

# Response to the Request for Information for **Federal Evidence Agenda on LGBTQI+ Equity**

**October 3, 2022**

---

**Submitted to:**

NSTC Subcommittee on Equitable Data  
Office of Science and Technology Policy  
Eisenhower Executive Office Building  
1650 Pennsylvania Ave. NW,  
Washington, DC 20504  
Attention: Meghan Maury, Senior Advisor for Data Policy

**Submitted by:**

Mathematica  
1100 First Street, NE, 12th Floor  
Washington, DC 20002-4221  
Phone: (202) 484-9220  
Fax: (609) 228-4958

At Mathematica, we believe diverse perspectives and varied lived experiences inform high-quality evidence. For more than 50 years we have worked side by side with federal agency staff, grantees, foundations, community organizations, and our other partners in the evidence community to uncover insights about what works and how best to improve public well-being. In this capacity, we often collect data ourselves, analyze federal survey data, and work with our agency partners to examine administrative data. We know data. We know the unique position the federal government holds in the evidence community. We know that improving the federal government's ability to make data-informed policy decisions is critical to advancing equity for the lesbian, gay, bisexual, transgender, queer, and intersex (LGBTQI+) communities. That's why we eagerly provide this response to the recent request for information on the Federal Evidence Agenda on LGBTQI+ Equity.

We are pleased that the National Science and Technology Subcommittee on Equitable Data intends to improve federal efforts to collect data related to sexual orientation and gender identity (SOGI), and variations in sex characteristics. We encourage the Subcommittee to do so by focusing on inclusive and flexible data collection strategies informed by intentional engagement with LGBTQI+ communities and a commitment to community-based research principles. By modeling after and partnering with community-based research, federal agencies can more accurately estimate disparities facing LGBTQI+ communities and promote important subsample analysis, particularly for transgender, nonbinary, and intersex people and LGBTQI+ youth. We also encourage the Subcommittee to facilitate collaboration between the federal government and community organizations to link community-based and federal data to have an even greater impact.

Thank you for giving Mathematica and other members of the evidence community an opportunity to provide input on the Federal Evidence Agenda on LGBTQI+ Equity and contribute to this important conversation. We look forward to engaging with the Subcommittee in the future.

## **Section 1: Describing Disparities**

**Question 1.1 - What disparities faced by LGBTQI+ people are not well-understood through existing Federal statistics and data collection? Are there disparities faced by LGBTQI+ people that Federal statistics and other data collections are currently not well-positioned to help the Government understand?**

Analyses of data from the Federal Reserve, [Census Bureau](#), and [U.S. Department of Justice](#) (DOJ), which recently began to collect SOGI data, reveal disparities in economic well-being, crime victimization, and access to public safety or law enforcement. However, the SOGI measures used in these and other federal data sets inhibit federal agencies' ability to accurately identify LGBTQI+ individuals, characterize the disparities they face, and focus federal resources to address challenges. There can be several types of limitations:

- 1. Existing data collection methodologies can obscure disparities experienced by people with transgender and nonbinary identities.** Transgender and gender nonconforming people face disparities in [health care](#) and [public safety](#) that federal surveys are poorly positioned to understand. Research has shown the importance of using [a two-step approach in treating SOGI as different constructs](#) that clearly distinguish between SOGI and reporting practices that follow a similar approach. For example, [limited response options](#) to gender identity questions can affect population estimates and misrepresent the identities and experiences of LGBTQI+ people in

relation to drug, tobacco, and alcohol use. These recommendations differ from a two-step approach used in recent [Census Bureau](#) and [DOJ](#) surveys, which first ask the respondent's sex assigned at birth ("male or female") and then ask if the respondent describes themselves as "transgender, male, female, or none." This approach will result in inconsistent answers from binary transgender people; for example, some transgender women might select "female" and others might select "transgender" depending on which term they more strongly identify with. The "none" option will include some nonbinary people, but others could identify as transgender, and still others could mark their sex assigned at birth, leaving no clear mechanism for reporting nonbinary identities. Including these questions in Census Bureau and DOJ products is a large improvement, but the response options limit the usefulness of the data collected to estimate disparities and understand the composition of the transgender, nonbinary, and gender nonconforming population.

- **Federal data collection rarely captures the experiences of intersex people.** Only the [Survey of Youth Transitioning from Foster Care](#) collected information on variations in sex characteristics in 2021. Analyses in [other countries](#) show intersex people experience significant discrimination in schooling and higher rates of suicidal ideation. The lack of federal data prevents our ability to understand the extent to which intersex people in the United States face these and other disparities.
- **Small sample sizes limit our ability to understand disparities faced by LGBTQI+ youth** LGBTQI+ youth are at heightened risk for violence, bullying, and crime; are overrepresented in the [juvenile justice](#) and [foster care](#) systems; and experience adverse [mental health and educational outcomes](#). Small sample sizes in key tools such as the Youth Risk Factor Surveillance System (YRBSS) limit youth experiences in federal data, making analysis of these disparities difficult. In addition, [audits of state juvenile facilities](#) report inconsistent quality of data on LGBTQI+ youth experiences and staff discomfort and unfamiliarity with SOGI topics limit its use to inform decisions.

**Question 1.2 - Are there community-based or non-Federal statistics or data collection that could help inform the creation of the Federal Evidence Agenda on LGBTQI+ Equity? Are there disparities that are better understood through community-based research than through Federal statistics and/or other data collection.**

The Subcommittee could consider modeling federal SOGI data collection processes after successful nonfederal and community-based research models that have addressed the data gaps identified in our response to Question 1.1. The Subcommittee could consider collaborating with these organizations to link federal survey, Census, and administrative data with data collected by community organizations to draw on the strengths of community-based research and longitudinal federal data:

- 1. Data that better reflect the unique experiences of trans, nonbinary, and intersex communities.** The [U.S. Transgender Survey](#) and [LGBTQ Women's Survey](#) include nuanced questions on different aspects of the respondents' own gender identity, their partners' gender identity, and how this intersects with their sexual orientation. Questions include more options than federal surveys, with multiple selections allowed and include many transgender, nonbinary, and intersex options.
- 2. Data collection that better reflects the lived experiences of LGBTQI+ communities.** The Trevor Project has far more [response options](#) for its questions on sexual orientation. Likewise, it

asks questions about conversion therapy, specific experiences of discrimination based on SOGI, and access to LGBTQ+ affirming spaces—key elements of LGBTQI+ youth health neglected in federal data collection.

- 3. Data that include chosen family more accurately reflects LGBTQI+ social ties.** Federal data sets have improved the data collection process for household and family characteristics of same-sex couples, particularly for households headed by same-sex couples. However, chosen family—meaning family-like ties that are outside of traditional legal and biological relationships—are important [social structures](#) within LGBTQI+ communities and federal data collection processes do not currently record them well. Community-based research that uses broader measures of family relationships demonstrate the significant social and economic importance of chosen family—particularly for LGBTQI+ [refugees](#), LGBTQI+ people with [disabilities](#) and [other health concerns](#), and [older](#) LGBTQI+ people.
- 4. Larger sample sizes of LGBTQI+ youth enable important subgroup analysis.** To obtain larger samples of LGBTQI+ youth and focus on their specific needs, the Trevor Project recruited a representative sample of 40,001 LGBTQI+ youth ages 13 to 24. This is a larger LGBTQI+ sample than YRBSS, while matching YRBSS on many demographic attributes. In addition, partnerships with Black, Latinx, and Asian Pacific Islander LGBTQI+ organizations enable these data to highlight intersectional identities.

**Question 1.3 - Community-based research has indicated that LGBTQI+ people experience disparities in a broad range of areas. What factors or criteria should the Subcommittee on SOGI Data consider when reflecting on policy research priorities?**

We believe those with lived experiences can best identify what research would best address their needs. As such, we recommend the Subcommittee co-create factors or criteria with members of LGBTQI+ communities to identify its policy research priorities. Engaging directly with diverse LGBTQI+ communities using [equitable evaluation practices](#) can result in research findings and dissemination approaches that better address community needs. In service of this collaboration, the Subcommittee should engage in several activities:

1. Consider using community-based participatory research (CBPR) practices to co-create their agenda with a diverse group of LGBTQI+ communities, including people of different races and ethnicities, gender identities, sex characteristic variations, ages, locations, religions, immigration statuses, disability statuses, people living with HIV/AIDs, and people experiencing homelessness.
2. Encourage federal agencies to fund research and evaluations that support the use of CBPR practices to co-create meaningful research objectives and outcomes. This could include requiring [community needs assessments](#) and advisory boards to address known barriers to engagement in federally funded evaluations.
3. Encourage federal agencies to consider the sustainability of proposed research and prioritize work that aims to generate action items that LGBTQI+ communities can lead, direct, and support.
4. Encourage federal agencies to invest in inclusive standards for evidence-based practice, to expand how the [field prioritizes evidence](#), and to think critically about innovative research to improve LGBTQI+ equity across intersectional identities

## Section 2: Informing Data Collections

**Question 2.1** In some instances, there are multiple surveys or data collections that could be used to generate evidence about a particular disparity faced by the LGBTQI+ community. In addition to factors like sample size, timeliness of the data, and geographic specificity of related data publications, what other factors should be considered when determining which survey would best generate the relevant evidence? Are there data collections that would be uniquely valuable in improving the Federal Government’s ability to make data-informed decisions that advance equity for the LGBTQI+ community?

The Subcommittee should explore including a standardized set of inclusive SOGI items in the demographics section of all federally funded, population-based surveys collecting demographic data and recommend oversampling LGBTQI+ populations. Existing population-based federal surveys already provide the nation’s most rigorous cross-sectional and longitudinal data on key substantive topics and include a consistent set of demographic questions that allow for disparity analysis. These surveys are well positioned to add an expanded set of standard and inclusive SOGI items, and to oversample SOGI populations, which will improve representation of LGBTQI+ communities in the data. Such standardization allows for (1) benchmarking across studies using high-quality research designs and methods, (2) assessing needs within LGBTQI+ communities related to the specific topical areas of the existing surveys, and (3) assessing change within the distinctive LGBTQI+ communities over time.

Surveys provide the potential for large sample sizes and data linking: we recommend the Subcommittee consider conducting a new, national federal data collection focusing on the LGBTQI+ communities and the breadth of topics affecting their lived experiences. Nongovernmental organizations fund and lead current surveys focused specifically on LGBTQI+ populations, such as the [U.S. Trans Survey](#) and the [National Survey on LGBTQ Youth Mental Health](#). They provide valuable data for exploring knowledge gaps and determining the unique disparities and existing inequalities experienced by the diverse communities under the LGBTQI+ population. However, these studies can lack the rigor and quality of large, nationally representative federal surveys, due to smaller sample sizes, nonprobability sample designs, and lower response rates. These are important factors for reducing bias in the data and producing high-quality data with the sample sizes necessary to disaggregate intersectional identities (for example, economic status and LGBTQI+ identity).

**Question 2.2 - To protect privacy and maintain statistical rigor, sometimes publicly released data must combine sexual and gender minority respondents into a single category. While this approach can provide valuable evidence, it can also obscure important details and differences. Please tell us about the usefulness of combined data, and under what circumstances more detailed data may be necessary.**

We recognize the need to uphold privacy concerns and maintain statistical rigor. The Subcommittee should make it clear to federal agencies that collecting and disseminating more detailed data about LGBTQI+ communities is a priority and should be part of federal surveys by default. The Subcommittee should take this opportunity to align relevant federal data collection practices with the National Academies of Sciences, Engineering, and Medicine [consensus recommendation](#) that clearly distinguishes between sexual orientation and gender identity, and reporting practices that

follow a similar approach. Detailed data reporting requirements can support better policy outcomes that capture more nuance.

**Question 2.3 - Are there any Federal surveys or administrative data collections for which you would recommend the Federal Government should not explore collecting SOGI data due to privacy risk, the creation of barriers to participation in Federal programs, or other reasons? Which collections or type of collections are they, and why would you make this recommendation?**

**Mathematica recommends using demographic questions, particularly inclusive SOGI items, in existing federal surveys to improve policy outcomes and better represent various subgroups (Question 2.1).** These surveys can handle privacy risks through imputation and data-masking procedures (Section 3). Comprehensive demographic data at the national level are essential for understanding and addressing, through policies, disparities experienced by marginalized populations. All federal, population-based surveys should incorporate inclusive SOGI items.

**The government should consider and mitigate any increased privacy risks and potential unintended harmful ramifications due to including expanded SOGI data.** Responding to inclusive SOGI items in administrative data can pose a higher risk to LGBTQI+ participants when compared to the perceived risk of responding to similar questions on a federal survey. For example, LGBTQI+ participants thought [SOGI items asked in administrative contexts](#) such as post office forms were more intrusive than in health contexts, because the relevance of SOGI information was not obvious. We recommend the government consider collecting data alongside expected data use and the risk to LGBTQI+ communities, making clear the relevance for and use of SOGI information. Further, using inclusive SOGI items could create barriers to program participation—for example, when a program application is deemed invalid if the gender identity of the applicant differs between their application and the government-issued ID used for verification. Similarly, linking administrative data to agencies or programs could create risk for LGBTQI+ people. State laws differ regarding degrees of discrimination for LGBTQI+ communities in updating identification documents, health care, education, housing, child welfare services, and more.

**Question 2.4 - How can Federal agencies best communicate with the public about methodological constraints to collecting or publishing SOGI data? Additionally, how can agencies encourage public response to questions about sexual orientation and gender identity in order to improve sample sizes and population coverage?**

**We recommend federal agencies use existing communication guidance.** All data collections contain survey errors and the potential for resulting bias and that should be communicated following the [federal plain language guidelines](#). The Centers for Disease Control provides a useful [Clear Communication Index Tool](#) for improving public communication related to disseminated products.

**Strategies for increasing item-level response to inclusive SOGI questions should follow the same best practices used when asking the public to respond to any question of a personal nature.** These practices include several steps:

1. Stating at the start of data collection, and again at the sensitive items, that responses are confidential and reported in summary form only (that is, not individually)
2. Clearly explaining the need for and use of the data



3. Indicating that respondents may skip questions they do not feel comfortable answering and by providing a response option such as “I don’t want to answer” to personal questions
4. Placing the demographic questions, including SOGI items, at the end of a survey
5. Considering the mode of data collection; research suggests [web-based surveys](#) have higher item response rates and improved data quality for personal or sensitive questions when compared to telephone or in-person surveys

**Question 2.5 - Data collection on vulnerable populations is often incomplete, creating challenges for techniques help identify missing SOGI data, and make statistically rigorous estimates for that missing data. How should qualitative information help agencies analyze what SOGI data might be missing?**

Qualitative information can provide critical insights into SOGI data. Agencies can take several steps:

1. **Engage with LGBTQI+ community advocates, experts, and community-based organizations to consider data collection.** Using focus groups and other qualitative participatory data collection methods can help agencies fully understand the challenges faced by relevant LGBTQI+ communities related to collecting data. This can help accurately capture identities and ensure data collection methods elicit the highest response rates for them. Consistent community engagement can help agencies stay up to date on evolving community preferences regarding SOGI terms.
2. **Implement qualitative coding of open-ended survey responses to help improve the understanding of emerging terms and how members of the population self-identify.** Culturally responsive and equitable evaluation practices emphasize the need to [combine qualitative methods with quantitative methods](#). Emergent themes can provide helpful insights to missing data. For example, some survey respondents might not have completed the question if the way they identify is not an option to select. Coding open-ended survey responses should illuminate common write-in responses, providing insight for updates to the survey that can increase responses.
3. **Hold reoccurring stakeholder or technical working group meetings with research and community experts.** Engagement provides multiple opportunities for partners’ input and helps agencies understand the population and how to frame SOGI [data collection materials](#), questions, and response options. Such meetings should not only inform language and framing that is appropriate, but also discuss outdated characterizations that should be discontinued.

### Section 3: Privacy, Security, and Civil Rights

**Question 3.2 - Unique risks may exist when collecting SOGI data in the context of both surveys and administrative forms. Please tell us about specific risks Federal agencies should think about when considering whether to collect these data in surveys or administrative contexts.**

**Collecting SOGI survey and administrative data for research or administration presents disclosure risks and requires careful review and revision to protect people’s confidentiality.** Federal agencies often make survey and administrative data available for use by researchers broadly to take advantage of the data that have the potential to grow knowledge on LGBTQI+ disparities. We recommend systematically reviewing data to [identify disclosure risks](#) that might arise either

through direct disclosure (such as data sets that include identifying information such as names, addresses, or unique ID numbers); or indirect disclosure (through variables that are stable over time, distinguishing, and knowable; variables that are rarely observed, such as SOGI data, or atypical; or through combinations of variables that could jointly reveal peoples' identities, such as being the only male teacher in an area when being neither male nor being a teacher is rare on its own). Most of the disclosure review process involves identifying and mitigating risk from potential indirect identifiers and this process requires substantial resources to achieve the goal of protecting people's confidentiality.

**Question 3.3 - Once SOGI data have been collected for administrative or statistical purposes, are there considerations that Federal agencies should be aware of concerning retention of these data? Please tell us how privacy or confidentiality protections could mitigate or change these concerns.**

**When programs collect SOGI data for administrative or statistical purposes, protecting against unauthorized or inadvertent disclosure is particularly crucial for LGBTQI+ youth.** LGBTQI+ youth face [bullying and harassment at school](#) and can experience [parental abuse and greater risk of suicide](#) after being outed to parents, teachers, and friends. However, the belief that the *topic matter* of SOGI poses inherent risk to youth is unfounded. This idea has threatened federal data collection before, with the [Bureau of Justice Statistics](#) proposing in 2018 to remove SOGI questions from the National Crime Victimization Survey due to “potential sensitivity ... for adolescents.” As [researchers have argued](#), decades of successful and age-appropriate sexual orientation questions asked on other school surveys in the United States and around the world contradict this stance. For example, self-administered federal youth surveys [like the YRBSS](#) conducted by the Centers for Disease Control and Prevention protect privacy and promote safe respondent SOGI reporting as recommended by [The Trevor Project](#).

**In addition to mitigating both direct and indirect disclosure risks, the Subcommittee should ensure the transfer and use of SOGI data for administrative purposes does not hinder people from seeking public services to which they are legally entitled.** A lack of comprehensive gender identity categories in federal surveys and [discrepancies](#) in sex or gender marker options across state and federal identification documents frequently cause inconsistencies between transgender people's documents and their sex or gender identities—even across the same individual's documents. As a result, [many transgender people face delays](#) in receiving Medicare, Medicaid, supplemental income, and employment approval.



---

**Mathematica Inc.**

Princeton, NJ • Ann Arbor, MI • Cambridge, MA  
Chicago, IL • Oakland, CA • Seattle, WA  
Tucson, AZ • Woodlawn, MD • Washington, DC

**EDI Global, a Mathematica Company**

Operating in Tanzania, Uganda, Kenya, and the United Kingdom

Mathematica, Progress Together, and the “spotlight M” logo are registered trademarks of Mathematica Inc.



[mathematica.org](https://mathematica.org)