

June 26, 2023

OMB Desk Officer for National Science Foundation
Office of Information and Regulatory Affairs
White House Office of Management and Budget
725 17th Street, NW
Washington, DC 20503

Suzanne H. Plimpton
Reports Clearance Officer
National Science Foundation
2415 Eisenhower Ave.
Alexandria, VA 22314

CC: Charles Barber, Chief Diversity and Inclusion Officer, National Science Foundation
The Honorable Tammy Baldwin, U.S. Senate
The Honorable Dianne Feinstein, U.S. Senate

Via reginfo.gov and email

RE: National Science Foundation; Notice of Submission for OMB Review; 2023 Survey of
Doctorate Recipients (**Federal Register Doc. 2023-13279**)

Dear OMB and NSF Officials:

Thank you for the opportunity to comment on the National Science Foundation's (NSF) information collection request related to the 2023 Survey of Doctorate Recipients (SDR). See 88 Fed. Reg. 40870 (June 22, 2023). My comments are concerned with the 2023 SDR's Sexual Orientation and Gender Identity (SOGI) Experiment Plan ([Attachment D5](#)). I appreciate that NSF has continued its recently revised practice of providing sufficient details regarding its proposed SOGI pilot methodology and question designs prior to OMB approval, as advocated in the March 3, 2023 [letter](#) to the NSF Director and as is intended by [5 CFR 1320.8\(d\)](#).

I commend NSF's National Center for Science & Engineering Statistics' (NCSES) extension of its SOGI measurement efforts to the SDR. As you know, SOGI data are critical to addressing the alarming [disparities](#) LGBTQI+ people are facing in STEM, including increasingly documented [career barriers](#), [underrepresentation](#), [retention failure](#), and [some of the highest levels of harassment in STEM](#). As our January 13, 2023 [letter](#) to the NSF Director pointed out, which was affirmed by a [subsequent letter](#) sent from Senators Tammy Baldwin, Dianne Feinstein, and 16 other U.S. senators, SOGI data in NCSES surveys are important for implementing federal policy, such as diversity-related funding provided by the CHIPS and Science Act, and for fulfilling the directives of [Executive Order 14075](#) to advance the use of SOGI data to address LGBTQI+ disparities, and [Executive Order 13985](#) to improve the equity of underserved communities.

The proposed pilot involving three different batteries of SOGI questions across a subset ($n = 12,000$) of the SDR population is a welcome next step and complements NCSES' current SOGI testing in the Survey of Earned Doctorates (SED) and National Survey of College Graduates (NSCG). However, there are a few methodological concerns that NCSES should address prior to launching the SDR.

Summary of Recommended Changes to the SDR SOGI Experiment Plan

- Battery 1 Sexual Orientation: “Straight, that is not gay or lesbian” → “Straight or heterosexual”
- Battery 3 Sexual Orientation: “Straight” → “Straight or heterosexual”
- GP1: “How comfortable or uncomfortable are you with providing your gender identity information in a U.S. federal government survey?” → “How comfortable or uncomfortable are you with providing your gender identity information in a U.S. federal government survey that produces official statistics? Your responses would be confidential and used for statistical purposes only.”
- SP1: “How comfortable or uncomfortable are you with providing your sexual orientation information in a U.S. federal government survey?” → “How comfortable or uncomfortable are you with providing your sexual orientation information in a U.S. federal government survey that produces official statistics? Your responses would be confidential and used for statistical purposes only.”
- GP3: “This information is important for ensuring equity for gender minorities” → “This information is important for ensuring equity for gender minorities (e.g., people who identify as transgender, non-binary)”
- SP3: “This information is important for ensuring equity for sexual minorities” → “This information is important for ensuring equity for sexual minorities (e.g., people who identify as gay, lesbian, bisexual)”
- Three respective probe items should be included for the SDR’s race, ethnicity, disability, and/or income questions.

Battery 1 Sexual Orientation Item: “Straight, That Is Not Gay or Lesbian” Should Be Changed to “Straight or Heterosexual” to Align with SED Cognitive Interview Data and Current SED Piloting

[Results](#) from NCSES’ cognitive interviews with the SED found the following (see p. 4):

In the sexual orientation questions, the clauses defining straight as “not lesbian or gay” or “not gay, lesbian, or bisexual” were confusing to some participants because they were not inclusive of other sexual orientations such as queer or asexual. SOGI minority participants and non-SOGI minority participants were equally confused by this. Twenty-eight percent of SOGI minority participants (11 of 39), all of whom were sexual minorities by definition, and 28% of non-SOGI minority participants (6 of 22) expressed confusion about the clause defining straight in the first sexual orientation question tested. For this population, the term straight was clearer without the defining clauses.

These results led NCSES to use the phrase “Straight or heterosexual” for all its [current SED pilot testing](#) (see items SO1-OMB-SO-REV1 and SO3-OMB-SO-REV2). NCSES is also proposing to remove this additional clause in its SDR Battery 3 sexual orientation item. Given its own SED cognitive interview results and to ensure consistency with its current SED testing, NCSES should revise Battery 1’s sexual orientation item from “Straight, that is not gay or lesbian” to “Straight or heterosexual”.

Battery 3 Sexual Orientation Item: “Straight” Should Be Changed to “Straight or Heterosexual”

For the proposed Battery 3 sexual orientation item, NCSES does heed its own SED cognitive interview results and has removed the phrase shown to be confusing to respondents. However, it is proposing to use only “Straight” rather than “Straight or heterosexual”, which creates an unnecessary discrepancy with its own [current SED pilot testing](#). This may result in ambiguity as to whether any potential quality issues identified with this question are due to omission of the negation phrase or due to omission of “or heterosexual”. While the term “straight” is widely known to respondents, the addition of “or heterosexual”

will only aid comprehension and offer an additional term that may resonate better with some non-sexual-minority respondents' self-identification. To be more complete and to increase consistency with its current SED pilot testing (as well as the suggested change to the Battery 1 sexual orientation item; see above), NCSES should expand "Straight" to "Straight or heterosexual" for the Battery 3 item.

GP1 and SP1 Probe Items Are Incomplete: NCSES Must Signal to Respondents That Data Will Only Be Used for Statistical Purposes, Consistent with Prior NSCG Probe Items

In its probe questions, NCSES proposes to ask respondents to answer the following questions: "How comfortable or uncomfortable are you with providing your gender identity information in a U.S. federal government survey?" (GP1) and "How comfortable or uncomfortable are you with providing your sexual orientation information in a U.S. federal government survey?" (SP1). NCSES should provide additional clarifying information, as respondents may wrongly assume that the federal government may use such data in an identifiable way. The proposed approach is also inconsistent with NCSES' previous SOGI comfortability probe items that were assessed on college graduates via Mechanical Turk (MTurk) for use with the NSCG. In that [study](#), NCSES asked college graduates to evaluate the following statements: "I would feel comfortable telling a federal government agency that produces official statistics that I am [fill GENDERID or fill SEXUALID]. Your responses would be confidential and used for statistical purposes only." To accurately signal to respondents the nature of the SOGI data used by the federal government and to be consistent with NCSES' [prior probe items with college graduates](#), GP1 and SP1 should be modified to the following: "How comfortable or uncomfortable are you with providing your gender identity [or sexual orientation] information in a U.S. federal government survey that produces official statistics? Your responses would be confidential and used for statistical purposes only."

NCSES Should Give Examples of "Gender Minorities" and "Sexual Minorities" For GP3 and SP3 Probe Items To Avoid Confusion, Consistent with Battery 2 SOGI Items and Current SED Piloting

Many respondents – LGBTQI+ or otherwise – are not accustomed to the terms "gender minorities" or "sexual minorities", as these are more routinely used by researchers. For the proposed GP3 and SP3 probe items, these terms may lead to unnecessary confusion: "This information is important for ensuring equity for gender minorities" and "This information is important for ensuring equity for sexual minorities". In the Battery 2 SOGI items, consistent with [current SED piloting](#) (see G2-Y/N-OPTION and SO2-Y/N-OPTION), NCSES proposes to provide examples for these terms: "Do you consider yourself to be a gender minority (e.g., transgender, non-binary)?" and "Do you consider yourself to be a sexual minority (e.g., gay, lesbian, bisexual)?" While respondents in the Battery 2 group will see these examples provided, those in the Battery 1 and Battery 3 groups will not. Besides reducing comprehension, this may also introduce a confound between Battery 2 vs. Battery 1 and Battery 3 groups. In all groups, GP3 should be expanded to "This information is important for ensuring equity for gender minorities (e.g., people who identify as transgender, non-binary)" and SP3 should be expanded to "This information is important for ensuring equity for sexual minorities (e.g., people who identify as gay, lesbian, bisexual)".

NCSES Should Include Probe Questions for Other Sensitive Demographic Items That Have Long Been Included in NCSES Surveys To Appropriately Benchmark SOGI Items

Federal statisticians have long regarded SOGI items as potentially sensitive questions, similar to race, ethnicity, disability, income, and other items. Heightened concerns of identifiability and data use are not unique to SOGI but would also apply to similarly sensitive questions. As described in our [January 13, 2023](#) and [March 3, 2023](#) letters to the NSF Director, a major deficiency in NCSES' SOGI measurement efforts over the years is that it does not appropriately benchmark quality assessment metrics of SOGI items to their comparable demographic questions that have long been included in NCSES surveys. This sets NCSES apart from the [approaches](#) of its peer statistical agencies, such as the Department of

Education's National Center for Education Statistics. Without any kind of quantitative comparison, NCSES is able to make subjective and potentially flawed decisions as to what constitutes a data quality issue vs. the inherent nature of a sensitive demographic question.

For example, if 85% of SDR respondents indicate they are very comfortable providing sexual orientation data to a federal statistical agency, NCSES could subjectively judge that 15% of respondents who are less than "very comfortable" is cause for exclusion. Yet if NCSES also were to measure respondents' comfortability with disclosing disability, race, ethnicity, or income to a federal statistical agency, it may find that only 70% of respondents report being very comfortable disclosing this comparable demographic information. Thus, NCSES' continued singling out of SOGI items and its lack of appropriate comparisons will only further the public appearance that NCSES is using a double standard in its consideration of SOGI. For instance, in January 2023, when discussing NCSES' decision to exclude sexual orientation from the 2023 NSCG and the issues raised in our letter, [Science](#) reported that "the agency [NCSES] provided no data to back up those assertions or show how the results compare with those for questions about, for instance, race, disability, or income."

In a letter dated February 24, 2023, the NSF Director made the comparison between NCSES' efforts to collect SOGI data and its efforts to collect non-degree certificate data. SOGI items are potentially sensitive demographic questions and should be compared accordingly. If NCSES will include SOGI items invariably and only intends to use the probe items to distinguish between whether Battery 1, Battery 2, or Battery 3 is optimal, then comparisons with other sensitive demographic questions are not needed. However, if NCSES will use these probe items to make a subjective call as to whether any of these question designs is even sufficient for inclusion in the SDR, then it should justify such a decision with appropriate comparisons by adding identical probe items for race, ethnicity, disability, and income.

I applaud NSF on taking this important step toward advancing LGBTQI+ equity in STEM.

Respectfully Submitted,



Jon Freeman
Associate Professor, Columbia University