report validation studies, covering 38 studies and 226 sensitive questions, estimated that 42% of sensitive behaviors are not reported in surveys (Lensvelt-Mulders, Hox, & Van Der Heijden, 2005).

This perpetual difficulty in producing valid and accurate prevalence estimates has been largely attributed to the fact

✉ Ronald E. Robertson
rrobertson@aibrt.org

[1] American Institute for Behavioral Research and Technology, 1035 East Vista Way, Ste. 120, Vista, CA 92084, USA

Springer

that people typically underreport sensitive or stigmatized information, while emphasizing or exaggerating socially valued attributes, attitudes or behavior (Barnett, 1998; Crowne & Marlowe, 1964; Krumpal, 2013; Lee, 1993; Paulhus, 2002; Rasinski, Willis, Baldwin, Yeh, & Lee, 1999; Singer, Von Thurn, & Miller, 1995; Stocké & Hunkler, 2007; Tourangeau, Rips, & Rasinski, 2000; Tourangeau & Smith, 1996)—a phenomenon that has been labeled "social desirability bias" (DeMaio, 1984; Paulhus, 1984, 1986). The strategy which people employ to manage their stigmatized social identities is known as "passing" (Kanuha, 1999) and is defined as "the management of undisclosed discrediting information about self" (Goffman, 1963, p. 42).

Among sensitive topics, surveys on sexual orientation are especially difficult to conduct for several reasons. First, unlike other sensitive behaviors (e.g., abortions or criminal history), there are no objective records of the public's sexual attractions, fantasies, behaviors, or identities, and objective measurements are virtually impossible to obtain due to the discreet and somewhat subjective nature of sexual orientation. Second, there are unique complications associated with the wording of sexual orientation-related questions and response options (Badgett, 2009; Savin-Williams, 2006; Vrangalova & Savin-Williams, 2012), as well as complications that can arise from the fluid and situationally dependent nature of a person's sexual orientation identity (Diamond, 2008; Kanuha, 1999; Katz-Wise, 2015; Mock & Eibach, 2012). Third, questions about one's sexual orientation are particularly sensitive due, in part, to a specific form of social desirability bias—heteronormativity (Warner, 1993). Heteronormativity is the preconception that heterosexuality is the socially accepted norm and that those who diverge from the norm are not only deviant but undesirable.

Although recent reports suggest that societal acceptance of the LGBT community is on the rise in the U.S. (Pew Research Center, 2013), the effects of heteronormativity are still prevalent, with 39% of LGBT adults in the U.S. reporting having been rejected by a family member or close friend because of their sexual orientation or gender identity, and 58% reporting having been the subject to slurs and jokes (Pew Research Center, 2013). Furthermore, 50% of 40,117 respondents in 40 countries believe homosexuality is "morally unacceptable" (Pew Research Center, n.d.), and in many countries people who are believed to be non-heterosexual have been subjected to indignations ranging from discrimination and hate crimes to enforced psychiatric treatments, torture, and execution (Itaborahy & Zhu, 2014; Kitzinger, 2005).

Obtaining Accurate Measures of Sexual Orientation

Given the particularly high sensitivity surrounding questions regarding one's sexual orientation (Coffman et al., 2016; Gates, 2013a, b) and the importance of obtaining accurate estimates of the non-heterosexual population for policy makers (Graham et al., 2011; Mayer et al., 2008; Sell & Holliday, 2014), an expert panel

of survey specialists and sexual orientation researchers was recently convened by the Ford Foundation to develop a guide to asking questions about sexual orientation on surveys. After a multi-year effort, these experts concluded that, when possible, sexual orientation-related questions should be placed on the self-administered portions of a survey (Badgett, 2009).

This advice is supported by the findings of researchers from a variety of fields who have found that in the U.S. the use of survey modes that increase a sense of anonymity in respondents can mitigate impression management (Dwight & Feigelson, 2000) and can increase reporting of sensitive or stigmatized beliefs, behaviors, and identities (Dillman, Smyth, & Christian, 2014; Durant, Carey, & Schroder, 2002; Krumpal, 2013; Villarroel et al., 2006). Researchers have also found that use of self-administered survey modes—such as self-administered questionnaires (SAQs), computer-assisted self-interviews (CASIs), and audio computer-assisted self-interviews (ACASIs)—can increase the likelihood of respondents reporting they have engaged in stigmatized or embarrassing behaviors when compared with interviewer-administered survey modes—such as face-to-face interviews (FTF), computer-assisted personal interviews (CAPI), and computer-assisted telephone interviews (CATI) (Aquilino & LoSciuto, 1990; Chang & Krosnick, 2010; Jones & Forrest, 1992; O'Reilly, Hubbard, Lessler, Biemer, & Turner, 1994; Tourangeau & Smith, 1996, 1998; Turner et al., 1998; Turner, Lessler, & Devore, 1992; Turner, Ku, Sonenstein, & Pleck, 1996).

Booth-Kewley, Larson, and Miyoshi (2007) found this to be especially true of computer-based self-administered surveys because "computers create an impersonal social situation in which individuals feel more anonymous, more private, more self-absorbed, less inhibited, and less concerned about how they appear to others" (p. 3), and this view has been echoed by others (e.g., Badgett, 2009; Baker, Bradburn, & Johnson, 1995; Bradburn et al., 1991; Buchanan, 2000; Gnambs & Kaspar, 2015; Joinson, 1999; Keeter, McGeeney, Igielnik, Mercer, & Mathiowetz, 2015; Lucas, Gratch, King, & Morency, 2014; Tourangeau & Smith, 1996; Tourangeau & Yan, 2007; Trau, Härtel, & Härtel, 2013). While some early studies found no significant advantage to computer-based self-administered surveys over traditional SAQs (Locke & Gilbert, 1995; Wright, Aquilino, & Supple, 1998), a meta-analysis published in 2000 of 30 studies yielding 77 effect sizes concluded that computer-based surveys were somewhat better than other measures (including SAQs) in increasing the self-disclosure of sensitive behaviors (Dwight & Feigelson, 2000), and a recent meta-analysis of 39 studies yielding 460 effect sizes confirmed this finding, adding that the effect was strongest for the most sensitive behaviors (Gnambs & Kaspar, 2015).

Despite the evidence in favor of using non-intrusive survey methodologies, and the availability of more accurate methods for assessing sexual orientation (Cerny & Janssen, 2011; Chivers, Rieger, Latty, & Bailey, 2004; Chivers, Seto, & Blanchard, 2007; Savin-Williams, 2006), the majority of national surveys conducted in the U.S. have utilized relatively intrusive methodologies to ask

about sexual identity, if they ask any sexual orientation questions at all (Sell & Holliday, 2014).

Taking the above factors into account, we set out to examine (1) the variability in 12 recent estimates of non-heterosexual identity prevalence in the U.S. as a function of the privacy and anonymity granted by the survey modes they employed, and (2) the comfort level that people report when asked to consider answering questions about their sexual orientation through 16 different survey modes.

Study 1: Non-Heterosexual Prevalence Estimates in the U.S.

We identified 12 nationwide surveys conducted between 2004 and 2015 that provided estimates of sexual orientation prevalence in the U.S. For each survey, we extract the available sexual orientation identity estimates and discuss the methods used to obtain them. Breakdowns of the estimates each survey produced by mode and response option are shown in Table 1.[1]

Method

We report the non-heterosexual estimates for each survey by taking the sum of explicit non-heterosexual identity survey responses ("Homosexual," "Gay or Lesbian," "Bisexual," or "Something else") and excluding ambiguous or non-responses ("Not sure," "Don't know," and refusals to answer). Though research has shown that the use of non-response options such as "I prefer not to answer" increases significantly when answers to sensitive questions are compared with answers to non-sensitive questions (Joinson, Woodley, & Reips, 2007), item non-response has not been shown to be a reliable indicator of item sensitivity, and non-responses cannot be assumed to be the result of social desirability (Beatty & Herrmann, 2002).

In order of the size of the prevalence estimates each survey produced, the surveys were: a 2004–2005 National Epidemiologic Survey on Alcohol and Related Conditions Wave 2 (NESARC), the 2013 National Health Interview Survey (NHIS), the 2004–2006 National Survey of Midlife Development in the United States (MIDUS II), the 2010 National Intimate Partner and Sexual Violence Survey (NISVS), a 2012 survey by the Gallup organization (Gates & Newport, 2012), the 2006–2010 National Survey of Family Growth (NSFG), the 2014 General Social Survey (GSS), the 2009–2012 National Health and Nutrition Exami-

nation Survey (NHANES), the 2010 National Survey of Sexual Health and Behavior (NSSHB), a 2015 online survey by the YouGov organization (Moore, 2015), a 2013 online survey (Coffman et al., 2016) administered on Amazon's Mechanical Turk (AMT)—an online subject pool frequently used by psychology and behavior researchers (Berinsky, Huber, & Lenz, 2012), and a 2012 survey that was freely available online (Epstein, McKinney, Fox, & Garcia, 2012).

Results

Among the 12 surveys we examined, we found that estimates of non-heterosexual identities were lower among surveys using interviewer-administered modes—with estimates ranging from 1.4 to 3.4% for FTF, CAPI, CATI, and telephone interviews—than among surveys using self-administered modes—with estimates ranging from 3.0 to 4.8% for SAQ, CASI, and ACASI (Table 1). Estimates obtained using Internet-based self-administered survey modes were the highest, with estimates ranging from 7.2 to 22.2% (Table 1). The prevalence of self-identified bisexuals appears to be especially predictable from the intrusiveness of the survey mode employed (Table 1). This could be due, in part, to a phenomenon known as "binegativity," which is defined as negative attitudes toward bisexuals from both males and females, and both heterosexuals and homosexuals (Feinstein, Dyar, Bhatia, Latack, & Davila, 2016; Rust, 2002). There are, however, several factors other than mode effects that could have contributed to the distortion of these estimates.

In terms of sampling, coverage, and response rates, the Coffman et al. (2016) and Epstein et al. (2012) surveys used self-selecting samples, MIDUS II was a follow-up survey that had a 30% attrition rate, and the NISVS and NSSHB had low response rates, which may indicate a higher degree of self-selection than is typical (Table 2). However, low response rates are not necessarily indicative of a biased sample (Chang & Krosnick, 2009; Curtin, Presser, & Singer, 2000; Keeter, Miller, Kohut, Groves, & Presser, 2000). For six of the 12 surveys we examined, we were unable to locate a response rate or a response rate was not provided by the researchers (Table 2). Given that the mean age of people who identify as LGBT in national surveys is often lower than the mean age of people who identify as heterosexual (Gates, 2014a), it is possible that the restricted age ranges of the NESARC Wave 2, MIDUS II, NSFG, and NHANES samples affected their estimates (Table 2). While the Coffman et al. (2016) study was posted on AMT, the Epstein et al. (2012) survey was freely available online and received both mainstream and LGB media coverage. Given that non-heterosexuals might be more likely to self-select to take an online test of this sort, the high figure obtained in this study probably overestimated the prevalence of non-heterosexuality in the general population.

There was also a high degree of variability in the question wording and response options provided by each survey: MIDUS

---

[1] Note that we have only included estimates from surveys that asked respondents directly about their sexual orientation. Data sources such as the American Community Survey and the US Census, which can be used to estimate sexual orientation preference inferentially, are omitted from our analysis. Such estimates are obtained by comparing data on the sex of the head of a given household, the sex of other members of the household, and the relationship between those other members of the household and the head of the household.

Table 1 Breakdown of non-heterosexual (NH) estimates by survey, mode, and response category

Survey	Mode	NH estimate[a]	Response options					
			Heterosexual; straight (%)	Homosexual; lesbian; gay	Bisexual	Other; something else	Don't know; not sure	Refused
NESARC (Wave 2)[b]	FTF	1.4%	98.1	0.8%	0.6%	NA	0.5%	NA
NHIS	CAPI	2.5%	96.6	1.6%	0.7%	0.2%	0.4%	0.6%
MIDUS II[b]	SAQ	3.0%	97.0	1.6%	1.4%	NA	NA	NA
NISVS[b]	CATI	3.4%	96.6	1.6%	1.8%	NA	NA	NA
Gallup[c,d]	Telephone	3.4%	92.2	NA	NA	NA	NA	4.4%
NSFG	ACASI	4.5%	94.5	1.5%	2.7%	0.3%	NA	1.0%
GSS	CASI	4.7%	94.0	1.9%	2.8%	NA	0.3%	1.7%
NHANES	ACASI	4.8%	93.6	1.8%	2.6%	0.4%	1.3%	0.2%
NSSHB	Online	7.2%	92.8	2.3%	3.4%	1.5%	NA	NA
YouGov	Online	9.0%	89.0	4.0%	4.0%	1.0%	NA	2.0%
Coffman et al.[c,e]	Online	11.3, 18.6%	88.7	NA	NA	NA	NA	NA
Epstein et al.	Online	22.2%	77.8	1.7%	17.7%	2.8%	NA	NA

FTF face-to-face, *CAPI* computer-assisted personal interview, *SAQ* self-administered questionnaires, *ACASI* audio computer-assisted self-interview, *CASI* computer-assisted self-interview

[a] Some rows may not add up to 100% due to rounding

[b] Some surveys did not provide "Other," "Don't know," or non-response options

[c] Some surveys did not allow for specific non-heterosexual subcategories to be identified

[d] Some surveys combined refusals to answer with "Don't know"

[e] Coffman et al. utilized ICT to produce two separate estimates

II and NHANES conflated sexual attraction with sexual orientation identity, and Gallup conflated gender with sexual orientation identity (Table 3). In addition, only NHIS and NHANES provided separate response options depending on the respondent's gender, and several of the surveys did not provide any type of "Other" response option (Table 3). Order effects could also have distorted the estimates, since only Gallup, GSS, and NHIS listed a non-heterosexual category as the first response option (Krosnick & Alwin, 1987). Unlike the rest of the surveys, which offered various sexual orientation labels as response options, Gallup and Coffman et al. (2016) only offered "yes," "no," or "don't know" responses. The questions preceding the sexual orientation question could also have affected the estimates, but what their impact might have been is uncertain (Bradburn & Mason, 1964; Lee, McClain, Webster, & Han, 2016).

In addition to the above areas, there were several areas of potential distortion that were unclear from the resources we were able to find. For example, whether participants were called by a computer or a person in the Gallup survey is not clear. When we contacted the Gallup organization to determine how participants were contacted, we were told that CATI, in which participants are asked questions over the phone by a human interviewer or an automated computer system and respond by pressing the number on their phone that corresponds to their answer, did not accurately describe Gallup's telephone interviewing process, and that

"to protect our proprietary systems and avoid divulging information to market competitors, we do not share our processes and methodology" (Gallup Client Support, personal communication, October 18, 2013). The NHIS also asked an unreported proportion of subjects the sexual orientation question over the phone (B. W. Ward, personal communication, April 13, 2015), which could have resulted in a mix of mode effects. The YouGov survey did not provide specific details about the sampling and weighting methods used, and the small sample size, as well as the relatively large margin of error of 4.2%, makes the validity of this estimate somewhat suspect (Moore, 2015).

Taking the above distortions into account, we believe that the six surveys that were least problematic and easiest to compare are NESARC, NHIS, NISVS, GSS, NHANES, and NSSHB. There are still notable differences among these surveys other than the mode of administration (including the year NESARC was conducted, the low response rate of NISVS, and the limited age range of NHANES), but given the paucity of data on non-heterosexual prevalence in the U.S. (Sell & Holliday, 2014), we believe these six surveys allow for the best possible comparison of survey mode as it affects estimates of non-heterosexual prevalence (Fig. 1). Utilizing a Welch's t test and a Cohen's d test modified for unequal variances (Bonett, 2008), we found that the mean estimate for surveys that involved an interviewer ($M = 2.4\%$, $SD = 1.0\%$) was significantly different from the mean estimate for surveys that

Table 2 Sampling and coverage details of U.S. non-heterosexual prevalence estimates

Survey	Year	Sample size	Ages	Response rate	Languages	Location	Sampling method	Coverage	References
NESARC (Wave 2)	2004–2005	34,653	20+	86.7%	N/A	Respondents' homes	Multistage stratified sample	Noninstitutionalized adults in the U.S.	Bostwick et al. (2010), NIAAA (2010)
NHIS	2013	34,557	18+	81.7%	English, Spanish	Respondents' homes	Multistage probability sample	Noninstitutionalized adults in the U.S.	Ward et al. (2014), CDC (n.d.-a)
MIDUS II	2004–2006	2560	35–84	70.0%	English	Respondents' homes	Multistage probability sample and RDD	Noninstitutionalized adults in the U.S.	Mock and Eibach (2012), Brim et al. (2011)
NISVS	2010	18,049	18+	27.5–33.6%	English, Spanish	Respondents' homes	RDD	Noninstitutionalized adults in the U.S.	Black et al. (2011), Walters et al. (2013)
Gallup	2012	121,290	18+	N/A	English, Spanish	Respondents' homes	RDD	Noninstitutionalized adults in the U.S.	Gates and Newport (2012)
NSFG	2006–2010	22,682	15–44	77.0%	English, Spanish	Respondents' homes	Multistage probability sample	U.S. households	Chandra, Mosher, and Copen (2013), NSFG (n.d.)
GSS	2014	2023	18+	N/A	English, Spanish	Respondents' homes and/ or over telephone	Two-stage probability sample	U.S. households	Smith, Marsden, and Hout (2015), Hinkins et al. (n.d.), NORC (n.d.)
NHANES	2009–2012	7070	18–59	N/A	English, Spanish, Mandarin, Cantonese, Vietnamese	Mobile examination center	Two-stage probability sample	Noninstitutionalized adults in the U.S.	Dahlhamer, Galinsky, Joestl, and Ward (2014), CDC (n.d.-b)
NSSHB	2010	5860	18+	52.6%	English, Spanish	Online	RDD and address-based sampling	U.S. households	Herbenick et al. (2010)
YouGov Omnibus	2015	1000	18+	N/A	English	Online	YouGov Omnibus survey panel	Noninstitutionalized adults in the U.S.	YouGov USA (n.d.), Moore (2015)
Coffman et al.	2016	2516	18+	N/A	English	Online	Self-selecting	Amazon's Mechanical Turk	Coffman et al. (2016)
Epstein et al.	2012	9444	18+	N/A	English	Online	Self-selecting	U.S. residents over 18	Epstein et al. (2012)

were self-administered ($M = 5.5\%$, $SD = 1.4\%$) ($t = 3.13$, $p < .05$; Cohen's $d = 1.97$, 95% confidence interval [CI] = 0.46, 3.50). Although we realize that in some respects we are comparing apples and oranges, the fact that there seems to be an orderly relationship in these studies between the intrusiveness of the survey mode and the magnitude of the prevalence found is suggestive. If we consider the nationally representative online NSSHB to be our ground truth—7.2% of the population identifies as non-heterosexual—then the prevalence of non-heterosexuals may have been underestimated by between 50% in the case of the relatively non-intrusive GSS CASI survey, to as much as 414% in the case of highly intrusive NESARC FTF survey (Fig. 1). However, the cause of these differences cannot be directly or solely attributed to the mode in which the survey was conducted (Dillman et al.,

Table 3 Question and response options used in U.S. non-heterosexual prevalence estimates

Survey	Sexual orientation questions	Answer choices and order	Preceding questions	References
NESARC (Wave 2)	Which of the categories on the card best describes you?	(1) Heterosexual (straight); (2) gay or lesbian; (3) bisexual; (4) not sure	N/A	Bostwick et al. (2010)
NHIS	Which of the following best represents how you think of yourself?	(1) Men: gay; women: lesbian or gay; (2) men: straight, that is, not gay; women: straight, that is, not lesbian or gay; (3) bisexual; (4) something else; (5) I don't know the answer	Demographics and health	Ward et al. (2014)
MIDUS II	How would you describe your sexual orientation? Would you say you are heterosexual (sexually attracted only to the opposite sex), homosexual (sexually attracted only to your own sex), or bisexual (sexually attracted to both men and women)?	(1) Heterosexual; (2) homosexual; (3) bisexual	Sexual history	Mock and Eibach (2012), Brim et al. (2011), Ryff et al. (2012)
NISVS	Do you consider yourself to be heterosexual or straight, gay or lesbian, or bisexual?	(1) Heterosexual or straight; (2) gay or lesbian; (3) bisexual	History of violent attacks and confrontations	Black et al. (2011), Walters et al. (2013)
Gallup	Do you, personally, identify as lesbian, gay, bisexual, or transgender?	(1) Yes; (2) no; (3) don't know/refused	N/A	Gates and Newport (2012)
NSFG	Do you think of yourself as …	(1) Heterosexual or straight; (2) men: homosexual or gay; women: homosexual, gay, or lesbian; (3) bisexual; (4) something else?	Sexual history and behavior	Chandra et al. (2013), NSFG (n.d.)
GSS	Which of the following best describes you?	(1) Gay, lesbian, or homosexual; (2) bisexual; (3) heterosexual or straight; (4) don't know; (5) refused	Random assignment: different sociopolitical attitudes	Smith et al. (2015), Hinkins et al. (n.d.)
NHANES	Do you think of yourself as…	(1) Men: heterosexual or straight (attracted to women); women: heterosexual or straight (attracted to men); (2) men: homosexual or gay (attracted to men); women: homosexual or lesbian (attracted to women); (3) bisexual (attracted to men and women); (4) something else; (5) not sure; (6) refused; (7) don't know	Sexual history, behavior, and STDs	Dahlhamer et al. (2014), (CDC 2009–2012)
NSSHB	Which of the following commonly used terms best describes your sexual orientation?	(1) Heterosexual/straight; (2) homosexual/gay or lesbian; (3) bisexual; (4) other	Sexual history and behavior, masturbation	Herbenick et al. (2010), Herbenick personal communication, January 30, 2014
YouGov Omnibus	With which group do you most closely identify?	(1) Heterosexual/straight; (2) lesbian/gay woman; (3) gay man; (4) bisexual; (5) other; (6) prefer not to say	N/A	YouGov USA (n.d.), Moore (2015)
Coffman et al.	Do you consider yourself to be heterosexual?	Direct response: (1) yes; (2) no Veiled response: ICT methodology	Demographics	Coffman et al. (2016)
Epstein et al.	Sexual orientation	(1) Straight; (2) gay or lesbian; (3) bisexual; (3) other	Demographics	Epstein et al. (2012)

2014). It is possible that other sensitive behaviors measured by these surveys would follow the same pattern. If comparable studies exist that do not fit this pattern, we are not aware of them.

Among the surveys we examined, only the GSS provided data that allowed us to compare mode effects within a survey.

We discovered from the researchers that a small portion of the GSS respondents were interviewed with a CATI mode rather than a CASI mode (T. Smith, personal communication, April 11, 2015). When we compared the non-heterosexual prevalence estimates between the participants surveyed by these
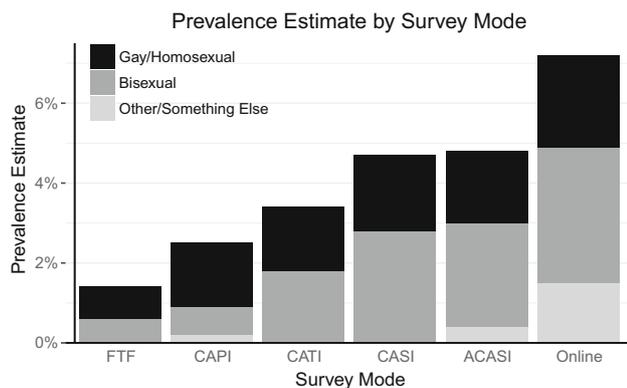
Fig. 1 Non-heterosexual prevalence estimates by mode and response option from six of the least problematic and easiest to compare US nationwide surveys. The FTF mode corresponds to the NESARC estimate, CAPI corresponds to the NHIS estimate, CATI corresponds to the NISVS estimate, CASI corresponds to the GSS estimate, ACASI corresponds to the NHANES estimate, and Online corresponds to the NSSHB estimate

two modes and employed weights provided by the GSS, we found that a larger portion of respondents reported a non-heterosexual identity with the CASI mode (4.7%) than with the CATI mode (1.8%) ($\chi^2[1] = 5.73$, $p < .05$), a finding that supports the notion that self-administered survey modes can produce larger and more accurate non-heterosexual prevalence estimates (Badgett, 2009).

The experimental design employed by Coffman et al. (2016) to produce two prevalence estimates is also relevant to the present study. Their participants were randomly placed into one of two treatment groups. In one group, item count technique (ICT) (Holbrook & Krosnick, 2010; Miller, 1984; Tsuchiya, Hirai, & Ono, 2007) was used to reduce social desirability bias and encourage honest responding. ICT, also known as an "unmatched count" or "list response" technique, is a within-subjects method in which a control group—the "Direct Report" group in this study—indicates how many of N questions are true, while an experimental group—the "Veiled Report" group in this study—indicates how many of $N + 1$ questions are true. By including a sensitive question in the Veiled Report group, researchers are able to estimate the population mean for the $N + 1$st item—the sensitive question—by comparing the mean number of affirmative responses in two groups. In this study, the $N + 1$st question was: "Do you consider yourself to be heterosexual?" and the only response options were "yes" or "no." Subjects in the Direct Report group answered the $N + 1$st on its own separate page after responding to the N list. Using this experimental design, the researchers found an 11.3% non-heterosexual prevalence estimate among the "Direct Report" group, and an 18.6% estimate among the "Veiled Report" group, a 64.2% increase (Coffman et al., 2016). The rationale behind ICT suggests that the larger value, 18.6%, is the more valid estimate (Miller, 1984).

The main problem with the Coffman et al. (2016) study is the sample. Although AMT has been shown to provide valid research results (Berinsky et al., 2012; Buhrmester, Kwang, & Gosling,

2011; Paolacci & Chandler, 2014; Peer, Vosgerau, & Acquisti, 2014), some evidence suggests MTurk samples are not representative of the US population (Chandler & Shapiro, 2016). That said, the ICT methodology employed by Coffman et al. avoids two problems inherent in most sexual orientation surveys: (1) the heteronormativity problem is reduced because participants do not have to answer the sexual orientation question directly, and (2) because the test is not identified as a sexual orientation test, participants are unlikely to self-select based on their interest in sexual orientation. On the surface, the survey contained entirely innocuous questions on a variety of subjects; the sexual orientation question was just slipped in.

From the estimates and methodologies we have gathered, it appears that little attention has been paid to the accumulating evidence on the importance of asking sensitive questions in particular ways (Badgett, 2009), and the dramatic variability we have found in estimates of sexual orientation prevalence in the U.S. suggests that most of the sexual orientation prevalence data available to policy makers may be distorted due to the use of survey methods known to produce distorted results.

This is alarming, because the need to collect accurate and reliable data on the LGBT population has been cited as a crucial step in researching and addressing disparities experienced in this diverse population (Bradford & Mayer, 2008; Cochran & Mays, 2006; Gates, 2013b; GLMA, 2001; Graham et al., 2011; Mayer et al., 2008; Sell & Becker, 2001; Sell & Holliday, 2014; Ward, Dahlhamer, Galinsky, & Joestl, 2014). These disparities have been found in areas such as physical health (Cochran & Mays, 2007), mental health (Bostwick, Boyd, Hughes, & McCabe, 2010; Conron, Mimiaga, & Landers, 2010; Diaz, Ayala, Bein, Henne, & Marin, 2001; Dilley, Simmons, Boysun, Pizacani, & Stark, 2010; Gilman et al., 2001; McLaughlin, Hatzenbuehler, & Keyes, 2010; Roberts, Austin, Corliss, Vandermorris, & Koenen, 2010), substance abuse (Conron et al., 2010; Dilley et al., 2010; Gilman et al., 2001; Gruskin, Greenwood, Matevia, Pollack, & Bye, 2007; McLaughlin et al., 2010), domestic violence (Walters, Chen, & Breiding, 2013), health insurance coverage (Gates, 2014b), and disability (Fredriksen-Goldsen, Kim, & Barkan, 2012). Such disparities have been found across all age groups of this population, from teenagers (Bradford & Mustanski, 2014; Kann et al., 2011; Mustanski, Van Wagenen, Birkett, Eyster, & Corliss, 2014; Remafedi, French, Story, Resnick, & Blum, 1998; Russell & Joyner, 2001) to the elderly (Fredriksen-Goldsen et al., 2012).

Study 2: Comfort Level by Survey Mode

How can we produce the most accurate estimates? To address this question, we surveyed a diverse group of people regarding how comfortable they would feel in answering questions about their sexual orientation given a wide range of survey methods—from, at one extreme, highly intrusive methods that preserve neither privacy nor anonymity to, at the other extreme, non-intrusive methods that preserve both. We hypothesized (1) that

self-reported comfort level would be predictable from the literature on sensitive topics and inversely related to the intrusiveness of the survey mode, and (2) that the comfort levels would be correlated with the non-heterosexual prevalence estimates obtained from the national sexual orientation surveys we analyzed.

Method

Participants and Procedure

On October 24, 2013, we administered our survey to 652 US participants through AMT with a solicitation entitled: "10–15 min opinion survey on comfortability with various research methods." Due to the nature of participation on AMT, no measure of response rate was collected. Each participant was paid US \$1 for participation in the study. The mean age was 33.60 years (SD = 11.44), there were slightly more male (58.9%) than female (41.1%) participants, and the sample was diverse demographically (Table 4).

Measures

We developed a list of survey mode scenarios that at one extreme were highly intrusive in both respects and that at the other extreme were not (see "Appendix"). We investigated eight interviewer-administered methods, such as face-to-face, CATI, and CAPI, and eight self-administered methods, such as SAQ, CASI, and online surveys. Each survey mode scenario had two versions: an anonymous (name is not required) version and a non-anonymous (name is required) version. The wording of these question pairs was otherwise identical ("Appendix").

In total, the questionnaire described 16 different survey modes that were listed in order, roughly, from least intrusive to most intrusive ("Appendix"). Each mode was described in a short paragraph, and participants were given the following instructions: "On a scale of −5 to +5 (where −5 means very uncomfortable and +5 means very comfortable), please rate how comfortable you would feel reporting accurate information about your sexual orientation (such as the opposite-sex or same-sex attractions you have felt or the opposite-sex or same-sex sexual encounters you have had) when asked in each of the ways listed below. Please note that some of these conditions are very similar, so read them carefully before answering."

Results

We examined the results of our survey in the context of the literature on mode effects and sensitive questions. We utilized nonparametric statistical tests such as Spearman's ρ, Mann–Whitney U, and Wilcoxon signed rank V throughout this study because our survey results lie on an ordinal scale. Subjects mean comfort level ratings varied greatly among survey modes and were consistent

with the literature on asking sensitive questions about sexual orientation. We found that anonymous survey modes elicited significantly higher mean comfort levels (M = 1.43, SD = 2.49) than non-anonymous modes (M = −0.21, SD = 3.33) in aggregate (V = 142,364; p < .001), as well as individually (Fig. 2). Among the anonymous modes we examined, participants reported the highest mean comfort level for online surveys (M = 3.93, SD = 2.05) and the lowest for CAPI-VR (M = −1.44, SD = 3.75). Among the non-anonymous modes we examined, the highest mean comfort level was again reported for online surveys (M = 1.02, SD = 3.61), and the lowest was reported for CAPI-VR (M = −1.81, SD = 3.74).

We also found support for the finding that self-administered surveys (M = 1.90, SD = 2.60) elicited more sensitive behavior reporting than interviewer-administered surveys (M = −0.70, SD = 3.36) (V = 162,012; p < .001). The finding that computers were valuable in eliciting higher rates of sensitive behavior reporting was also consistent with our data, with the CASI scenario having a significantly higher mean comfort level rating (M = 3.14, SD = 2.52) than its paper-and-pencil counterpart SAQ (M = 2.31, SD = 2.98) (V = 44,575; p < .001). As with other studies that sampled from a highly literate population (e.g., Gnambs & Kaspar, 2015), we did not find a significant difference for audio-enhanced computer surveys: the mean difference between CASI and ACASI modes was 0.2 (V = 9777; p = .99). The anonymous online survey mode not only elicited the highest mean comfort level rating from our participants (M = 3.93, SD = 2.05), but was also significantly higher than CASI surveys (V = 35,273; p < .001), supporting the notion that an online survey mode can enhance feelings of privacy compared with CASI, because no interviewer is involved in the administration of an online survey (Gnambs & Kaspar, 2015).

We found only one significant difference in comfort level for gender, race, education, income, sexual orientation, or age groups (Table 4). Specifically, a significant difference emerged among age groups for comfort level with non-anonymous survey modes. A significant negative correlation also emerged between continuous age and comfort level (Spearman's ρ = −0.13, p < .001). This relationship is supported by previous findings concerning age and non-heterosexual orientation prevalence estimates (Gates, 2014a).

To examine the effects of age, gender, and sexual orientation on mean comfort for anonymous, non-anonymous, self-administered, and interviewer-administered survey mode groups, we constructed four regression models. Given the ordinal nature of our response variable and the bimodal distribution in responses, we utilized logistic regressions and converted the mean comfort for each methodology to a binary response variable, where subjects received a one if they reported a mean comfort with that group of modes over zero (comfortable), and a zero otherwise (uncomfortable). We also collapsed the smallest demographic groups for age (45–64 and 65–74) and sexual orientation ("other" and "unsure"). Only the non-anonymous and the interviewer-

Table 4 Demographic characteristics and mean comfort level differences

Demographic		Percent	All methods			Anonymous			Non-anonymous		
			M	SD	Significance	M	SD	Significance	M	SD	Significance
Age	18–24	22.2	0.79	2.72	$\chi^2 = 182.3$	1.83	2.35	$\chi^2 = 82.3$	0.45	3.15	$\chi^2 = 103.2*$
	25–44	60.4	0.66	2.81		1.99	2.38		0.18	3.35	
	45–64	16.3	−0.40	2.85		1.18	2.67		−1.23	3.35	
	65–74	1.1	1.53	2.44		3.16	1.79		1.02	3.40	
Gender	Male	58.9	0.70	2.71	$U = 55,766.5$	1.98	2.30	$U = 54,603.5$	0.23	3.24	$U = 55,982.0$
	Female	41.1	0.28	2.95		1.63	2.62		−0.28	3.48	
Ethnicity	White	75.9	0.65	2.84	$\chi^2 = 175.7$	1.94	2.44	$\chi^2 = 49.5$	0.17	3.35	$\chi^2 = 85.2$
	Asian	9.7	0.21	3.14		1.57	2.65		−0.39	3.75	
	Black	5.7	0.53	2.81		1.74	2.40		0.04	3.47	
	Other	1.8	−1.76	2.32		−0.42	2.15		−2.62	2.89	
	Hispanic	5.1	0.63	2.39		1.94	2.24		0.14	3.51	
	American Indian	0.6	−1.09	2.69		0.43	2.76		−2.25	3.37	
	Unknown	1.2	−0.40	2.03		1.41	1.90		−1.46	2.62	
Sexual Orientation	Straight	90.5	0.51	2.83	$\chi^2 = 205.0$	1.84	2.46	$\chi^2 = 82.4$	−0.01	3.36	$\chi^2 = 80.0$
	Gay/lesbian	2.9	1.68	2.64		2.74	2.11		1.21	3.22	
	Bisexual	5.2	0.48	2.45		1.57	2.24		0.25	2.90	
	Other	0.9	0.45	2.46		2.02	1.65		0.05	3.12	
	Unsure	0.5	−3.35	1.27		−1.52	2.64		−4.48	0.50	
Education	None	0.9	−0.38	1.56	$\chi^2 = 159.6$	0.76	1.89	$\chi^2 = 63.3$	−0.62	1.74	$\chi^2 = 54.5$
	High School	31.6	0.34	2.89		1.69	2.45		−0.22	3.45	
	Associates	12.7	0.65	2.89		1.79	2.64		0.39	3.27	
	College	45.7	0.71	2.78		1.99	2.42		0.21	3.29	
	Masters	7.4	0.20	2.69		1.67	2.23		−0.44	3.36	
	Doctorate	1.4	0.76	2.90		2.57	2.29		−0.06	3.65	
	Unknown	0.3	−2.88	0.00		−0.36	0.71		−4.79	0.30	
Income	<$10,000	7.7	0.53	2.89	$\chi^2 = 155.2$	1.75	2.35	$\chi^2 = 72.2$	0.00	3.47	$\chi^2 = 68.0$
	$10,000–14,999	5.7	0.06	2.52		1.32	2.30		−0.40	3.01	
	$15,000–19,999	5.8	0.02	2.72		1.10	2.56		−0.31	3.10	
	$20,000–29,999	12.4	0.65	3.04		1.74	2.70		0.22	3.53	
	$30,000–39,999	13.2	0.59	2.73		1.96	2.19		0.12	3.32	
	$40,000–49,999	13.5	0.72	2.77		1.97	2.47		0.34	3.19	
	$50,000–74,999	19.5	0.49	2.86		1.96	2.40		−0.13	3.47	
	$75,000–99,999	11.7	0.61	2.95		2.06	2.60		−0.02	3.48	
	$100,000–149,999	4.4	1.39	2.94		2.70	2.27		0.80	3.62	
	>$150,000	2.9	0.54	2.53		1.64	2.00		0.20	3.04	
	I prefer not to say	3.1	−0.62	2.33		0.98	2.48		−1.41	2.88	

$*p < .05$

Mann–Whitney U tests were used to assess differences in average comfort level for gender because there were only two possible groups, and Kruskal–Wallis χ^2 tests were used to assess differences among the other demographic groups. Bonferroni corrections for multiple hypothesis testing were applied to the resulting p values

administered survey mode groups returned significant coefficients, and diagnostic plots of deviance residuals as well as Hosmer–Lemeshow goodness of fit tests (Hosmer & Lemeshow, 2000) suggested that both logistic response functions were appropriate (non-anonymous, $\chi^2 = 6.16$, $p = .10$; interviewer-administered, $\chi^2 = 0.45$, $p = .93$). Coefficients were adjusted for dispersion and are presented as odds ratios with 95% confidence intervals in Table 5. For non-anonymous survey modes, being in the 25–44 age group reduced the odds of being comfortable by 0.64 relative to the 18–24 group. For interviewer-administered

Fig. 2 The 16 survey modes described in the questionnaire grouped by anonymous and non-anonymous versions. Significant differences were found between anonymous and non-anonymous versions of each methodology. *Error bars* represent standard error of the mean
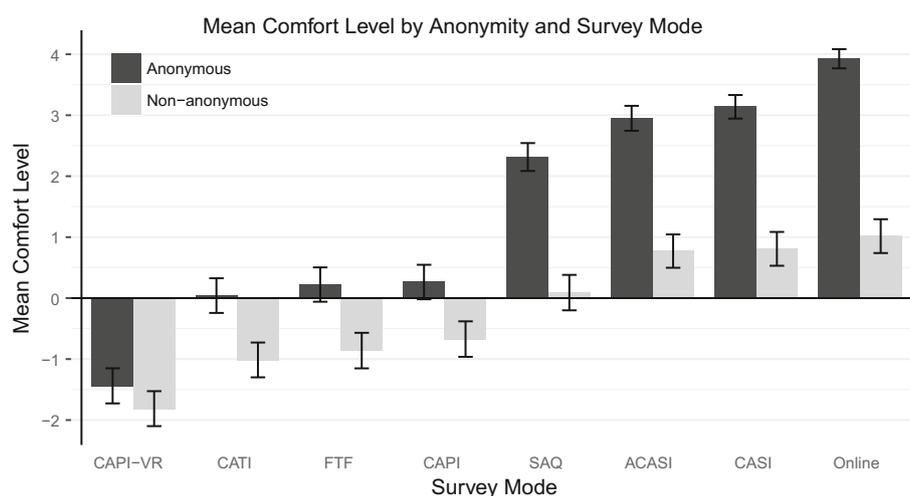

Table 5 Logistic regressions predicting mean comfort

Survey mode group	Predictor	OR	95% CI
Non-anonymous	Intercept	0.746*	(0.601, 0.953)
	25–44	0.637*	(0.449, 0.943)
	45–74	0.803	(0.622, 1.060)
	Female	0.796	(0.578, 1.121)
	Bisexual	1.379	(0.665, 2.762)
	Gay	1.915	(0.732, 4.870)
	Other or unsure	0.721	(0.150, 2.897)
Interviewer-administered	Intercept	0.603***	(0.489, 0.782)
	25–44	0.699	(0.488, 1.036)
	45–74	0.843	(0.649, 1.114)
	Female	0.920	(0.660, 1.291)
	Bisexual	0.950	(0.451, 1.944)
	Gay	2.880*	(1.072, 7.427)
	Other or unsure	0.441	(0.067, 1.979)

OR odds ratio, *CI* confidence interval

* $p < .05$; *** $p < .001$

survey modes, a self-reported sexual orientation of gay increased the odds of being comfortable by 2.88 relative to a sexual orientation of straight. This latter finding suggests that people who openly identify as gay might be comfortable discussing the details of their sexual orientation with an interviewer because they've already declared their preferences, whereas someone who identifies as straight might feel uncomfortable disclosing potentially dissonant sexual preferences to an interviewer.

When the prevalence estimates obtained from the six comparable studies we mentioned previously (Fig. 1) were compared to the mean comfort levels we obtained for their respective survey modes, we found a suggestive correlation (Spearman's $\rho = 0.77$, $p = .05$). Generally speaking, we found that the higher the comfort level of a survey mode, the higher the non-heterosexual prevalence estimate it produced.

Discussion

Given that the most common surveys of sexual orientation prevalence in the U.S. conducted in recent years employ relatively intrusive methodologies, we conclude that those surveys may greatly underestimate the prevalence of non-heterosexuals and, therefore, that current public policy making in this area may be based on inaccurate estimates. Our findings suggest that integrating non-intrusive administration modes, such as anonymous online surveys, into national data collection efforts will produce more accurate estimates of non-heterosexual prevalence and that the methodology least likely to underestimate the prevalence of non-heterosexuals has the following characteristics: (1) It is self-administered online; (2) it is fully automated; (3) it assures people that their participation is completely anonymous; (4) it uses recruit-

ment methods that do not compromise anonymity; and (5) it does not give participants the feeling of being observed or recorded.

We have taken liberties in the present study by comparing surveys that differ from one another not only in intrusiveness but also in other significant ways, including coverage, sampling methods, and non-response rate (Dillman et al., 2014). We defend these comparisons only by saying that these are, to our knowledge, all of the large-scale national surveys conducted in the U.S. over the past decade or so that have asked about sexual orientation in samples drawn from the general population, and that the limited manner in which we have compared these studies is not unreasonable.[2]

Given the wide range of prevalence estimates that have been obtained from such studies, we believe it is important that standards be set that will allow future studies to be conducted with a higher degree of validity than appears to exist in studies to date. While we realize that the primary purpose of each of the surveys we examined was not necessarily to produce accurate estimates of sexual orientation prevalence, we believe that any studies that ask about sexual orientation should be informed by relevant research (Badgett, 2009; Savin-Williams, 2006; Vrangalova & Savin-Williams, 2012).

The validity of our survey results is limited by our sample, which was drawn from a pool of highly experienced online survey takers (AMT). It is possible that people with limited or no online experience would express less comfort when asked about the possibility of taking tests online and that people with little or no experience taking surveys would express less comfort over the possibility of taking any surveys at all. However, a growing body of research suggests that data gathered on AMT are as valid as data collected in other settings (Berinsky et al., 2012; Casler, Bickel, & Hackett, 2013; Mason & Suri, 2012) and perhaps even superior to other commonly used subject pools (Casler et al., 2013; Chandler & Shapiro, 2016). Follow-up research should repeat our procedure with (1) a sample of people with experience taking surveys but with little online experience, (2) a sample of people who have little or no experience taking surveys, and (3) include control questions on non-sensitive behaviors or sensitive behaviors unrelated to sexuality. Based on research related to the present study (Booth-Kewley et al., 2007; Coffman et al., 2016; Gnambs & Kaspar, 2015; Villarroel et al., 2006), however, we conjecture that, in both cases, the overall pattern will be similar to the one we found—namely, that the more intrusive the survey mode, the less likely people will be to disclose sensitive information.

It is also possible that the comfort level ratings we found were affected by the order in which the items were presented; we did not vary the order (Appendix). We chose to use a fixed order for our questions in order to minimize confusion; the fixed order, we hoped, would draw attention to the specific restrictions we were adding to each question. A random order, we felt, could easily cause participants to overlook specific restrictions (since we were showing them long, compound sentences). Relevant research on order effects (e.g., Tourangeau, Couper, & Conrad, 2013) suggests that the fixed order of our questions might have distorted the ratings we found by between 0.20 and 0.30 points on an 11-point scale, but our comfort levels spanned a 6-point range (Fig. 1). We believe, therefore, that order was a contributing factor in the effect we observed but not a fatal confound. This issue should be explored in follow-up research.

Besides providing participants with a sense of anonymity and privacy, online surveys provide researchers with additional benefits such as reduced costs, rapid data collection, and easy access to large samples, both national and international (Rhodes, Bowie, & Hergenrather, 2003; Wright, 2005). Internet penetration in the USA is currently at 88.5 percent and is still increasing (Internet Live Stats, 2016), which means that Internet-based research will become even more advantageous and valid in future years. Despite these advantages, there are potential limitations to the validity of data collected through online surveys. Such surveys do not necessarily produce representative samples, which can limit the generalizability of the results (Rhodes et al., 2003; Wright, 2005). Some studies have found, however, that the validity of data collected from online surveys is comparable to that of data collected through SAQ (Fortson, Scotti, Ben, & Chen, 2006; Millar & Dillman, 2011; Nathanson & Reinert, 1999), telephone interviews (Chang & Krosnick, 2002, 2009; Graham & Papandonatos, 2008; Keeter et al., 2015), and CAPI (Ramo, Hall, & Prochaska, 2011)—the methods used in several of the national surveys we examined.

In spite of the limitations of the present study, we believe we have provided reasonably strong support for several key ideas that should be of interest to people working on sexual orientation issues, especially to people responsible for formulating relevant public policy: (1) survey methodology is critically important in determining non-heterosexual prevalence, (2) intrusive methodologies will likely lead to underestimates, and (3) for better or worse, technology is increasing our ability to elicit sensitive information from individuals.

Compliance with Ethical Standards

---

[2] National studies measuring non-heterosexual prevalence conducted in other countries, such as the UK (Dahlgreen & Shakespeare, 2015; Joloza, Evans, O'Brien, & Potter-Collins, 2010), were not included in our study. There is reason to believe, however, survey issues regarding sexual orientation are similar in other countries to the issues we raise in the present paper (e.g., see Burkill et al., 2016; Langhaug, Sherr, & Cowan, 2010).

declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Appendix: Questionnaire Content

Mode	Questionnaire item
Online	In any setting where Internet access is available to you (on your personal computer at home, a public library computer, etc.), you read questions about your sexual orientation on a computer screen and respond by entering your answer on a keyboard or a touch screen. No interviewer is present, and you are not required to give your name or contact information
Online (name required)	In any setting where Internet access is available to you (on your personal computer at home, a public library computer, etc.), you read questions about your sexual orientation on a computer screen and respond by entering your answer on a keyboard or a touch screen. No interviewer is present, but you are required to give your name and contact information
CASI	In a university setting or research facility— somewhere you've never been before— you read questions about your sexual orientation on a computer screen. You respond by entering your answers on a keyboard or touch screen. No interviewer is present, and no one in that setting knows your name
CASI (name required)	In a university setting or research facility— somewhere you've never been before— you read questions about your sexual orientation on a computer screen. You respond by entering your answers on a keyboard or touch screen. No interviewer is present, but you gave your name and contact information when you entered the setting
ACASI	In a university setting or research facility— somewhere you've never been before— you are presented with questions about your sexual orientation on a computer screen as you simultaneously hear an audio recording of the questions through a pair of headphones. You respond by entering your answer on a keyboard or a touch screen. No interviewer is present, and no one in that setting knows your name

Appendix continued

Mode	Questionnaire item
ACASI (name required)	In a university setting or research facility— somewhere you've never been before— you are presented with questions about your sexual orientation on a computer screen as you simultaneously hear an audio recording of the questions through a pair of headphones. You respond by entering your answer on a keyboard or a touch screen. No interviewer is present, but you gave your name and contact information when you entered the setting
SAQ	In a university setting or research facility— somewhere you've never been before— you answer questions about your sexual orientation on a piece of paper and return it to the administrator when you are done. You were not required to write your name or contact information on the paper, and no one in that setting knows your name
SAQ (name required)	In a university setting or research facility— somewhere you've never been before— you answer questions about your sexual orientation on a piece of paper and return it to the administrator when you are done. You were required to write your name and contact information on the paper
CATI	At home, you are called by an interviewer who says that he or she works for a university or a research organization. You are not asked to provide your name. The caller asks you questions about your sexual orientation and enters your answers into a computer
CATI (name required)	At home, you are called by an interviewer who says that he or she works for a university or a research organization. The caller asks you for your name, and you give it. The caller then asks you questions about your sexual orientation and enters your answers into a computer
CAPI	In a university setting or research facility— somewhere you've never been before— you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and enters your answers into a computer. No one in that setting knows your name
CAPI (name required)	In a university setting or research facility— somewhere you've never been before— you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and enters your answers into a computer. You gave your name and contact information when you entered that setting

Appendix   continued

Mode	Questionnaire item
CAPI (video recorded)	In a university setting or research facility—somewhere you've never been before—you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and enters your answers into a computer. No one in that setting knows your name, but a video recording is being made of your session
CAPI (name required and video recorded)	In a university setting or research facility—somewhere you've never been before—you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and enters your answers into a computer. You gave your name and contact information as part of the interview, and a video recording is being made of your session
Face-to-face	In a university setting or research facility—somewhere you've never been before—you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and takes notes as you answer. No one in that setting knows your name
Face-to-face (name required)	In a university setting or research facility—somewhere you've never been before—you are face-to-face with an interviewer whom you don't know. He or she asks you questions about your sexual orientation and takes notes as you answer. You gave your name and contact information as part of the interview

References

Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly, 58,* 210–240.

Aquilino, W. S., & LoSciuto, L. A. (1990). Effects of interview mode on self-reported drug use. *Public Opinion Quarterly, 54,* 362–393.

Badgett, M. V. (2009). *Best practices for asking questions about sexual orientation on surveys*. Los Angeles, CA: The Williams Institute.

Baker, R. P., Bradburn, N. M., & Johnson, R. A. (1995). Computer-assisted personal interviewing: An experimental evaluation of data quality and cost. *Journal of Official Statistics, 11,* 413–431.

Barnett, J. (1998). Sensitive questions and response effects: An evaluation. *Journal of Managerial Psychology, 13,* 63–76.

Beatty, P., & Herrmann, D. (2002). To answer or not to answer: Decision processes related to survey item nonresponse. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 71–85). New York, NY: Wiley.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20,* 351–368.

Black, M. C., Basile, K. C., Breiding, M. J., Smith, S. G., Walters, M. L., Merrick, M. T., & Stevens, M. R. (2011). *National intimate partner and sexual violence survey*. Atlanta, GA: Centers for Disease Control and Prevention.

Bonett, D. G. (2008). Confidence intervals for standardized linear contrasts of means. *Psychological Methods, 13,* 99–109.

Booth-Kewley, S., Larson, G. E., & Miyoshi, D. K. (2007). Social desirability effects on computerized and paper-and-pencil questionnaires. *Computers in Human Behavior, 23,* 463–477.

Bostwick, W. B., Boyd, C. J., Hughes, T. L., & McCabe, S. E. (2010). Dimensions of sexual orientation and the prevalence of mood and anxiety disorders in the United States. *American Journal of Public Health, 100,* 468–475.

Bradburn, N. M., Frankel, M. R., Hunt, E., Ingels, J., Schoua-Glusberg, A., Wojcik, M., & Pergamit, M. R. (1991, May 16–19). *A comparison of computer-assisted personal interviews (CAPI) with personal interviews in the national longitudinal survey of labor market behavior-youth cohort*. Lecture presented at the annual meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Bradburn, N., & Mason, W. (1964). The effect of question order on responses. *Journal of Marketing Research, 1,* 57–61.

Bradford, J. B., & Mayer, K. H. (2008). Demography and the LGBT population: What we know, don't know, and how the information helps to inform clinical practice. In H. Makadon, K. Mayer, J. Potter, & H. Goldhammer (Eds.), *The Fenway guide to lesbian, gay, bisexual, and transgender health* (pp. 25–41). Philadelphia, PA: American College of Physicians.

Bradford, J., & Mustanski, B. (2014). Health disparities among sexual minority youth: The value of population data. *American Journal of Public Health, 104,* 197. doi:10.2105/AJPH.2013.301818.

Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., & Shweder, R. A. (2011). *National Survey of Midlife Development in the United States (MIDUS), 1995–1996 (ICPSR02760-v8)*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Buchanan, T. (2000). Potential of the Internet for personality research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 121–140). San Diego, CA: Academic Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5.

Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., et al. (2016). Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles. *PLoS One, 11*(2), e0147983. doi:10.1371/journal.pone.0147983.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior, 29,* 2156–2160.

CDC. (n.d.-a). *National health interview survey questionnaires, datasets, and related documentation 1997 to the present*. Retrieved September 2, 2015 from http://www.cdc.gov/nchs/nhis/quest_data_related_1997_forward.htm.

CDC. (n.d.-b). *National health and nutritional examination survey (NHANES) 2011–2012 questionnaire data overview*. Retrieved from http://www.cdc.gov/nchs/nhanes/nhanes2011-2012/quexdoc_g.htm.

Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS). (2009–2012). *National health and nutrition examination survey questionnaire (or examination protocol, or laboratory protocol)*. Hyattsville, MD: U.S. Department of Health and Human Services, CDC.

Cerny, J. A., & Janssen, E. (2011). Patterns of sexual arousal in homosexual, bisexual, and heterosexual men. *Archives of Sexual Behavior, 40,* 687–697.

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology, 12,* 53–81.

Chandra, A., Mosher, W. D., & Copen, C. E. (2013). Sexual behavior, sexual attraction, and sexual identity in the United States: Data from the 2006–2010 national survey of family growth. In A. K. Baumle (Ed.), *International handbook on the demography of sexuality* (pp. 45–66). Dordrecht: Springer.

Chang, L., & Krosnick, J. A. (2002). *A comparison of the random digit dialing telephone survey methodology with Internet survey methodology as implemented by Knowledge Networks and Harris Interactive*. Columbus, OH: Department of Psychology, Ohio State University.

Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the Internet comparing sample representativeness and response quality. *Public Opinion Quarterly, 73,* 641–678.

Chang, L., & Krosnick, J. A. (2010). Comparing oral interviewing with self-administered computerized questionnaires: An experiment. *Public Opinion Quarterly, 74,* 154–167.

Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A sex difference in the specificity of sexual arousal. *Psychological Science, 15,* 736–744.

Chivers, M. L., Seto, M. C., & Blanchard, R. (2007). Gender and sexual orientation differences in sexual response to sexual activities versus gender of actors in sexual films. *Journal of Personality and Social Psychology, 93,* 1008–1121.

Cochran, S. D., & Mays, V. M. (2006). Estimating prevalence of mental and substance-using disorders among lesbians and gay men from existing national health data. In A. M. Omoto & H. S. Kurtzman (Eds.), *Sexual orientation and mental health: Examining identity and development in lesbian, gay, and bisexual people* (pp. 143–165). Washington, DC: American Psychological Association.

Cochran, S. D., & Mays, V. M. (2007). Physical health complaints among lesbians, gay men, and bisexual and homosexually experienced heterosexual individuals: Results from the California Quality of Life Survey. *American Journal of Public Health, 97,* 2048–2055.

Coffman, K. B., Coffman, L. C., & Ericson, K. M. M. (2016). The size of the LGBT population and the magnitude of anti-gay sentiment are substantially underestimated. *Management Science*. doi:10.1287/mnsc.2016.2503.

Conron, K. J., Mimiaga, M. J., & Landers, S. J. (2010). A population-based study of sexual orientation identity and gender differences in adult health. *American Journal of Public Health, 100,* 1953–1960.

Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly, 64,* 413–428.

Dahlgreen, W., & Shakespeare, A. (2015, August 16). *1 in 2 young people say they are not 100% heterosexual*. Retrieved from https://yougov.co.uk/news/2015/08/16/half-young-not-heterosexual/.

Dahlhamer, J. M., Galinsky, A. M., Joestl, S. S., & Ward, B. W. (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.

DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. Turner & E. Martin (Eds.), *Surveying subjective phenomena* (Vol. 2, pp. 257–282). New York, NY: Sage.

Diamond, L. M. (2008). *Sexual fluidity: Understanding women's love and desire*. Cambridge, MA: Harvard University Press.

Diaz, R. M., Ayala, G., Bein, E., Henne, J., & Marin, B. V. (2001). The impact of homophobia, poverty, and racism on the mental health of gay and bisexual Latino men: Findings from 3 U.S. cities. *American Journal of Public Health, 91,* 927–932.

Dilley, J. A., Simmons, K. W., Boysun, M. J., Pizacani, B. A., & Stark, M. J. (2010). Demonstrating the importance and feasibility of including sexual orientation in public health surveys: Health disparities in the Pacific Northwest. *American Journal of Public Health, 100,* 460–467.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: Wiley.

Durant, L. E., Carey, M. P., & Schroder, K. E. (2002). Effects of anonymity, gender, and erotophilia on the quality of data obtained from self-reports of socially sensitive behaviors. *Journal of Behavioral Medicine, 25,* 439–467.

Dwight, S. A., & Feigelson, M. E. (2000). A quantitative review of the effect of computerized testing on the measurement of social desirability. *Educational and Psychological Measurement, 60,* 340–360.

Epstein, R., McKinney, P., Fox, S., & Garcia, C. (2012). Support for a fluid-continuum model of sexual orientation: A large-scale Internet study. *Journal of Homosexuality, 59,* 1356–1381.

Feinstein, B. A., Dyar, C., Bhatia, V., Latack, J. A., & Davila, J. (2016). Conservative beliefs, attitudes toward bisexuality, and willingness to engage in romantic and sexual activities with a bisexual partner. *Archives of Sexual Behavior, 45,* 1535–1550.

Fortson, B. L., Scotti, J. R., Ben, K. S. D., & Chen, Y. C. (2006). Reliability and validity of an Internet traumatic stress survey with a college student sample. *Journal of Traumatic Stress, 19,* 709–720.

Fredriksen-Goldsen, K. I., Kim, H. J., & Barkan, S. E. (2012). Disability among lesbian, gay, and bisexual adults: Disparities in prevalence and risk. *American Journal of Public Health, 102,* e16–e21.

Fu, H., Darroch, J. E., Henshaw, S. K., & Kolb, E. (1998). Measuring the extent of abortion underreporting in the 1995 national survey of family growth. *Family Planning Perspectives, 30,* 128–138.

Gates, G. J. (2013a). Demographics and LGBT health. *Journal of Health and Social Behavior, 54,* 72–74.

Gates, G. J. (2013b). Geography of the LGBT population. In A. K. Baumle (Ed.), *International handbook on the demography of sexuality* (pp. 229–242). Dordrecht: Springer.

Gates, G. J. (2014a). *LGBT demographics: Comparisons among population-based surveys*. Los Angeles, CA: The Williams Institute.

Gates, G. J. (2014b). *In U.S., LGBT more likely than non-LGBT to be uninsured*. Retrieved from http://www.gallup.com/poll/175445/lgbt-likely-non-lgbt-uninsured.aspx.

Gates, G. J., & Newport, F. (2012). *Special report: 3.4% of U.S. adults identify as LGBT*. Retrieved from http://www.gallup.com/poll/158066/special-report-adults-identify-lgbt.aspx.

Gay and Lesbian Medical Association (GLMA). (2001). *Healthy People 2010: Companion document for lesbian, gay, bisexual and transgender (LGBT) health*. San Francisco, CA: GLMA.

Gilman, S. E., Cochran, S. D., Mays, V. M., Hughes, M., Ostrow, D., & Kessler, R. C. (2001). Risk of psychiatric disorders among individuals reporting same-sex sexual partners in the national comorbidity survey. *American Journal of Public Health, 91,* 933–939.

Gnambs, T., & Kaspar, K. (2015). Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis. *Behavior Research Methods, 47,* 1237–1259.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York, NY: Simon and Schuster.

Graham, R., Berkowitz, B., Blum, R., Bockting, W., Bradford, J., de Vries, B., … Makadon, H. (2011). *The health of lesbian, gay, bisexual, and transgender people: Building a foundation for better understanding*. Washington, DC: Institute of Medicine.

Graham, A., & Papandonatos, G. (2008). Reliability of Internet- versus telephone-administered questionnaires in a diverse sample of smokers. *Journal of Medical Internet Research, 10,* e8.

Gruskin, E. P., Greenwood, G. L., Matevia, M., Pollack, L. M., & Bye, L. L. (2007). Disparities in smoking between the lesbian, gay, and bisexual population and the general population in California. *American Journal of Public Health, 97,* 1496–1502.

Herbenick, D., Reece, M., Schick, V., Sanders, S. A., Dodge, B., & Fortenberry, J. D. (2010). Sexual behavior in the United States: Results from a national probability sample of men and women ages 14–94. *Journal of Sexual Medicine, 7,* 255–265.

Hinkins, S., Crego, A., English, N., Pedlow, S., & Wolter, K. M. (n.d.). *2010 national sample frame*. Retrieved from http://www.norc.org/Research/Projects/Pages/2010-national-sample-frame.aspx.

Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports tests using the item count technique. *Public Opinion Quarterly, 74,* 37–67.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.

Hyman, H. (1944). Do they tell the truth? *Public Opinion Quarterly, 8,* 557–559.

Internet Live Stats. (2016). *Internet users in the U.S.A. (2016*)*. Retrieved from http://www.internetlivestats.com.

Itaborahy, L., & Zhu, J. (2014). *State-sponsored homophobia, a world survey of laws: Criminalisation, protection and recognition of same-sex love* (9th ed.). Geneva: International Lesbian, Gay, Bisexual and Intersex Associates.

Johnson, T., & Fendrich, M. (2005). Modeling sources of self-report bias in a survey of drug use epidemiology. *Annals of Epidemiology, 15,* 381–389.

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers, 31,* 433–438.

Joinson, A. N., Woodley, A., & Reips, U. D. (2007). Personalization, authentication and self-disclosure in self-administered Internet surveys. *Computers in Human Behavior, 23,* 275–285.

Joloza, T., Evans, J., O'Brien, R., & Potter-Collins, A. (2010). *Measuring sexual identity: An evaluation report*. Newport, UK: Office for National Statistics.

Jones, E. F., & Forrest, J. D. (1992). Underreporting of abortion in surveys of U.S. women: 1976 to 1988. *Demography, 29,* 113–126.

Jones, R. K., & Kost, K. (2007). Underreporting of induced and spontaneous abortion in the United States: An analysis of the 2002 National Survey of Family Growth. *Studies in Family Planning, 38,* 187–197.

Kann, L., Olsen, E. O., McManus, T., Kinchen, S., Chyen, D., Harris, W. A., & Wechsler, H. (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. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 60*(7), 1–133.

Kanuha, V. (1999). The social process of passing to manage stigma: Acts of internalized oppression or acts of resistance? *Journal of Sociology and Social Welfare, 26*(4), 27–46.

Katz-Wise, S. L. (2015). Sexual fluidity in young adult women and men: Associations with sexual orientation and sexual identity development. *Psychology & Sexuality, 6,* 189–208.

Keeter, S., McGeeney, K., Igielnik, R., Mercer, A., & Mathiowetz, N. (2015, May 13). *From telephone to the web: The challenge of mode of interview effects in public opinion polls*. Retrieved from http://www.pewresearch.org/2015/05/13/from-telephone-to-the-web-the-challenge-of-mode-of-interview-effects-in-public-opinion-polls/.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64,* 125–148.

Kitzinger, C. (2005). Heteronormativity in action: Reproducing the heterosexual nuclear family in after-hours medical calls. *Social Problems, 52,* 477–498.

Kleck, G., & Roberts, K. (2012). What survey modes are most effective in eliciting self-reports of criminal or delinquent behavior? In L. Gideon (Ed.), *Handbook of survey methodology for the social sciences* (pp. 417–439). New York, NY: Springer.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and Web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly, 72,* 847–865.

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly, 51,* 201–219.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47,* 2025–2047.

Langhaug, L. F., Sherr, L., & Cowan, F. M. (2010). How to improve the validity of sexual behaviour reporting: Systematic review of questionnaire delivery modes in developing countries. *Tropical Medicine & International Health, 15,* 362–381.

Lee, R. M. (1993). *Doing research on sensitive topics*. Thousand Oaks, CA: Sage.

Lee, S., McClain, C., Webster, N., & Han, S. (2016). Question order sensitivity of subjective well-being measures: Focus on life satisfaction, self-rated health, and subjective life expectancy in survey instruments. *Quality of Life Research, 25,* 2497–2510.

Lensvelt-Mulders, G. J., Hox, J. J., & Van Der Heijden, P. G. (2005). How to improve the efficiency of randomised response designs. *Quality & Quantity, 39,* 253–265.

Locke, S. D., & Gilbert, B. O. (1995). Method of psychological assessment, self-disclosure, and experiential differences: A study of computer, questionnaire, and interview assessment formats. *Journal of Social Behavior and Personality, 10,* 255–263.

Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37,* 94–100.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44,* 1–23.

Mayer, K. H., Bradford, J. B., Makadon, H. J., Stall, R., Goldhammer, H., & Landers, S. (2008). Sexual and gender minority health: What we know and what needs to be done. *American Journal of Public Health, 98,* 989–995.

McLaughlin, K. A., Hatzenbuehler, M. L., & Keyes, K. M. (2010). Responses to discrimination and psychiatric disorders among Black, Hispanic, female, and lesbian, gay, and bisexual individuals. *American Journal of Public Health, 100,* 1477–1484.

Millar, M. M., & Dillman, D. A. (2011). Improving response to Web and mixed-mode surveys. *Public Opinion Quarterly, 75,* 249–269.

Miller, J. D. (1984). *A new survey technique for studying deviant behavior*. Doctoral dissertation, George Washington University.

Mock, S. E., & Eibach, R. P. (2012). Stability and change in sexual orientation identity over a 10-year period in adulthood. *Archives of Sexual Behavior, 41,* 641–648.

Moore, P. (2015). *A third of young Americans say they aren't 100% heterosexual*. Retrieved from https://today.yougov.com/news/2015/08/20/third-young-americans-exclusively-heterosexual/?sid=5388f1ffdd52b8ed110008bc&wpsrc=slatest_newsletter.

Morral, A. R., McCaffrey, D., & Iguchi, M. Y. (2000). Hardcore drug users claim to be occasional users: Drug use frequency underreporting. *Drug and Alcohol Dependence, 57,* 193–202.

Mustanski, B., Van Wagenen, A., Birkett, M., Eyster, S., & Corliss, H. L. (2014). Identifying sexual orientation health disparities in adolescents: Analysis of pooled data from the Youth Risk Behavior Survey, 2005 and 2007. *American Journal of Public Health, 104,* 211–217.

Nathanson, A. T., & Reinert, S. E. (1999). Windsurfing injuries: Results of a paper-and Internet-based survey. *Wilderness & Environmental Medicine, 10,* 218–225.

National Institute on Alcohol Abuse and Alcoholism (NIAAA). (2010). *U.S. alcohol epidemiological data reference manual* (2nd ed., Vol. 8). Retrieved from http://pubs.niaaa.nih.gov/publications/NESARC_DRM2/NESARC2DRM.pdf.

National Survey of Family Growth (NSFG). (n.d). *2006–2010 NSFG questionnaires*. Retrieved from http://www.cdc.gov/nchs/nsfg/nsfg_2006_2010_questionnaires.htm.

NORC at the University of Chicago. (n.d.). *2010 national sample frame*. Retrieved from http://www.norc.org/Research/Projects/Pages/2010-national-sample-frame.aspx.

O'Reilly, J., Hubbard, M., Lessler, J., Biemer, P., & Turner, C. F. (1994). Audio computer assisted self-interviewing: New technology for data collection on sensitive issues and special populations. *Journal of Official Statistics, 10,* 197–214.

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science, 23,* 184–188.

Parry, H. J., & Crossley, H. M. (1950). Validity of responses to survey questions. *Public Opinion Quarterly, 14,* 61–80.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 143–165). New York, NY: Springer.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 51–77). Mahwah, NJ: Lawrence Erlbaum Associates.

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46,* 1023–1031.

Pew Research Center. (2013, June 13). *A survey of LGBT Americans.* Retrieved from http://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/.

Pew Research Center. (n.d.). *Global views on morality.* Retrieved from http://www.pewglobal.org/2014/04/15/global-morality/.

Preisendörfer, P., & Wolter, F. (2013). Asking sensitive questions: An evaluation of the randomized response technique versus direct questioning using individual validation data. *Sociological Methods & Research, 42,* 321–353.

Ramo, D. E., Hall, S. M., & Prochaska, J. J. (2011). Reliability and validity of self-reported smoking in an anonymous online survey with young adults. *Health Psychology, 30,* 693–701.

Rasinski, K. A., Willis, G. B., Baldwin, A. K., Yeh, W., & Lee, L. (1999). Methods of data collection, perceptions of risks and losses, and motivation to give truthful answers to sensitive survey questions. *Applied Cognitive Psychology, 13,* 465–484.

Remafedi, G., French, S., Story, M., Resnick, M. D., & Blum, R. (1998). The relationship between suicide risk and sexual orientation: Results of a population-based study. *American Journal of Public Health, 88,* 57–60.

Rhodes, S. D., Bowie, D. A., & Hergenrather, K. C. (2003). Collecting behavioural data using the World Wide Web: Considerations for researchers. *Journal of Epidemiology and Community Health, 57,* 68–73.

Roberts, A. L., Austin, S. B., Corliss, H. L., Vandermorris, A. K., & Koenen, K. C. (2010). Pervasive trauma exposure among U.S. sexual orientation minority adults and risk of posttraumatic stress disorder. *American Journal of Public Health, 100,* 2433–2441.

Russell, S. T., & Joyner, K. (2001). Adolescent sexual orientation and suicide risk: Evidence from a national study. *American Journal of Public Health, 91,* 1276–1281.

Rust, P. R. (2002). Bisexuality: The state of the union. *Annual Review of Sex Research, 13,* 180–240.

Ryff, C., Almeida, D. M., Ayanian, J. S., Carr, D. S., Cleary, P. D., Coe, C., … Mroczek, D. K. (2012). *National Survey of Midlife Development in the United States (MIDUS II), 2004–2006 (ICPSR04652-v6).* Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Savin-Williams, R. C. (2006). Who's gay? Does it matter? *Current Directions in Psychological Science, 15,* 40–44.

Sell, R. L., & Becker, J. B. (2001). Sexual orientation data collection and progress toward Healthy People 2010. *American Journal of Public Health, 91,* 876–882.

Sell, R. L., & Holliday, M. L. (2014). Sexual orientation data collection policy in the United States: Public health malpractice. *American Journal of Public Health, 104,* 967–969.

Singer, E., Von Thurn, D. R., & Miller, E. R. (1995). Confidentiality assurances and response a quantitative review of the experimental literature. *Public Opinion Quarterly, 59,* 66–77.

Smith, T. W., Marsden, P. V., & Hout, M. (2015). *General social surveys, 1972–2014* [data file]. Retrieved from http://publicdata.norc.org:41000/gss/documents/BOOK/GSS_Codebook.pdf.

Stocké, V., & Hunkler, C. (2007). Measures of desirability beliefs and their validity as indicators for socially desirable responding. *Field Methods, 19,* 313–336.

Tourangeau, R., Couper, M. P., & Conrad, F. G. (2013). "Up means good": The effect of screen position on evaluative ratings in web surveys. *Public Opinion Quarterly, 77,* 69–88.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response.* Cambridge, MA: Cambridge University Press.

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions the impact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60,* 275–304.

Tourangeau, R., & Smith, T. W. (1998). Collecting sensitive information with different modes of data collection. In M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. Clark, J. Martin, W. L. Nicholls II, et al. (Eds.), *Computer assisted survey information collection* (pp. 431–454). New York, NY: Wiley.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133,* 859–883.

Trau, R. N., Härtel, C. E., & Härtel, G. F. (2013). Reaching and hearing the invisible: Organizational research on invisible stigmatized groups via web surveys. *British Journal of Management, 24,* 532–541.

Tsuchiya, T., Hirai, Y., & Ono, S. (2007). A study of the properties of the item count technique. *Public Opinion Quarterly, 71,* 253–272.

Turner, C. F., Forsyth, B. H., O'Reilly, J. M., Cooley, P. C., Smith, T. K., Rogers, S. M., & Miller, H. G. (1998). Automated self-interviewing and the survey measurement of sensitive behaviors. In M. P. Couper, R. P. Baler, J. Bethlehem, C. Z. Clark, J. Martin, W. Nichols, & J. M. O'Reilly (Eds.), *Computer assisted survey information collection* (pp. 455–473). New York, NY: Wiley.

Turner, C. F., Ku, L., Sonenstein, F. L., & Pleck, J. H. (1996). Impact of ACASI on reporting of male-male sexual contacts: Preliminary results from the 1995 national survey of adolescent males. In R. B. Warnecke (Ed.), *Health survey research methods* (pp. 171–176). Hyattsville, MD: National Center for Health Statistics.

Turner, C. F., Lessler, J. T., & Devore, J. (1992). Effects of mode of administration and wording on reporting of drug use. In C. F. Turner, J. T. Lessler, & J. C. Groerer (Eds.), *Survey measurement of drug use: Methodological studies* (pp. 177–220). Rockville, MD: National Institute on Drug Abuse.

Villarroel, M. A., Turner, C. F., Eggleston, E., Al-Tayyib, A., Rogers, S. M., Roman, A. M., … Gordek, H. (2006). Same-gender sex in the United States impact of T-ACASI on prevalence estimates. *Public Opinion Quarterly, 70,* 166–196.

Vrangalova, Z., & Savin-Williams, R. C. (2012). Mostly heterosexual and mostly gay/lesbian: Evidence for new sexual orientation identities. *Archives of Sexual Behavior, 41,* 85–101.

Walters, M. L., Chen, J., & Breiding, M. J. (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.

Ward, B. W., Dahlhamer, J. M., Galinsky, A. M., & Joestl, S. S. (2014). *Sexual orientation and health among U.S. adults: National health interview survey, 2013. National Health Statistics Reports (no. 77).* Atlanta, GA: Centers for Disease Control and Prevention.

Warner, M. (Ed.). (1993). *Fear of a queer planet: Queer politics and social theory* (Vol. 6). Minneapolis, MN: University of Minnesota Press.

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication.* doi:10.1111/j.1083-6101.2005.tb00259.x.

Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly, 62,* 331–353.

YouGov U.S. (n.d.). *YouGov Omnibus.* Retrieved from https://today.yougov.com/find-solutions/omnibus/.