

Design, Development and Testing of the NHIS Sexual Identity Question

Kristen Miller

J. Michael Ryan

Questionnaire Design Research Laboratory, Office of Research and Methodology,
National Center for Health Statistics

October 2011

This report describes research to develop and evaluate a sexual identity question for the National Health Interview Survey. Development and then evaluation of the question is based on findings from cognitive testing studies conducted by the Questionnaire Design Research Laboratory (QDRL), specifically, 7 previous testing projects as well as this current study which, taken together, consisted of a total of 386 in-depth cognitive interviews. (For final reports of previous projects, see Q-Bank at <http://wwwn.cdc.gov/QBANK/Home.aspx>). Additionally, data from the 2002 and 2006 National Survey of Family Growth (NSFG) were examined to further investigate findings from past cognitive interviewing studies. This report, first, defines the construct to be measured and then outlines known question design problems with existing sexual identity measures. It then presents a revised version of the question and the rationale for the new design. Finally, the report describes the method used to test the new question along with a detailed discussion of the findings. Results suggest that the new design will generate improved estimates of the lesbian, gay and bisexual population.

BACKGROUND

Construct Definition, Question Design and Question Response Problems

Prior to designing a survey question, it is necessary to identify the specific construct intended for capture by the particular survey question. For this project, the intended construct is sexual identity, which must be differentiated from other terms used to characterize the sexuality of populations. While the word ‘sexual orientation’ is most often used in today’s lexicon, the term itself is more of a catch-all term that does not specifically pertain to an actual, measurable phenomenon. In its essence the term has come to describe an aspect (or a conglomeration of aspects) that include a person’s history of sexual behavior, how they conceptualize and summarize their attractions toward opposite and same-gender people, and how they have come to understand and label their own selves. These three concepts—attraction, behavior and identity—although inter-related, pertain to different aspects of sexuality and are typically asked as separate questions in survey questionnaires.¹ Additionally, the three differing constructs may be of varying relevance to a particular research study. For example, a study intending to examine the spread of sexually transmitted diseases would likely be more

¹ QDRL has examined the performance of identity, behavior and attraction questions in previous testing projects. In these studies, findings reveal that these concepts, particularly attraction, are also complex phenomena and that they can be understood differently across groups of respondents. See Q-Bank for reports on specific behavior and attraction questions.

interested in respondents’ sexual histories as opposed to the label that individuals use to describe themselves.

In the context of the Healthy People 2020 directive that mandates health monitoring of disparity populations, the construct of sexual identity is the most appropriate because it most succinctly conveys an individual’s relationship to the disparity population. Sexual identity is best conceptualized as a concept of self that is formed within a social context and defines for individuals their relationship to other individuals, groups, and sociopolitical institutions within that context (Rust, 1993). Furthermore, identities are instrumental in organizing peoples’ lives and their everyday interactions, which hold important implication for individuals’ behaviors and others’ actions toward them (Cast, 2003). In the context of health, sexual identity is informative in understanding respondents’ access to health care and, subsequently, the quality of care they are provided. It is also informative in understanding risk factors such as diet, exercise, stress and smoking patterns as these factors are closely linked to community as well as self-conception. It is important to note that although individuals may conceptualize their identity within a framework of who they have sex with or who they are attracted to, behavior and attraction in and of themselves do not constitute identity. It is the meaning—specifically the interpretations that the individuals assign those behaviors and experiences—that defines how they ultimately conceptualize their identity (Plummer, 1981; 1995).

Measuring sexual identity on a survey questionnaire, however, presents unique challenges. Sexual identity is a complex concept that is rooted in social and political contexts and can change over the course of an individual’s life. Consequently, individuals’ sexual identities do not necessarily conform to discrete, objective and uniformly-defined categories. Additionally, as previous QDRL study of sexual identity questions revealed, the construct, itself, can differ substantially across various sexuality subgroups (Ridolfo et al, forthcoming). While the concept of ‘sexual identity’ holds a particularly distinct and salient meaning for those identifying as lesbian, gay, bisexual or transgender, many non-minority respondents do not hold salient sexual identities. Instead, these respondents (who for all intents and purposes would be categorized as being heterosexual), often dis-identify from a gay identity, possessing what is referred to as a ‘not-me’ identity (McCall, 2003). Rather than identifying as heterosexual, these respondents typically identify as ‘not gay’ or ‘normal.’ Table 1 summarizes the construct differences between minority and non-minority respondents.

Non-Minority Patterns	LGBT Patterns
Lack of salient sexual identity	Highly salient
No concept of sexual identity but rather dis-identification	Identity rooted in complex process of negotiating and forming a

<p>“not me,” “I’m normal,” “soy mujer,” “I don’t know”</p>	<p style="text-align: right;">sexual identity</p> <p style="text-align: center;">Shifting sexual identity; For transgender respondents, intersection of gender and sexuality</p>
--	---

This lack of construct comparability may generate relatively disparate data across minority and non-minority groups, though more significantly, it generates different types of response patterns. Table 2, then, illustrates that for both minority and non-minority respondents, misclassification and missing data errors can occur, however for different reasons. Non-minority respondents who do not identify with a particular sexual identity are not always familiar with the response categories, specifically, the terms ‘heterosexual,’ ‘homosexual’ and ‘bisexual.’ For example, previous cognitive interviewing studies found that respondents can confuse the words ‘homosexual’ and ‘heterosexual,’ believing that ‘heterosexual’ is the equivalent of being gay and that ‘homosexual’ is the equivalent of being straight. Additionally, some cognitive interviewing respondents, not knowing the terminology, surmise that the term ‘bisexual’ means ‘heterosexual,’ concluding that ‘bi’ means two: one man and one woman. This lack of understanding contributes to relatively high rates of missing data or misclassification (Ridolfo et al, forthcoming).

Table 2. Misclassification and missing data errors for minority and non-minority respondents

Non-Minority Patterns	Result	LGBT Patterns
<i>Because doesn’t know terminology</i>	High rate of ‘something else’	<i>Because uses another label</i>
<i>Because doesn’t know terminology</i>	High rate of ‘Don’t know’	<i>Because shifting sexual identity</i>
<i>Because believes implies heterosexuality</i>	Misclassification into ‘bisexual’	<i>Because interprets question as attraction or behavior</i>

These types of problematic response patterns can be contrasted with those found among LGBT respondents. While the problematic response patterns for non-minority respondents center on the lack of a salient sexual identity, problematic response patterns for LGBT respondents are rooted within the complex process of negotiating and forming

a sexual identity. The problematic response patterns found among LGBT respondents, then, relate to shifting sexual identities and use of non-traditional categories (e.g. queer, same-gender-loving), and for transgender respondents, the complex intersection between gender and sexuality. (For more detailed discussion, see Ridolfo et. al, forthcoming). Regarding the implication of question design, the contrast of problematic response patterns suggest that potential design solutions may be at odds for the two groups; while simplifying the question and providing concrete definitions related to sexual behavior and attraction may be the best for non-minority respondents, this solution would likely create more response problems for LGBT respondents. Previous QDRL work, however, has shown the importance of utilizing categories that respondents use in their everyday lives to describe themselves—a solution that is beneficial for both minority and non-minority respondents. As opposed to the more abstract, scientific labels (i.e. ‘homosexual’ and ‘heterosexual’) which respondents do not always understand and do not use to describe themselves, using the terms ‘straight,’ ‘lesbian,’ and ‘gay’ does indeed improve question performance for many respondents of both populations.

It is impossible to know the extent of misclassification in the survey data that is depicted in Table 2. Additionally, it is impossible to determine the extent to which misclassification is improved with the addition of the more meaningful categories. However, as previous QDRL work (Ridolfo, et al., forthcoming) has shown, it is possible to glean insight by examining those cases that fall into the missing categories, specifically, the respondents who refused or answered ‘don’t know’ or ‘something else.’ Table 3 below, which compares the 2002 NSFG and 2006 NSFG survey data, illustrates that survey data collected using the more abstract labels are associated higher rates of ‘something else,’ ‘refused’ and ‘don’t know’ responses. In the 2002 NSFG, in which respondents were only asked about being heterosexual, homosexual, bisexual or something else, a full 6.2 percent of the sample fell into the missing categories. With the simple addition of the terms ‘straight,’ ‘gay’ and ‘lesbian,’ (categories that respondents are more likely to use in their everyday lives) missing rates fell to 1.6 percent.

Table 3. Comparison of 2002 and 2006 NSFG sexual identity measures

	Heterosexual	Homosexual,	Bisexual,	Something else?	Don’t Know/Refused
	Heterosexual or straight,	Homosexual, gay, (or lesbian,)			
NSFG 2002-03	89.6	1.7	2.4	4.2	2.0
NSFG 2006-08	94.5	1.3	2.5	0.4	1.2

Most noteworthy, those missing cases in 2002 did not occur randomly. As illustrated in Table 4, those respondents with lower education were more likely to have ‘something else,’ ‘refused’ and ‘don’t know’ responses. In 2002, a full 14.4% of women with less than a high school education, in comparison to 2.1% of those with more than a high school diploma were more likely to be missing. The degree of this relationship drops significantly in the 2006 data—to 3.8 percent with less than a high

school diploma compared to 1.0 percent for those with more than a high school degree. While women have higher rates of missing data, it is important to note that the relationships mimic the same patterns among the data for men.

Table 4. Distribution of missing data by education in the 2002 and 2006-08 NSFG

Missing data	Men			Women		
	Less than high school	High school	More than high school	Less than high school	High school	More than high school
2002	11.4% (n=1361)	8.0% (n=1505)	2.1% (n=2055)	14.4% (n=1702)	7.9% (n=2167)	2.1% (n=3767)
2006	3.1% (n=1883)	1.6% (n=1590)	0.7% (n=2637)	3.8% (n=1960)	1.2% (n=1844)	1.0% (n=3522)
	2002: Rao-Scott Chi-square (2) = 63.47, p<.05 2006: Rao-Scott Chi-square (2) = 17.14, p<.05			2002: Rao-Scott Chi-square (2) = 240.28, p<.05 2006: Rao-Scott Chi-square (2) = 39.42, p<.05		

Note: Missing data = something else, refused and don't know responses

By and large, as illustrated in the above tables, the 2006 NSFG design for sexual identity represents a marked improvement from the 2002 design. However, response problems remain. Perhaps most problematic in the 2006 design, the Spanish version of the questionnaire provided no translation for the word 'straight' because there is no comparable word in Spanish. Interestingly, as shown in Table 5, while the rates of missing decreased most dramatically in 2006 for English interviews, the rate of missing for Spanish language respondents continues to be relatively high at 8.9 percent and 9.3 percent for Spanish-speaking men and women, respectively.

Table 5. Percentage of missing data by language and ethnicity in the 2002 and 2006-08 NSFG

Missing data	Men			Women		
	Spanish Hispanic Interview	English Hispanic Interview	English Non-Hispanic Interview	Spanish Hispanic Interview	English Hispanic Interview	English Non-Hispanic Interview
2002	12.1% (n=359)	10.6% (n=763)	5.1% (n=3793)	12.9% (n=558)	9.5% (n=1031)	5.6% (n=6037)
2006	8.9% (n=451)	1.3% (n=947)	1.0% (n=4708)	9.3% (n=546)	1.2% (n=1053)	1.1% (n=5716)

Note: Missing data = something else, refused and don't know responses

Additionally problematic, previous cognitive interview findings reveal that the word 'straight' is not always understood as intended among English-speaking respondents, who interpret the word to mean 'straight-laced.' For those respondents who also believe the word 'heterosexual' means being gay, simply inserting the word 'straight' does not alleviate problems with misclassification. Similarly, addition of the word 'straight' does not clarify the word 'bisexual' for respondents who believe the term implies

heterosexuality. To be sure, addition of the word ‘straight’ to the English version, alone, does not resolve comprehension problems entirely.

Examination of the relationship between education and missing rates also suggests that a problem remains. As illustrated in Table 4, women with less than a high school diploma are 1.8 times more likely than women with a high school diploma and 6.9 times more likely than women with more than a high school education to be missing. For men, those without a diploma were 1.4 times more likely than men with a high school diploma and 5.4 times more likely than those with more than a high school education to be missing. Interestingly, those women without a diploma were also 1.5 times more likely than those with a high school diploma and 1.9 times more likely than those with more than a high school education to answer ‘bisexual.’ Men without a diploma were 1.2 times more likely than those with a high school diploma and 2 times more likely than those with more than a high school education to answer ‘bisexual.’ The relationship between education and identifying as bisexual either reflects a true relationship or reflects a remaining comprehension problem. However, given the relatively high rates of missing that are related to education (which alone indicates problems with the measure), these authors believe it is more likely to be misclassification of those answering ‘bisexual.’ This argument, additionally, gains strength with the cognitive interview finding that some respondents believe the term implies being heterosexual.

Development of an Improved Sexual Identity Question

In designing a new question, the 2006 NSFG version was used as a point of departure because it was regarded as the best performing question to date on a survey. In order to improve upon the 2006 NSFG version, it was determined that the goals for the new question would be to 1) reduce misclassification of non-minority respondents, 2) reduce rates of ‘don’t know’ and ‘something else’ and 3), particularly for those respondents who do fall into ‘something else,’ be able to sort non-minority from minority sexual identity cases. Given these goals, revisions were based on findings from previous QDRL studies which were presented in the above discussion. The revised question and the follow-up questions are presented below. Revisions were based on the following design principles: 1) use labels that respondents use to refer to themselves, 2) do not use labels that some respondents do not understand—particularly if those misunderstood terms are not required by any other group of respondents, and 3) use follow-up questions to meaningfully categorize those respondents answering ‘something else’ and ‘don’t know.’

Do you think of yourself as:

[For men:] Gay

[For women:] Lesbian or gay

[For men:] Straight, that is, not gay

[For women:] Straight, that is, not lesbian or gay

Bisexual

Something Else (*Go to A*)

Don’t Know (*Go to B*)

<p>A. [If 'something else' is selected] By something else, do you mean that...</p> <p>You are not straight, but identify with another label such as queer, trisexual, omniseual or pan-sexual</p> <p>You are transgender, transsexual or gender variant</p> <p>You have not figured out your sexuality or are in the process of figuring it out</p> <p>You do not think of yourself as having sexuality</p> <p>You do not use labels to identify yourself</p> <p>You made a mistake and did not mean to pick this answer</p> <p>You mean something else (<i>Go to C</i>)</p>
<p>B. [If 'don't know' is selected] You did not enter an answer for the question. That is because you:</p> <p>You don't understand the words</p> <p>You understand the words, but you have not figured out your sexuality or are in the process of figuring it out</p> <p>You mean something else</p>
<p>C. [If 'you mean something else' is selected]</p> <p>What do you mean by something else? Please type in your answer</p> <p>_____</p>

The following discussion outlines the design principles and presents the rationale behind the new version.

1) Use labels that respondents use to refer to themselves.

For minority group respondents, the word 'gay' (and, for women, the phrase 'lesbian and gay') was retained, while the word 'homosexual' was dropped. By far, as was discovered in the 386 interviews conducted in the QDRL since the year 2000, the most common labels used by sexual minorities to describe themselves are 'gay,' 'lesbian,' and 'bisexual.' In many cases, these respondents described the term 'homosexual' as being overly scientific or connoting sexual deviance, not an affirmative label for gay and lesbian people. Besides 'lesbian' or 'gay,' other less common words were used by sexual minority respondents to describe themselves, including 'queer,' 'same-gender loving,' 'pansexual,' and 'trisexual.' For these respondents who do not identify with the more commonly used categories, the category 'something else' was provided and, from this question, they would be taken to a follow-up question that would allow them to clarify how they identify. For the non-minority category label, the word 'heterosexual' was replaced with the phrase, 'not gay,' because many non-gay respondents did not associate themselves with a sexual identity, but rather maintained a 'not-me' identity.

2) Do not use labels that some respondents do not understand—particularly if those misunderstood terms are not required by any other respondent.

The term 'heterosexual' was dropped because many respondents, particularly those of lower socio-economic status, did not accurately understand the term. Significantly, in some cases, it was understood to mean 'being gay.' In no case did a

respondent identify as heterosexual and require the heterosexual label; respondents either thought of themselves as not gay or as straight, or know that the word ‘straight’ implied heterosexuality. The degree of confusion over the terminology along with the potential for misclassification required that the new question at least be tested without the word ‘heterosexual.’ The term ‘bisexual’ was also often misunderstood by lower SES non-minority respondents. However, this term could not be replaced because it is a term that is commonly used by bisexual respondents. Furthermore, it was hoped that the preceding category, ‘straight, that is, not gay,’ would be clear enough that non-minority respondents would not be compelled to erroneously choose ‘bisexual’ thinking it implied heterosexuality.

3) *Use follow-up questions to meaningfully categorize those respondents answering ‘something else’ and ‘don’t know.’*

Because missing data continued to be a problem in 2006 NSFG (specifically, the percentage of missing cases (1.2%) was essentially the same (1.3%) as those reporting as gay or lesbian), it was determined that an improved version would attempt to successfully classify these respondents into meaningful categories. While it is hoped that use of ‘not gay’ and the elimination of the terms ‘homosexual’ and ‘heterosexual’ will generate less cases of missing data, we do know that some minority respondents fall into ‘something else’ or are still figuring out their sexual identity. A follow-up question was added, then, to sort these cases into the proper non-minority category.

One potentially controversial aspect of this revised question design is the response ordering. The general practice in question design is to place the most chosen response option first. One reason for this is the primacy effect, that is, due to cognitive burdening, respondents spend more time processing earlier options (Krosnick and Alwin, 1987). There are at least two reasons, however, why this should not be the case for the question. The first, and most obvious, reason is that because the heterosexual option reads ‘straight, that is, not gay,’ and, therefore, requires the ‘gay’ option to be listed previously. The second reason, and related to the first, is that respondents engage in satisficing, that is they look for the first option that is reasonable, even if not optimal (Tourangeau, Rips, and Rasinski, 2000). By placing the non-minority response lower in the list it encourages respondents to more deeply consider previous response options.

METHOD OF EVALUATION FOR THE REVISED QUESTION

The method used to examine the performance of the newly revised sexual identity question was cognitive interviewing. Cognitive interviewing is the primary method used by the federal statistical community to ensure data quality. It is also the only known method that can provide insight into question validity, that is, insight into the phenomena that a question actually captures—the substance that makes the statistic. The aim of cognitive interviewing is to investigate how survey questions perform when asked of respondents, specifically, how respondents understand a question and how they go about forming an answer. Cognitive interviewing is a qualitative method that

provides rich, contextual information regarding the ways respondents 1) interpret a question, 2) consider and weigh out relevant aspects of their lives and, finally, 3) formulate a response based on that consideration. As such, cognitive interviewing provides in-depth understanding of the ways in which a question operates, the kind of phenomena that it captures, and whether or not it ultimately serves the scientific goal. Findings from a cognitive interviewing project typically lead to recommendations for improving a survey question, or results can be used in post-survey analysis to assist in data interpretation. (For more thorough discussion of cognitive interviewing, see Miller, 2011).

Recruitment and Respondent Demographics

To test the newly revised question, the QDRL conducted 139 cognitive interviews: 94 in English and 45 in Spanish. These interviews were conducted on-site at the QDRL interview lab in Hyattsville, Maryland as well as at several off-site locations including The DC Center for the LGBT Community, Mpoderate (a center for Latino gay male and transgender youth), Casa de Maryland, and a rented office building located in the Colombia Heights neighborhood of Washington, D.C. English speaking respondents were recruited through the QDRL database, newspaper advertising, flyers and by word-of-mouth. Spanish speaking respondents were recruited through flyers, by word-of-mouth, and with the assistance of several non-profit organizations catering to the Latino community.

Table 5 presents respondent demographics for the study. An attempt was made to capture a broad range of respondents but particular emphasis was placed on recruiting gay and lesbian respondents as well as a range of those reporting ‘something else,’ specifically, those who identify as transgender, queer or who are still in the process of figuring out their sexuality.

Table 5. Respondent Demographics

Interviews Completed:	139	
	Count	Percentage
Gender		
Male	65	46.8%
Female	66	47.5%
More Complicated	8	5.8%
Sexual Identity		
Straight	86	61.9%
Gay or Lesbian	24	17.3%
Bisexual	9	6.5%
Something Else	19	13.7%
Education		
Less than HS degree	23	16.5%
High School Degree/GED	38	27.3%

Some college, no degree	22	15.8%
Associates Degree	17	12.2%
Bachelors	21	15.1%
Graduate School	17	12.2%
Race		
White	32	23.0%
Black	62	44.6%
Indian American	7	5.0%
Asian	4	2.9%
Other	18	12.9%
Latino	49	35.3%
Language		
English	94	67.6%
Spanish	45	32.4%
Age		
Under 25	21	15.1%
26-40	45	32.4%
41-60	48	34.5%
Over 60	16	11.5%

Interviewing Procedures

Respondents were scheduled for specific interview times (with the exception of a few “drop-ins”) and reported to a set location for their interview. Interviews lasted between 30 and 90 minutes with the typical interview lasting from 45-60 minutes. All interviews were audio recorded using both a cassette recorder as well as a sound recording program on the computer. Respondents were asked to check an anonymous consent form before the interview began and were also asked to give their oral consent once the taping began. At the conclusion of the interview all respondents were given \$50 as remuneration.

Unlike other QDRL interviewing projects, the questionnaire for this project was administered using an audio-computer assisted self-interview (ACASI) system. (Although not relevant to the findings of this report, the ACASI system was also being tested as one piece of this overall project). Respondents were asked to answer 8-10 demographic questions using the ACASI system and without any assistance from the interviewer. At the conclusion, respondents were asked each item and were then asked to explain their answer. Typical follow-up questions included, “How so?” and “Why do you say that?” If a respondent’s answer seemed vague or unclear, the interviewer asked: “Can you give an example to describe what you are talking about?” Specifically for the sexual identity question, respondents were also asked how they typically referred to themselves and were also asked about other words (i.e. ‘heterosexual’ and ‘homosexual’) that were not appearing in the question. The culminating text from the interview related how respondents

understood or interpreted each question and also outlined the types of experiences and behaviors respondents considered in providing an answer.

Data Analysis

Data from the interviews were analyzed using qualitative techniques, specifically, the constant comparative method (Lincoln and Guba, 1985; Strauss and Corbin, 1990; Creswell, 1998). Analysts used Q-Notes, an analysis software tool developed by the National Center for Health Statistics. As data were entered into the Q-Notes software, patterns of question interpretation and cognitive processing problems were identified. Some analyses, specifically assessment of question performance and identification of problems, were conducted simultaneously with interviews. This iterative process allowed for the question to be improved should any problems have arisen. By and large the question performed well, that is, respondents' answers were consistent with their identity, and only one time was a minor change made to the question. The Spanish translation, however, received a relatively significant alteration. (This will be discussed in the Findings section below).

After interviews were conducted, more intensive analyses were conducted so as to more systematically identify patterns of interpretation. The first step of data analysis involved reviewing the data and identifying the analytic themes as well as the thematic categories that made up each theme. For example, the theme 'respondent interpretation of heterosexual' was identified as an important analytic theme, and the categories underneath this theme reflected all of the different ways in which respondents conceptualized the term 'heterosexual.' Then, each interview was coded to reflect the particular interpretation. Categories were created as new interpretative patterns were discovered.

In order to specify the dimensionality of the themes and categories, respondents' narratives were compared, resolving any discrepancies and noting similarities. Additionally, the relationship of the themes and categories were examined. These core themes served as the unifying link between all patterns and denoted a working theory that depicts the phenomena captured by the survey questions.

FINDINGS

In comparison to previous versions of the sexual identity question (including the 2006 NSFG version), data from the cognitive interviews indicate that this newly developed version is a noticeable improvement. In all but 10 of the 139 interviews, respondents selected the response category that best reflected their sexual identity. That is, respondents' answers were based on the ways in which they conceptualize their own sexuality. (This will be fully discussed below). This was true for all age and socio-economic groups. Notably, almost all heterosexual respondents opted for the 'straight, that is, not gay,' response option with no difficulty.

The presence of the ‘something else’ category along with the follow-up question also proved to be a successful revision. All respondents who opted for this category were able to effectively classify themselves within one of the provided options. Unlike previous versions of the question, none of these respondents were heterosexual; non-minority respondents answered by selecting the ‘straight, that is, not gay’ category. Thus, we believe that the revision of the heterosexual category resolves the missing data problem, including heterosexuals choosing the ‘something else’ category. It should also be noted that only one Spanish-speaking respondent selected the ‘don’t know option’ because she was not familiar with the terminology. The ‘something else’ option was most frequently chosen by transgender respondents, who then selected the transgender option in the follow-up question. Other respondents who selected the ‘something else’ option included those respondents who identify as queer, do not use labels to identify themselves, have not figured out their sexuality or do not consider themselves to have a sexuality.

Question Interpretation

As was found in previous QDRL studies, many non-LGBT respondents did not possess salient sexual identities². Additionally, as in previous studies, for these respondents it was not so much an association with a particular sexual identity that mattered as it was a disassociation from a gay identity. When asked about their sexual identity, many respondents simply said that they are “not gay.” During probing a number of respondents indicated that they chose this option specifically because it said “not gay” and that this is what made the question easy for them to answer. One respondent, for example, felt that it was insulting to gay people to call oneself straight – “you’re just not gay” she noted. Another respondent was asked if she would use the word ‘straight’ to describe herself in her everyday interactions. After pausing for a moment the respondent answered, “I would just say that I am not involved in a gay relationship.” She went on to say “I don’t know why they use that word....cuz really to me the word is ‘not gay.’ I don’t know why people define it as straight and gay.” Another respondent said that he was confused by the category ‘straight,’ but when saw ‘that is, not gay,’ he knew immediately which category applied to him. Another respondent who identified as straight said that to her this meant that she “doesn’t mess around or do things out of the ordinary.” Another respondent said plainly that to be straight means she “don’t act like they do.”

Almost all sexual minorities answered this question based on their conception of self, that is, how they identify themselves. As found in previous studies, these respondents consider their sexual identity to be a central component of their sense of self. Respondents based their conceptualization on a number of factors – membership in a larger community, political activism, various personality characteristics, and

² It is important to note that this question was asked within the context of other demographic questions. It is known from past research (Miller, Ridolfo, and Maitland, forthcoming) that the context of a sexual identity question may impact the way respondents interpret the question.

relationship status. What is true of all of these factors, however, is that they are all various mechanisms through which respondents make sense of their sexual identity.

One way in which respondents framed their sexual identity was through membership in a larger community. Many of the respondents saw themselves as members of a larger sociopolitical group, and they conceptualized their identity based on an affiliation with a larger LGBT community. One respondent, for example, said that they define gay simply as “the whole community.” Another mechanism by which respondents informed their sense of self was through what they perceived to be political activism. There was a clear theme among many of the minority-identified respondents that their sexual identity was strongly tied to a sense of political activism. In a culture where homosexuality has been and continues to be heavily politicized, this sort of activist affiliation is seen to be a logical base for identity development. One respondent, for example, who identified as ‘something else’ said that they do not really like to use labels but that they feel that they should do so in order to educate people.

Several of the respondents viewed certain personality traits as expressions of their (and others’) identity as a sexual minority. One male respondent who identifies as gay, for example, made sense of his identity based on his perception of characteristics he finds to be inherent to gay people. He said that to be gay means to be happy living a certain lifestyle which involves “being free, ecstatic, dramatic, full of zest and flavor.” He went on to mention all of the artistic gifts that gay people have been given. He further noted that it had nothing to do with sex as he has not had sexual relations in five years yet he still identifies as gay. Another respondent said that to be gay means that he can’t think like a straight person. He said that straight people are more closed minded and focused on machismo while gay people are more open minded.

The sex of one’s relationship partner was another mechanism by which some respondents’ made sense of their sexual identity. One female respondent, who identifies as ‘something else,’ for example, is currently in a relationship with a man but has been in relationships with women before. She said at the time of her relationship with a self-identified lesbian, she identified herself as “Maria-sexual,” based on the name of her partner. She makes sense of her identity not based on behavior or attraction but rather based on the relationship that she is in at the time.

Transgender people often have a difficult time fitting into either the heterosexual or the LGB community, although they feel a greater affinity for the latter. For this reason, many trans respondents referred to the gay community in broader, more encompassing terms than LGB or heterosexual respondents. Thus, a number of transgender respondents conceived of the term “gay” as both an individual identity as well as an umbrella term for a larger community of sexual minorities (the exact composition of that community varied among respondents). One transgender respondent said that although gay can specifically refer to a man who is masculine it can also be used to refer to “the whole community.” Another transgender respondent wanted to choose the term transgender but since it was not available chose gay because she felt that this was the closest option for her since it would include her in the LGBT

community. Another transgender respondent said that she thinks of the term gay as being in the middle of a big circle of other terms like bisexual and transsexual and that 'gay' is the word used to describe all of these things. She said that 'gay' is the generic word used to describe all of these other terms, but that it is not specific enough and she would not identify this way. Instead, she identifies specifically as transsexual.

Cases of Response Problems

Of the 10 respondents who did not answer according to their sexual identity (and which could be considered error) 3 were sexual minority respondents and the other 7 were Spanish-speaking respondents. Of the sexual minority respondents, 2 interpreted the question as a behavior question as opposed to an identity question and, consequently, answered bisexual. One woman, for example, who identifies as 'queer,' answered bisexual because she surmised that a CDC survey must be asking about her behavior, not her self-conceptualization. The other respondent not basing his answer on identity was transgender and answered according to the clinical records where his gender transitioning occurred, which was bisexual. While these cases do represent what would be considered error, it was deemed imprudent to make a revision to the question because any 'fix' would likely generate other types of error. Rather, these authors believe it may be more prudent to embed this question among other demographic or self-identification questions as opposed to other behavioral questions. Such a context may cue respondents to base the answer on their self-identification as opposed to their behavioral history.

The 7 Spanish-speaking respondents who answered incorrectly were respondents who did not understand the word 'gay,' but were more familiar with the term 'heterosexual.' Since the word 'gay' (along with the term 'straight') is also an English-derived term, some of the Latino respondents were unable to make sense of the phrase 'no es gay.' For these respondents, absence of the term 'heterosexual' generated more (as opposed to less in comparison to their English-speaking counterparts) response problems. For example, one Latino who answered something else later revealed that he is heterosexual but that he did not see that option listed for this question. Similarly, a Latina respondent answered bisexual, but during probing revealed that, because she had to think very quickly and did not see the option for heterosexual, chose bisexual.

Additionally for Spanish-speaking respondents, because the word 'heterosexual' was not listed, other terms, specifically 'bisexual' and 'lesbiana o gay,' were misinterpreted. For example, one Latina who answered 'bisexual' explained during probing that bisexuals are those who only sleep with men. Realizing her mistake she said "oh no! Bisexual means that they have sex with both men and women. I'm heterosexual!" She went on to say that the response categories did not include the option she was looking for – heterosexual. Another Spanish-speaking respondent answered 'lesbian or gay' because he was not sure what the word is for men who only like women. He couldn't remember if it was bisexual or heterosexual so he just chose the first response category listed. To resolve this response problem, the Spanish translation was modified shortly after these Spanish interviews, and that it is believed

that this modification will minimize, if not eliminate, these instances. It should be noted that none of the Spanish speaking respondents had difficulty selecting the response category that best reflected their sexual identity after the word 'heterosexual' was added.

Interpretation of the Term 'Heterosexual'

Perhaps most controversial about this revised question in comparison to previous questions about sexual identity is the non-appearance of the term 'heterosexual' as a response option. For English interviews, we found no evidence to suggest the presence of response error or any response difficulty because the word 'heterosexual' was not listed. This was true for all English-speaking demographic groups across heterosexuals. Even those respondents who said that they used the word 'heterosexual' to self-identify were also familiar with the word 'straight' or related to the concept 'not gay.' In no case did an English speaking respondent indicate that they did not know how to answer because the word 'heterosexual' was not there.

Consistent with previous studies, in the follow-up probing, it was found that many lower socio-economic non-minority respondents either did not know or misunderstood the term 'heterosexual.' For example, when asked what heterosexual meant one English-speaking respondent said, "Who?," and then asked the interviewer what that word meant and how to pronounce it. Another female respondent noted that she was familiar with the term heterosexual but wasn't entirely sure what it meant. A number of respondents confused the term 'heterosexual' with being homosexual and with being bisexual. For example, when asked what heterosexual meant one respondent answered that "it means men who like men." One female respondent explained that heterosexual means you can go with both men and women. Another respondent said that it is "somebody who goes both ways." Yet another respondent pointedly replied that "heterosexual means the same thing as bisexual."

Indeed, many of those who knew the definition of heterosexual remained unsure. When asked why he chose the answer he did, one respondent said it was because he identifies as "heterosexual or as someone who only likes women, unless I'm wrong about the definition of heterosexual." Another respondent who was also unsure said that they were fairly confident it meant the same thing as straight but they weren't totally sure about that. This last respondent emphasized the point that even for those respondents who might know the term 'heterosexual,' the use of more common language is a more guaranteed way to ensure respondent comprehension of response options.

Although sexual minorities tended to be more familiar with sexual identity-related terms, there were instances, particularly related to the word 'heterosexual,' when they were not. One lesbian, for example, said that the word she would use for someone who likes the opposite sex is straight. When probed whether there was another word for this she said "I think the word is heterosexual, but maybe it's homosexual." She said that

either way it didn't matter to her because these words are basically for people who "deal with" the opposite sex.

For English-speakers, even among those who knew the term 'heterosexual,' there was still a clear preference for the word 'straight.' Several respondents noted that the term 'heterosexual' (and on occasion, but not always, the term 'homosexual') is a very scientific term and not what they use in everyday language. One respondent noted that he thought he had heard the term 'heterosexual' in science class. An English speaking male responded that he uses the word 'straight' to describe himself normally and only uses the term heterosexual at school and when asked directly if he is a heterosexual or not. Most importantly, even among those who do use the word 'heterosexual,' straight was also understood.

In sum, the reasons for omitting the word 'heterosexual' in a response option of the English version of the question are three-fold: 1) it is not the word that most people use in their everyday speech, 2) it is not required, as people understand the word 'straight,' and 3) many people are confused, do not understand, or misunderstand the word 'heterosexual.' The word 'straight,' although considered by some respondents to be slang, was understood by all English speaking respondents and, equally as important, understood to mean what is implied by a heterosexual identity. The usage of the word 'straight' and the removal of the word 'heterosexual' in combination with the phrase 'not gay,' therefore, were found to greatly reduce conceptual confusion among respondents.

As previously noted, the above findings did not hold true for Spanish speaking respondents. Because there is no word for 'straight' in Spanish (although many Spanish speakers who had been living in the United States for a while were familiar with this term), the option simply read 'no es gay.' 'No es gay' was not clear because the term 'gay' is also an English term that is not always understood (as these few cases illustrate). For Spanish speakers, the term 'heterosexual' was found to be much more commonly used and understood. That is, as far as usage and familiarity, the term 'heterosexual' in Spanish is comparable to the term 'straight' in English. For example, two respondents noted that they would have chosen 'heterosexual' had they seen this option but since they did not see it instead chose 'something else' and 'bisexual.'

Even among Spanish speaking heterosexuals who did not have problems selecting the response option that best reflected their sexual identity, there was a strong sentiment that the presence of the word 'heterosexual' would have made the question easier to answer. One respondent, for example, when asked how she understood 'no lesbiana o gay' said "this is maybe where heterosexual goes?" Another respondent when asked what other words he would use to describe 'no es gay' said that heterosexual was the most common word used for this. Yet another said that she found 'no es gay' to be confusing and instead would have chosen the words 'heterosexual' or 'straight.' Our data suggests that, although not always used, the term 'heterosexual' is more commonly used among Spanish speakers. In response to this finding, the response option was

changed from 'no es gay' to 'heterosexual, o sea, no es gay.' This was then tested on 18 respondents, none of which had error or response difficulty.

Interpretation of the Terms 'Gay,' 'Lesbian,' and 'Homosexual'

As revealed in the comparison of the 2002 and 2006 NSFG data, the addition of the term 'gay' appears to increase conceptual clarity among respondents. Because this is the word used most commonly by both sexual minorities and non-sexual minorities, it was scarcely unknown or misunderstood. The term 'lesbian' was also commonly understood by respondents with no cases of conceptual confusion among either English or Spanish speaking respondents. The term was generally understood to mean the same thing as gay with the exception of one respondent who reported that she uses the word lesbian to refer to herself but does not use the word 'gay.' For example, when a respondent who reported that she uses both the word 'gay' as well as 'lesbian' to describe herself was asked which she preferred, she responded "I would choose lesbian, but it's still the same." Alternatively, another female respondent said that she uses both gay and lesbian to refer to herself but that she has a slight preference for gay. Another respondent said that she defines herself as a lesbian but that the term gay would also apply to her since it is a broader term encompassing "both men and women who like the same gender." Thus, although there was variation in preference for the term gay or the term lesbian, there was no conceptual confusion created by the term 'lesbian.'

We found compelling evidence not to use the term 'homosexual' in the response options. Like the term 'heterosexual,' 'homosexual' was often misunderstood or not known by respondents. One English speaking respondent, for example, knew the term gay but not the terms heterosexual or homosexual. Another female respondent explained that, to her, being homosexual means being attracted to the opposite sex. In addition, and like the term 'heterosexual,' even when the term 'homosexual' was understood it was often seen as an overly clinical term or, unlike the term 'heterosexual,' seen in a pejorative light. One Spanish speaking respondent noted that to refer to gays she uses the term "chicos gays" because the word 'homosexual' is "stronger" and has a negative connotation. An English speaking respondent said that he only hears the word 'homosexual' used when speaking disparagingly of people, for example with reference to a "homosexual agenda." Another gay male acknowledged that homosexual does not have any inherently bad meaning but that people "don't use it properly.... and they say it with disdain."

Some respondents acknowledged that they might use the word 'homosexual' but only in certain circumstances. For example, one gay male respondent said that he might identify as homosexual to a foreigner "who might not know what gay means." Another said he uses the word 'homosexual' "only in the context of jokes." Another context for the usage of the word homosexual seems to be generational. For example, one 88 year old respondent said that her grandchildren always correct her when she uses the word homosexual and tell her that the word is just gay. This latter point illustrates the larger

point that even when homosexual is the preferred word choice, respondents are still familiar with the term gay.

There were a small number of Spanish speaking respondents who noted that the term ‘gay’ might not always be understood by other Spanish speaking respondents because it is an English word. One respondent, for example, said that he uses the term ‘gay’ with his friends in the States but ‘homosexual’ with his friends back in El Salvador. This would be consistent with our finding that those who had lived in the United States longer were also more likely to understand the term ‘straight,’ another English language slang (although this term was not used on the Spanish version, it was still mentioned by several respondents). One reason for this is that each country has its own specific slang for gay people, most of which are fairly insulting. One respondent, for example, said that the lower class in his country use the terms “maricones” or “culeros” and only the upper class really uses the term gay. Although this potential source of error should be noted, in none of our 45 Spanish language interviews did we encounter a respondent who was unable to select the sexual identity that best represented them because of the presence of the term ‘gay’ (or the absence of the term ‘homosexual’).

While a few confused the term gay with being heterosexual or bisexual, a fair number of non-minority respondents believed that the term ‘gay’ meant taking on some characteristic of being transgender, that is acting, dressing, or taking on the characteristics of the opposite gender. One respondent, for example, talked about gays as “men who wear ladies clothes.” This was echoed by another respondent who said that gay men dress like females and wear bras and skirts. Another respondent defined being a lesbian as someone with the body of a woman but the attitudes of a man. Yet another respondent answered that “a gay” is a man who dresses like women and likes men while another said a gay person is someone “trying” to be male or female, especially men who “try to play a female role.” This misunderstanding was also found among Spanish speaking respondents. One heterosexual Spanish speaking respondent, for example, said that gay means when a man wants to be a woman or to act like a woman. Another Spanish speaking respondent said that gay men are biologically men but want to be women and are not well defined in their identity. Again, it is difficult to determine if these are actually instances of conceptual confusion or, more likely, misunderstanding due to homophobia. Either way, they did not appear to impact respondents’ ability to properly select the response option that best reflected their sexual identity.

To a certain extent, some heterosexual respondents—particularly Spanish-speaking respondents—conflated being gay with a dimension of gender identity. For example, one respondent said “if you’re male, you’re straight. If you’re female, you’re straight.” Another noted that “I’m normal. I’m a woman. I’m feminine” thus expressing not only a confusion with gender identity but the reference to “normal” also implies an association with a “not-me” identity. A number of other respondents answered this question by simply saying “soy hombre” (I am a man) or “soy mujer” (I am a woman). The underlying theme of these respondents can be summed up by one who said that heterosexuals “don’t feel like a man one day and a woman the next.”

Interpretation of the Term ‘Bisexual’

Although there was some confusion over the meaning of the term ‘bisexual’, it did not lead to response error problems because those people that did not know what the term meant did know the category with which they identified – i.e. gay or not gay – and so knew for sure that bisexual was not for them. In the English speaking cases where bisexual was chosen as response error, it was done not because of confusion over the term, but rather because these respondents thought the question was asking about behavior rather than identity (as discussed above).

That said, there were some respondents who did not know what the term ‘bisexual’ means. One respondent, for example, said she had heard of the term but added “I don’t quite understand what it means.” Even many respondents who knew the meaning of the word ‘bisexual’ still had definitions typically rooted in being gay or heterosexual. For example, one respondent said that being bisexual meant being “heterosexual and attracted to the same sex.” This respondent started with an understanding of heterosexual and then built from it.

Other respondents confused the meaning of bisexual with either gay, heterosexual, or transgender. One respondent, for example, said that it was just another term for gay – “sounds like the same thing to me,” he said. Another verified that it meant the same to her as heterosexual. One male respondent who identifies as gay but is married to a woman said that the word bisexual is just a “cover word” for people who think the word gay means something bad. Another respondent said that bisexual is either someone who watches a couple have sex or a woman that enjoys sex with men and women but they are not certain which one. Another respondent said that a bisexual person tries to be “a woman and a man at the same time.”

Even among those who understood the general concept of bisexuality, there was still sometimes confusion over its precise meaning. One transgender respondent, for example, revealed that although he has sex with both men and women, he does not consider himself bisexual because he thinks bisexual means that half the time you are attracted to men and half the time you are attracted to women whereas he is attracted to women 80% of the time and men only 20% of the time.

Some respondents knew the concept of bisexuality, but not the word. One female heterosexual, for example, said that you can like a man and a woman at the same time but she was not sure what the word to describe this would be. An elderly female respondent seemed to understand the concept but not be familiar with the word. She said that someone is either gay or not gay (reinforcing our earlier point of the not-me identity) but that someone might be somewhere in the middle. She assumed, however, that this person would then select ‘don’t know.’

Confusion over the term bisexuality was also found among Spanish speaking respondents. One such respondent said that bisexual meant someone who likes women but also “likes gays.” Another Spanish respondent said that bisexuals have a personal

conflict on how to define themselves. Another, unable to clearly articulate a definition, could only say that a bisexual is “someone who is a human being.” This is further evidence that sexuality is conceived of differently among Spanish speakers.

Behavior seemed to be much more prevalent in respondent’s conception of bisexuality. One heterosexual female, for example, said that bisexual means people who sleep with their own sex and the opposite sex. She said that unlike being gay, being bisexual necessarily involves sex. Another common response was that bisexual implies “going both ways” with follow up references to sexual activity with both males and females. Along these lines one lesbian respondent said that bisexual means “when you don’t know which sex you want to be with and you just take them both.”

Interpretation of the Response Category ‘Something Else’

The response option for ‘something else’ was well understood by those who identified as something else. Many transgender respondents, for example, selected something else on the basis of their transgender identity. Several of the trans respondents noted that the first thing they looked for was a ‘transgender’ response option³ but when they did not find this option, these respondents then chose ‘something else’ assuming that that is what it meant. There were also respondents who identify as queer, do not use labels to identify themselves, and are asexual – all sub-options of the ‘something else’ response category - who were also able to accurately select this category as the one that best reflected their sexual identity.

Even many of the non-transgender respondents felt that ‘something else’ implied some variation of an understanding of transgender. One respondent, for example, said that something else is for those people who don’t know what they want to be – male or female – and that they have not found their sexuality yet. Another respondent felt that maybe it was for people who didn’t want to openly identify as gay or who were transgender or “lost” and don’t really know what they are. Others noted that it was a category for people who are not a lesbian or a homosexual. A gay male respondent said that “there are so many letters now” and so it gives people a chance to pick something different.

Some respondents, especially those who did not identify as ‘something else’ had varying initial conceptions of what the ‘something else’ category could possibly mean or simply had no idea what it might imply. A heterosexual female, for example, said that something else made no sense to her because either you are straight or you are not. Another heterosexual respondent thought that “maybe they like dogs.” Another female respondent said that something else could be a hermaphrodite. She said that she knew a couple of hermaphrodites and that these are people born “with both sexes, both organs,” and then their parents decide if they want to raise them as a boy or a girl. Another

³ This was certainly not the case for all trans respondents as some chose “gay or lesbian” or “straight, that is not gay” without debate.

respondent said it was for someone who doesn't know if they like men or women and is the same as the 'don't know' option.

CONCLUSIONS

Overall, our analysis of the 139 cognitive interviews leads to at least four main conclusions:

- The absence of the word 'heterosexual' on the English language question is helpful to reduce response difficulty. It is important to use common vernacular in order to reduce conceptual confusion. Thus, while the absence of the term 'heterosexual' did not lead to any confusion among respondents in any demographic, its presence did.
- The presence of the word 'heterosexual' on the Spanish language question helps respondents make sense of other response categories. Since there is no conceptual translation for the word 'straight' in Spanish the presence of 'heterosexual,' a word more commonly used by Spanish speakers than English ones, is useful to provide context not only for this option but for the others as well.
- For many heterosexuals the concept of sexual identity is not salient. They do not so much identify with being heterosexual as they dis-identify with being gay. To this end, the addition of 'that is, not gay' was useful in helping these respondents select the optimal response category.
- Due to the presence of the 'not gay' wording, it is necessary to put this response category lower than the 'gay' category. This is not only logically more correct, it also encourages respondents to more deeply consider previous response options.

REFERENCES

- Cast, A.D. (2003). "Identities and behavior," In P.J. Burke, T.J. Owens, R.T. Sherpe, & P.A. Thoits (Eds.), *Advances in Identity Theory and Research*. New York: Kluwer Academic/Plenum Publishers: pp. 41-53.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design; Choosing Among Five Traditions*. London, New Delhi, Thousand Oaks, Sage Publications.
- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51: pp. 201 - 219.
- Lincoln, Y. S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications, Inc.
- McCall, G.J. (2003). "The me and the not-me: Positive and negative poles of identity," In P.J. Burke, T.J. Owens, R.T. Sherpe, & P.A. Thoits (Eds.), *Advances in Identity Theory and Research*. New York: Kluwer Academic/Plenum Publishers: pp. 11-25.
- Miller, K. (2011). "Cognitive Interviewing," In J. Madans, K. Miller, A. Maitland, and G. Willis (Eds.), *Question Evaluation Methods: Contributing to the Science of Data Quality*. Hoboken, NJ: Wiley Publishers.
- Plummer, K. (1981). "Homosexual categories: Some research problems in the labeling perspective of homosexuality," In K. Plummer (Ed.), *The Making of the Modern Homosexual* Totowa, NJ: Barnes and Noble Books: pp. 53-75.
- Plummer, K. (1995). *Telling sexual stories: Power, change, and social worlds*. New York: Routledge.
- Q-Bank. <http://wwwn.cdc.gov/QBANK/Home.aspx>
- Ridolfo, H., Miller, K. Maitland, A. (Forthcoming). *Sexuality Research and Social Policy*.
- Rust, P.C. (1993). 'Coming out' in the age of social constructionism: Sexual identity formation among lesbian and bisexual women. *Gender & Society*, 7(1): pp. 50-77.
- Strauss, A., and Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications, Inc.
- Tourangeau, R., Rips, L.J., and Rasinski, K. 2000. *The Psychology of Survey Response*. Cambridge University Press.