Best Practice for Demographic Data Collection & Reporting

Evaluator's Guide

By Kimberly Magoon,
Mary-Jo Robinson,
Alexandra Kissling,
Victory Ozeua

August 2022

Introduction

This guide is a reference for best practice in collecting and reporting of race, ethnicity, as well as sexual orientation, gender identity and expression (SOGIE)[1] in evaluation work. Sources chosen for this review include national leaders for specific demographics as well as those that represent topic areas in which a large portion of public health and human services work are conducted.

While there are other sensitive demographics that may be considered in this work, such as disability/ability status, marital status, education level, and income, this document focuses on the four most commonly collected/ reported variables at this time. Future document revisions may include additional categories of data.

The primary goal of this document is to move evaluators toward more culturally and linguistically appropriate standards (CLAS) for data collection and reporting via establishing guidelines for the work we do for our clients and the people we receive information from. While standardization of demographic data collection and reporting is helpful in comparing information across settings, minimizing reporting burden, and optimizing data sharing, this is not the purpose of this document.

It is acknowledged that not only does best practice evolve and change over time, but people may have unique and distinct preferences for how they would like to be identified. Therefore, best practices cited in this document should be considered a minimum standard or default for data collection and reporting. It is paramount that, whenever possible, data collection or reporting begin with conversations with clients, stakeholders, and people whose demographic characteristics will be described to outline evaluation and reporting parameters accurately and respectfully.

Guide Structure

This document begins with an overview of health equity and inclusive language. Then, synthesizing information from nearly 40 agencies and organizations, the document offers quick reference examples for inclusive nomenclature, question structure, and reporting of race, ethnicity, gender identity, and sexual orientation.

Health Equity & Inclusive Language

Many organizations have begun to formally recognize and emphasize the importance of addressing inclusivity for all people. The U.S. Centers for Disease Control and Prevention (CDC) suggests the use of a health equity lens for ensuring non-stigmatizing, bias-free language.[2]

A health equity lens includes four main components:

Systemic, social inequities have long-stigmatized groups of people and have linked them to negative circumstances or outcomes. We, as consultants and evaluators, have a responsibility to take precaution to **avoid implying that a person, community, or population is responsible for increased risks of adverse outcomes.** Therefore,

health disparities should be tied to and conceptualized by social determinants of health (SDOH). The Office of Disease Prevention and Health Promotion (ODPHP) outlines five categories for SDOH: economic stability, education, neighborhood and build environment, health and health care, and social and community context. See their report on HealthyPeople.gov for more information.[3]

✓ **Language should describe people first**—not as a condition or circumstance in which the person happens to have experienced. People are always people first, described as an identifier second. For example, in terms of sexual orientation, we should refer to people who are heterosexual, gay, lesbian, bisexual, etc. rather than gay people. Further, in the case of substance use, evaluators should avoid terminology such as drug users or abusers and instead refer to people who use substances or people who experience substance use disorder, whichever is more appropriate.

✓ Because people may belong to or identify with several groups at the same time (e.g., race, ethnicity, SOGIE), **health equity in data collection and reporting must account for intersectionality** of groups. For example, someone who identifies as Black, female, and heterosexual will likely have a different experience than someone who identifies as Black, male, and gay. A singular identification of race does not guarantee a homogenous experience, and data collection/reporting should contextualize and account for this.

✓ **Incorporation of communities who are being engaged and described is paramount** for inclusive data collection and reporting. While it may not always be possible to include community members in every project, best practice dictates that attempts are made, when possible, throughout the life of the project from planning to implementation and reporting. Engagement can strengthen partnerships, ensure cultural and linguistically appropriate practices, and promote health equity.

Data Collection Best Practice

Though there have been attempts to standardize data collection categorizations nationally, one of the most notable resources is "The Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity" in 1997 by the Office of Management and Budget under the president and the subsequent update in 2016.[4] Although much has been learned, even in the last six years, regarding inclusivity and health equity, there are four basic principles of inclusive data collection that continue to remain true.[5]

1. **Every question should have a purpose.** Consider what you want to do with the data and how it will inform the project's objectives and goals. While it may be interesting to know something more about a population, sometimes it isn't necessary for the analysis and findings. Data should only be collected for variables that will be useful in analysis and reporting.

2. **Self-identification is always the preferred method of obtaining demographic information.** As will become clear in the section describing categorical best practices, sources vary slightly in their preferences for collecting information around gender identity, sexual orientation, race, and ethnicity. Not only is it preferred that people respond to questions about themselves, but it is also important that populations who will be responding to those questions are involved in the data collection tool development whenever possible. It is also best practice to compensate them for their time and effort.

3. **Information for demographic categories should be collected separately.** Most leading authorities agree that questions about race and ethnicity should be asked separately as should questions about a person's gender identity and sexual orientation. This delineation of terms allows for a more accurate, comprehensive picture of a population and validates, for example, that someone can identify with a culture or ethnicity and as a member of one or more races in the same way they can identify as a specific gender that does not dictate to whom they are sexually attracted.

4. **Evaluators should consider the need for granularity.** While collecting nuanced information is important for accurately understanding impact and contextualizing results, the depth of granularity in reporting should always be supported by sufficient sample size to promote confidentiality and prevent identification/ misuse of the knowledge. Therefore, information on the validity and reliability of the data should be included, whenever possible.

Understanding these principles and viewing data collection of specific demographics through this lens, the following sections summarize literature review findings for race, ethnicity, gender identity, and sexual orientation and offer examples of inclusive questions for use in surveys, interviews, and other data collection tools. For a detailed description of the methodology used to develop this guide, please see Appendix A.

Race and Ethnicity

This section synthesizes findings from the literature review into a recommended best practice for data collection related to race and ethnicity. Race has been traditionally defined as a set of genetic or physical characteristics ascribed to a group of people, but scientific inquiries have indicated that there is little evidence to support this. The National Institute of Health (NIH) and other authorities report that there are often more genetic variations among racial groups than between them. Therefore, race, like ethnicity, is much more of a social construct than a biological one.[67] For our purpose, we will use both terms to describe a shared identity based on common national, geographic, tribal, religious, linguistic, and cultural beliefs and customs.[8]

Defining race and ethnicity as social constructs reinforces the previously outlined self-identification best practice for demographic data collection. It is critically important for accuracy that demographics are self- defined characteristics—not physically identifiable things that can be merely observed. This also means that individuals can and often do belong to more than one racial or ethnic group. Therefore, sources all agree that data collection should support the ability of people to describe themselves as more than one racial or ethnic category.

As definitions have evolved, so have the best practices for data collection. Some leading agencies for race and ethnicity, such as the National Association for the Advancement of Colored People (NAACP), do not offer very nuanced descriptors, using only Black, African American, White, and Hispanic terminology.[9] However, the U.S. Census Bureau has collected information on race and ethnicity since 1790.[10] Over the years, descriptions for race have become more nuanced to reflect the diversity of the U.S. population.[11] Currently, census data classifies individuals into the following categories: White, American Indian/Alaska Native, Asian, African American, Native Hawaiian/Pacific Islander, and Some Other Race. For ethnicity, census data classifies individuals in as "Hispanic or Latino" or "Not Hispanic or Latino."[12] The National Prevention Network (NPN) and American Evaluation Association (AEA) prescribe similar descriptors.

Moreover, the descriptors for race and ethnicity used by the U.S Census Bureau constitute the primary data adopted by many organizations for their categorization. For example, the U.S. CDC, Office of Policy Development and Research, and the U.S. Department of Health and Human Services (HHS), Office of Minority Health use the U.S. Census Bureau data to enhance existing resources and support their classification of race and ethnicity.

However, the American Civil Liberties Union (ACLU) and the Federal Drug Administration (FDA) suggest yet more nuanced categorization. For example, the ACLU recommends classifications for Middle East andNorth Africa regions, Asian American, Black, and African populations.[13] For Asian American populations, they suggest checkboxes for Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, and "other Asian" categories. The ACLU also suggests using separate checkboxes for Native Hawaiian and Pacific Islanders. For Black and African populations, opportunities should also be provided to self-identify. More specifically for Black immigrant populations, individuals should have the ability to indicate more than one national origin or subgroup. Further, they recommend that people who identify as American

Indian, Native American, or Alaskan Native should be allowed to further self-describe as identities are often specifically tied to a tribal nation.

The FDA also recommends allowing individuals to separately and more distinctly define ethnicity as non- Hispanic, Latino/a, or Spanish origin or as a specific type of Hispanic origin (e.g., Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or "other" self-described category).

Recommended Questions for Race and Ethnicity

Based on review of institutional descriptors, a modified version of the U.S Census Bureau's demographic classification is recommended as best practice, utilizing the two-part question design recommended by the FDA.

Are you of Hispanic, Latino/a/x, or Spanish origin? (check one or more boxes)

- ☐ No, not of Hispanic, Latino/a/x, or Spanish Origin [Exclusive choice]
- ☐ Yes, Mexican, Mexican American, or Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, another Hispanic, Latino or Spanish origin (e.g., Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian) [Open text box]
- ☐ Prefer not to answer

What is your race? (check one or more boxes **AND** specify origin)

- ☐ White (e.g., German, Irish, English, Italian) [Open text box]
- ☐ Black or African (e.g., African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali) [Open text box]
- ☐ American Indian/Alaska Native (e.g., Navajo Nation, Blackfeet tribe, Mayan, Aztec) [Open text box]
- ☐ Middle East / North Africa (MENA) region (e.g., Lebanese, Egyptian, Syrian, Moroccan) [Open text box]
- ☐ Asian American
- ☐ Chinese
- ☐ Filipino
- ☐ Asian Indian
- ☐ Vietnamese
- ☐ Korean
- ☐ Japanese
- ☐ Other Asian (e.g., Pakistani, Cambodian, Hmong) [Open text box]
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Chamorro
- ☐ Other Pacific Islander (e.g., Tongan, Fijian, Marshallese, Palauan, Tahitian, Chuukese) [Open text box]
- ☐ Some other race [Open text box]
- ☐ Prefer not to answer

Recommended Reporting for Race and Ethnicity

As we have seen, there are numerous terms used to describe race. However, we also know it is not always ethically responsible to report data in such finite categories, especially when sample sizes are limited, and people could be specifically identified. Therefore, it may make sense to talk about people in more general ways or in larger groups.

Below are a few common examples:

- People of Color (POC) generally refers to people who identify as a race other than White.

- BIPOC specifically refers to people who identify as Black, Indigenous, People of Color and BIWOC refers to people who identify as Black, Indigenous, Women of Color.

- Indigenous communities, peoples, and nations are defined as those that have "historical continuity with pre-invasion and pre-colonial societies that developed on their territories, consider themselves distinct from other sectors of the societies now prevailing on those territories, or parts of them. They form at present non-dominant sectors of society and are determined to preserve, develop and transmit to future generations their ancestral territories, and their ethnic identity, as the basis of their continued existence as peoples, in accordance with their own cultural patterns, social institutions and legal system."[14]

- American Indian/Alaskan Native (AIAN) are types of indigenous communities within North America.

- MENA is used to describe people from the Middle East and North Africa, such as Lebanon, Egypt, Syria, and Morocco.

Regardless of terminology, it is best to describe people as they would like to describe themselves and include them in the process of evaluation and reporting.

Gender Identity

The following section outlines recommended practices for the use of demographic questions to describe gender identity. Though the terms are often used interchangeably, sex refers to a person's biological composition and anatomy, whereas gender refers to a person's broader social and cultural identify which may or may not align with biological markers.[15] Data analysis typically focuses on sex, but in most cases, evaluators are interested in gender, not sex. Gender encompasses identity (a core element of a person's individual sense of self), expression (how an individual signals their gender to others through behavior and appearance), and social and cultural expectations (related to social status, characteristics, and behavior that are associated with sex traits).[16]

The increasing visibility of people who identify as transgender (people whose current gender identity is different from the sex they were assigned at birth) or intersex (people whose sex traits do not all correspond to a single sex) requires asking gender identity questions in ways that go beyond the binary categorization of male and female. Researchers using the General Social Survey, one of the longest running and influential U.S. national surveys, have found that interviewer-reported sex, self-reported sex, and self-reported gender identity may not match within the same respondent.[17] A growing desire within the scientific community to measure sex and gender more accurately has contributed to calls for standards on how sex and gender questions should be asked. More accurate and specific measurement of gender identity is essential for identification of at-risk populations and understanding the challenges they face.

When determining which gender or sex-related demographic question(s) to include in survey research, the National Academy of Sciences recommends considering the purpose behind asking this question. If the research questions of interest are more related to social relationships, then use gender questions. Similarly, the American Evaluation Association suggests asking four questions: What do you want to assess (e.g., sexual orientation, gender constructs)? What is your sample (e.g., general population or more specific)? What are your measurement options (e.g., open ended branches, prompt definitions)? What can you do with it (capacity, expertise, use)?[18]

Leading gender researchers find that using a two-part question (such as asking for both current gender identity and sex at birth) avoids misinterpretation and misclassification.[16-18] There are many benefits to using a multi-part gender question. First, there is a distinction between sex assigned at birth and current gender identity, and a multi-part question captures respondents who identify as transgender with a broad definition. Some surveys, such as Youth Risk Behavior Surveillance System (YRBSS), have used a single step standalone question that asks, "Are you transgender?"[19] A multi-step question improves on this model with the ability to cross-tabulate two items, producing counts of people who identify as cisgender men or transgender women, for example.

After reviewing over 20 measures of gender from top national surveys such as National Longitudinal Study of Adolescent and Adult Health, Behavioral Risk Factor Surveillance System, and the National Health Interview Study, the National Academy of Sciences state that the questions comprising their multi-step gender identity measure should always be paired together.[20] Measuring sex assigned at birth alone does not take into account multidimensional anatomical, genetic, and physiological sex traits and the independent role of gender in shaping people's health and life experiences. Sex and gender identity produce combined biological, environmental, and contextual experiences. Social structures influence behaviors that both create and magnify differences that may appear to be rooted in biology.

While useful, there are some challenges to be noted with using a two-step gender question written in this way. Though the question is "forced choice," the response options are not necessarily mutually exclusive. Additionally, the write-in gender identity field will have to be cleaned and coded for reporting. Newer terms, such as gender-fluid (does not identify with a fixed gender) are not included in the list. Further, asking about sex assigned at birth is a sensitive question for some people and likely is not appropriate when privacy and confidentiality cannot be assured (e.g., small sample studies, employment contexts). Finally, sex assigned at birth does not consider the nonbinary options that have recently become available on birth certificates.

If also collecting race data, researchers may choose to include Two-Spirit as a response option. Two-Spirit refers to a broad array of tribally specific gender and sexual orientation identities that cannot be reflected in the English language or western thought.[21] The U.S. Department of Health and Human Services, Indian Health Service clarifies that people who identify as Two-Spirit are not considered men or women but occupy a distinct, alternative gender status.[22] Because Two-Spirit is deeply tied to both gender identity and indigenous experience, it is recommended that this question be only asked of those who report AIAN status.

There may be situations in which it is appropriate to identify intersex individuals. If identifying this population is relevant to the research, the National Academy of Sciences recommends doing so with a standalone measure and to avoid including intersex as a response category to a binary measure of sex.

Recommended Questions for Gender Identity

Based on a review of institutional descriptors, we have adapted the National Academy of Sciences approach and recommended this adaptation as best practice.

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

- [ ] Female
- [ ] Male
- [ ] Don't know
- [ ] Prefer not to answer

What is your current gender identity?

- [ ] Female
- [ ] Male
- [ ] Transgender

☐ [If collecting race data and respondents may identify as AIAN]: Two-Spirit

☐ Nonbinary

☐ I use a different term [Open text box]

☐ Don't know

☐ Prefer not to answer

NOTE: *We do not recommend collecting sex assigned at birth as a standalone item in any data collection context because sex and gender identity are intertwined and social structures influence behaviors that appear to be rooted in biology. However, it may be appropriate to collect gender as a standalone question, depending on the intent of the evaluation and when evaluators are clear that a person's gender identity does not necessarily equate or dictate a person's sex. As reiterated throughout this guide, ideally every new engagement includes some conversation with clients and populations of study to clarify appropriate terminology and goals for data collection.*

Recommended Reporting for Gender Identity

There are two important considerations for reporting gender data. First, include any write-in categories to tabulations of responses. Monitoring write-ins helps researchers identify whether existing categories reflect those used in the general population or no longer make sense.

Second, as with race, reporting data in small, finite categories may be ethically irresponsible. It may be appropriate to report a general "transgender" identity over more detailed subcategories, such as "transgender man" or "transgender woman" because the sample sizes for these groups are likely to be small and may need to be aggregated to ensure the confidentiality of individuals participating in the survey.

Again, it is ultimately best practice to allow people to indicate how they would like to be described and include them in the process of evaluation planning and reporting, when possible.

Sexual Orientation

The following section synthesizes findings from the literature review regarding best practices for collecting information related to sexual orientation. Given that people who identify as LGBTQ have often been marginalized and, therefore, face economic, health, and other disparities, it is widely agreed that data on sexual orientation should be collected across human service efforts to identify and address specific disparities appropriately. However, best practices regarding how specifically to collect this information has been debated.

Funders for LGBTQ Issues (Funders) recommends collecting survey data on sexual orientation for the same reasons that many institutions collect data on race, ethnicity, and gender identity, not only to understand disparities, but also to understand progress and to ensure that institutions reflect the diversity of communities served.[23] Funders draws upon the William's Institute of the University of California, Los Angeles (UCLA) School of Law's best practice guide to ask about sexual orientation. This guide was informed by a panel of national experts and evidence-based research and recommends using a trivariate approach with heterosexual or straight, gay or lesbian, and bisexual options. The Williams Institute and Funders both recommend this language because the terms utilized are the most easily understood by most people, providing the most accurate data.[24]

Other common, evidence-based sources, particularly public health related sources, utilize a multi-question approach regarding sexual orientation based on the understanding that sexual behavior does not always align with sexual orientation. The Youth Risk Behavior Surveillance System (YRBSS) is one such source.

YRBSS monitors health-related behaviors among youth across the country. YRBSS is operationalized via a number of national and state surveys distributed by schools, health

agencies, and local and tribal governments.[25] YRBSS defines sexual orientation as consisting of two components: sexual identity and sex of their sexual contacts. Therefore, surveys include questions about both components to identify sexuality minority youth. The 2015, 2017, and 2019 YRBSS surveys ascertained sexual identity using a similar trivariate option as the Williams Institute and Funders, but they also included an option for those who were unsure.

A second question for those three survey years asked respondents about the sex of their sexual contacts, including I have never had sexual contact, Females, Males, and Females and Males.

Additionally, the Trevor Project, the world's largest suicide prevention and crisis intervention organization for young people who identify as LGTBQ, produced a report in 2021 titled *Measuring Youth Sexual Orientation and Gender Identity*. This report provides findings from a literature review, internal lessons learned regarding surveying LGBTQ youth, and best practices and recommendations for measuring LGBTQ identities. Based upon the literature review findings and survey results from previous years, the Trevor Project recommends a set of two questions, the first being open-ended in order to allow youth flexibility and autonomy in describing their sexual orientation and the second in a structured format that enables coding for effective data analysis (e.g., straight or heterosexual, gay, lesbian, bisexual, queer, pansexual, asexual, I am not sure, I don't know what this question means, and decline to answer).

Based upon write-in answers on previous years' surveys, the Trevor Project expanded its list of options for the second question to capture as much of the LGBTQ youth population as possible. However, the Trevor Project also recognizes that such an expansive list may not always be appropriate and therefore suggests careful consideration about the population being surveyed and the purposes of the instrument.

More generally, the Trevor Project urges researchers to be intentional about survey item selection based upon the data needs of the researcher.[26]

While different entities have utilized different language and approaches to asking about sexual orientation, it is widely agreed that questions about sexual orientation should be asked separately from gender identity questions, and, when necessary, questions need to reflect the reality that sexual behavior may not always align with sexual orientation. The most recent literature suggests that providing individuals with an opportunity to self-describe their sexual orientation may provide insight into categories that may otherwise be missed and exclude portions of the population.

Recommended Questions for Sexual Orientation

Based on a review of the literature, the following survey question is recommended for collecting information on sexual orientation in surveys:

1. Which of the following best describes you?

☐ Heterosexual (straight)
☐ Gay or Lesbian
☐ Bisexual
☐ Not Sure
☐ Other [Open text box]
☐ Prefer not to answer

However, data should be carefully considered across projects, and if a significant percentage of individuals selects "Other" and provides answers such as Pansexual or Asexual, future surveys should be adapted to provide a more expansive list when appropriate for the target population.

In some cases, it may be helpful to provide the population with the opportunity to self-identify

their sexual orientation in a noncategorical, open-ended manner. This may further encourage comfortability, inclusion, and autonomy and provide additional qualitative data that may otherwise not be captured. For those cases, surveys may include the Trevor Project's open-ended introductory question:

2. Open Response Question: Sexual orientation is a person's emotional, romantic, and/or sexual attractions to another person. There are many ways a person can describe their sexual orientation and many labels a person can use. How would you describe your current sexual orientation in your own words? [Open text box]

Recommended Reporting for Sexual Orientation

According to the Williams Institute at UCLA, responsible analysis and reporting of sexual orientation data requires consideration of three factors:

1. Careful consideration of differences among non-heterosexual responses: When the size of the sample allows for meaningful analysis at the subgroup level, subgroups such as lesbian/gay, bisexual, man, and woman should be analyzed separately to accurately understand differences in social, health, economic, and other outcomes.

2. Separate demographic subgroup analyses of outcomes: Researchers should consider additional subgroups based on characteristics such as race, as evidence suggests that LGBTQ outcomes can vary substantially across a variety of demographic subgroups.

3. Careful consideration of contextual factors: Researchers should try to understand and provide appropriate context to aid in the accurate and unbiased interpretation of results.[27]

Appendix A: Methodology

PCG employed a three-step project approach for development of data collection and reporting standards. This document is designed to be a guide for evaluators by evaluators. First, the PCG evaluation team made a list of agencies, organizations, and funders who are:

✓ Sources frequently referenced in evaluation work (e.g., Substance Abuse and Mental Health Services (SAMHSA), Centers for Disease Control and Prevention (CDC), Child Welfare League of America (CWLA), Casey Family Programs)

✓ Managers for large data sets frequently used in evaluation (e.g., Youth Risk Behavior Surveillance System, Behavioral Risk Factor Surveillance System, Census)

✓ National advocacy entities for race, ethnicity, sexual orientation, and gender identity (e.g., American Civil Liberties Union (ACLU), National Association for the Advancement of Colored People (NAACP), Trevor Project, Human Rights Campaign)

✓ National leaders for evaluation or writing best practice (e.g., American Evaluation Association, American Public Health Association, National Academy of Sciences, Associated Press Style Guide)

Next, the project team systematically reviewed each source for publicly available reports, data collection tools, dashboards, and policies regarding language use around race, ethnicity, and SOGIE. The team recorded information on how each source used language around the highlighted topics and noted any best practices or policies promoted by the source. Additionally, during this review the team was asked to take note of related partners identified by the initial list of sources. Related resources were also searched for information pertaining to the race, ethnicity, and SOGIE. Approximately 40 sources were identified and reviewed in total.

Finally, the team reviewed the body of work by topic, comparing language usage and best practice among sources. Each data collection section was written as a synthesis of this process to put forth our most inclusive recommendations for question development and reporting of each topic area.

About PCG

Public Consulting Group LLC (PCG) is a leading public sector solutions implementation and operations improvement firm that partners with health, education, and human services agencies to improve lives. Founded in 1986 and headquartered in Boston, Massachusetts, PCG employs approximately 2,000 professionals worldwide—all committed to delivering solutions that change lives for the better. PCG helps state, county, and municipal human services agencies respond to change, enhance operations, improve access to services, optimize processes and compliance, and achieve their performance goals in order to better serve populations in need. PCG's seasoned professionals offer proven solutions to help agencies design programs, services, and systems; increase program revenue; cut costs; and improve regulatory compliance with state and federal regulations.

To learn more about demographic data collection and reporting, contact us today.

📞 (800) 210-6113   @ info@pcgus.com   🖥 www.publicconsultinggroup.com/humanservices/

Endnotes

1   SOGIE is an emerging inclusive and expansive term that describes a wide spectrum of people, including people who are heterosexual, rather than simply focusing on the LGBTQ (Lesbian, Gay, Bisexual, Transgender, and Queer) or the LGBTQIA+ (Lesbian, Gay, Bisexual, Transexual, Questioning or Queer, Intersex, Asexual, and others) community in ways that can lead to bias and perpetuated stereotypes. Many states, such as New York and entities such as the United Nations are shifting to the use of SOGIE language in reporting of information regarding sexual orientation, gender identity and expression.1 Using SOGIE language allows researchers to detail subgroup outcomes, such as lesbian and gay or bisexual outcomes, within the broader context of all people.

2   U.S. Centers for Disease Control and Prevention. (2021). "Health Equity Guiding Principles for Inclusive Communication." Retrieved from https://www.cdc.gov/healthcommunication/Health_Equity.html.

3   Office of Disease Prevention and Health Promotion. HealthyPeople.gov. (2022). "Social Determinants of Health | Healthy People 2020." Retrieved from https://www.healthypeople.gov/2020/topics-objectives/topic/social- determinants-of-health.

4   National Archives. Federal Register. (2016). "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity." Retrieved from https://www.federalregister.gov/documents/2016/09/30/2016-23672/standards- for-maintaining-collecting-and-presenting-federal-data-on-race-and-ethnicity.

5   Alchemer. (2021). "How to Write Better Demographic Survey Questions (with Examples)." Retrieved from https://www.alchemer.com/resources/blog/how-to-write-better-demographic-questions/?sg=1.

6   National Human Genome Research Institute. (2022). "Race." Retrieved from https://www.genome.gov/genetics- glossary/Race.

7   Gendered Innovations. (2021). "Race & Ethnicity." Retrieved from https://genderedinnovations.stanford.edu/terms/race.html.

8   Gendered Innovations. (2021). "Race & Ethnicity." Retrieved from https://genderedinnovations.stanford.edu/terms/race.html.

9   NAACP. (2022). "Race & Justice." Retrieved from www.naacp.org.

10   United States Census Bureau. (2017). "2015 National Content Test Race and Ethnicity Analysis Report." Retrieved from https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis- reports/2015nct-race-ethnicity-analysis.pdf.

11   United States Census Bureau. (2017). "2015 National Content Test Race and Ethnicity Analysis Report." Retrieved from https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis- reports/2015nct-race-ethnicity-analysis.pdf.

12   United States Census Bureau. (2021). " Improvements to the 2020 Census Race and Hispanic Origin Question Designs, Data Processing, and Coding Procedures." Retrieved from https://www.census.gov/newsroom/blogs/random-samplings/2021/08/improvements-to-2020-census-race-hispanic- origin-question-designs.html.

13   ACLU. (2022). "Coalition Letter Urging OMB to Revisit Standards on Federal Data on Race and Ethnicity." Retrieved from https://www.aclu.org/letter/coalition-letter-urging-omb-revisit-standards-federal-data-race-and- ethnicity.

14   Carroll, S.R., Rodriguez-Lonebear, D. and Martinez, A. (2019). "Indigenous Data Governance: Strategies from United States Native Nations. *Data Science Journal*, 18(1), p.31." Retrieved from https://doi.org/10.5334/dsj-2019- 031.

15   The Trevor Project. (2021). "Guide to Being and Ally to Transgender and Nonbinary Youth." Retrieved from https://www.thetrevorproject.org/wp-content/uploads/2021/07/Guide-to-Being-an-Ally-to-Transgender-and-Nonbinary- Youth.pdf.

16   National Academies of Sciences, Engineering, and Medicine. (2022). *Measuring Sex, Gender Identity, and Sexual Orientation*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26424.

17   Lagos, D., & Compton, D. L. (2021). Evaluating the use of a two-step gender identity measure in the 2018 General Social Survey. *Demography,* 58(2), 763-772.

18   GenIUSS Group. (2014). Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys. Williams Institute. Los Angeles, CA.

19   Johns, M. M., Lowry, R., Andrzejewski, J., Barrios, L. C., Demissie, Z., McManus, T., ... & Underwood, J. M. (2019). Transgender identity and experiences of violence victimization, substance use, suicide risk, and sexual risk behaviors among high school students—19 states and large urban school districts, 2017. *Morbidity and Mortality Weekly Report*, 68(3), 67. http://dx.doi.org/10.15585/mmwr.mm6803a3

[20] Lagos, D., & Compton, D. L. (2021). Evaluating the use of a two-step gender identity measure in the 2018 General Social Survey. Demography, 58(2), 763-772.

[21] Wilson, A. (1996). How we find ourselves: Identity development and two spirit people. *Harvard Educational Review*, 66(2), 303-318.

[22] U.S. Department of Health and Humans Services, Indian Health Service. (n.d.) "Two-Spirit." Retrieved from https://www.ihs.gov/lgbt/health/twospirit/

[23] Funders for LGBTQ Issues. (2019). "Data Collection." Retrieved from https://lgbtfunders.org/resources/best- practices-for-foundations-on-collecting-data-on-sexual-orientation-and-gender-identity/

[24] Funders for LGBTQ Issues. (2019). "Data Collection." Retrieved from https://lgbtfunders.org/resources/best- practices-for-foundations-on-collecting-data-on-sexual-orientation-and-gender-identity/

[25] U.S. Centers for Disease Control and Prevention. (2020). "Youth Risk Behavior Surveillance System (YRBSS)." Retrieved from https://www.cdc.gov/healthyyouth/data/yrbs/index.htm

[26] DeChants, J., Green, A.E., Price, M.N, & Davis, C. (2021). Measuring Youth Sexual Orientation and Gender Identity. West Hollywood, CA: The Trevor Project.

[27] UCLA School of Law Williams Institute. 2022. Best Practices for Asking Questions about Sexual Orientation on Surveys. Retrieved from https://williamsinstitute.law.ucla.edu/publications/smart-so-survey/