Measuring Sexual Harassment in STEM: Implementation Plan

Fors Marsh Project Team
Rachel Walker-Kulzick, PhD
Anna Scolese, MPH
Jordan Loeffelman, PhD
Elizabeth Richards, MA
Leah Fleming, PhD
Marcus Maher, MAS
Marieke Visser, PhD
Laura Severance, PhD

Table of Contents

Table of Contents

Executive Summary 3
Chapter 1: Project Overview 7
Chapter 2: Recommended Constructs of Interest 11
Chapter 3: Considerations for Survey Measures of Sexual Harassment and Related Constructs 17
Chapter 4: Option 1 – Field a New Survey 31
Chapter 5: Option 2 – Field a Supplemental Survey 38
Chapter 6: Option 3 – Add to an Existing Survey 45
Chapter 7: Summary and Future Research Opportunities 52
References 57
Appendix A: Sexual Harassment Measures 61
Appendix B: Message Board Methodology Overview and Key Themes 79
Appendix C: Message Board Protocol 86
Appendix D: Considerations and Recommendations in Surveying Minors 98

Executive Summary

Central to the National Science Foundation's (NSF) National Center for Science and Engineering Statistics' (NCSES) mission is a comprehensive understanding of the STEM workforce and its rapidly changing composition and needs. Although women are achieving STEM degrees at an unprecedented rate, they remain underrepresented in the STEM workforce (NCSES, 2023). In fact, in 2020, women received about half of STEM degrees, yet they only comprised about 35% of the STEM workforce (NCSES, 2023). Likewise, LGBTQIA+ individuals remain underrepresented in the STEM enterprise (Cech & Pham, 2017).

Numerous factors underlie the underrepresentation of women and sexual and gender minorities (SGM), with sexual harassment and gender-based discrimination being chief among them. Harassment and discrimination exert influence throughout the entire career cycle, starting with the selection into STEM programs and careers, continuing through job performance and advancement, and affecting transfers between or out of STEM jobs (Lytle & Shin, 2020). Because women—especially those with multiple marginalized identities (e.g., SGM, people of color)—are more likely to experience sexual harassment, the cumulative effect of these outcomes at an institutional level results in a less diverse field in terms of gender and racial/ethnic identity1 (Beal, 2008; Bowleg et al., 2003; National Academies of Science, Engineering, and Medicine [NASEM], 2018; NASEM, 2020; Richey et al, 2019; Schuyler et al., 2020). In 2018, NASEM published a report detailing the impact of sexual harassment within the STEM enterprise. Using results from large university systems' climate surveys, the authors described the harassment experienced and its impacts on individuals, such as limiting leadership and growth opportunities, pushing individuals to leave institutions, or causing them to leave the field altogether (i.e., pipeline loss; NASEM, 2018). In 2020, NASEM concluded that sexual harassment is one of the chief drivers of the underrepresentation of women—particularly women with multiple marginalized identities—in STEM (NASEM, 2020).

However, as noted by the Committee on National Statistics, there are no high-quality national data sets available to further our understanding of the extent and implications of sexual harassment and its related constructs (e.g., heterosexist harassment) in the STEM workforce (NASEM, 2018). Beyond sexual harassment, it is important to also capture the incidence of behaviors that may occur in conjunction with sexual harassment and contribute to the underrepresentation of people from marginalized backgrounds in STEM. For example, previous research indicated that individuals with multiple marginalized identities are at an increased risk for sexual harassment (Beal, 2008; Bowleg et al., 2003, NASEM, 2020). Therefore, understanding how the intersection of multiple, overlapping identities might affect experiences of sexual harassment within STEM is equally crucial.

This report provides NCSES with an implementation plan that aims to guide the development and, ultimately, the measurement of sexual harassment and related constructs in STEM. Systematically measuring the incidence of these behaviors is a crucial first step to not only understanding the persistence of the problem, but also identifying key risk and protective factors associated with these experiences. Ultimately, gathering this information will enable policymakers to create and adapt programs to prevent sexual harassment (and similar problematic workplace behaviors), appropriately respond to these behaviors to create a welcoming environment for all, enhance the talent attracted to STEM careers, and retain that talent.

We conducted the research informing this implementation plan in three phases: (1) a detailed literature review; (2) a review of NCSES surveys; and (3) qualitative message boards in the form of online message boards. Each of these research milestones contributes to the recommendations we propose in this fourth phase of research, the implementation plan.

[1] Demographic information regarding sexual orientation is currently not collected, so the impact of harassment on SGM individuals within STEM is not known. However, one study found that sexual minority women and gender minorities in astronomy and planetary science experience more workplace harassment compared to cisgender, straight women (Richey et al., 2019). This finding underscores the need for more research on the experiences of SGMs in STEM.

Literature Review

We began this effort by conducting an extensive literature review focused specifically on providing an understanding of five key topic areas: (1) definitions, operationalizations, and methodologies for measuring sexual harassment; (2) theories of gender and harassment; (3) racial and ethnic harassment in STEM; (4) intersectionality in experiences of harassment; and (5) the impact of harassment on both victims and institutions. This phase of research informed the topic selection and protocol development for the qualitative message boards, specific areas of focus for the survey review (e.g., considerations related to placement of a sexual harassment scale), and informed gaps and methods for exploration and discussion in the implementation plan. Specifically, the literature review informed additional constructs for consideration beyond sexual harassment (e.g., heterosexist harassment) that may be valuable to measure alongside sexual harassment. The literature review also enabled us to identify validated scales that researchers have used and/or adapted in a variety of contexts. It informed recommendations related to the most appropriate ways to measure sexual harassment and related constructs, particularly considering the target population. More information about the literature review is available in Literature Review: Sexual Harassment in the STEM Enterprise.

NCSES Survey Review

We assessed six NCSES surveys as potential vehicles for sexual harassment questions, leveraging knowledge gained from the literature review phase and specifically considering the number of potential questions that would be needed to measure sexual harassment and related constructs in STEM. Our review focused on the history and purpose of each survey, sample, design, and survey content. The survey review offers baseline information on current NCSES survey infrastructure for measuring sexual harassment. We incorporated elements of the survey review into our implementation plan, particularly using it to inform potential paths forward for NCSES in measuring sexual harassment. More information about the survey review and its findings is available in NSF Sexual Harassment Survey Review.

Qualitative Message Boards

As a third phase, we conducted qualitative research by leveraging online message boards to gain an understanding of the perspectives of STEM students and professionals, identify additional salient constructs of interest, and detail specific concerns related to fielding a sexual harassment survey. We directly incorporate findings from the message boards into the implementation plan. Specifically, findings from the message boards informed topics and populations for further exploration and provided direct recommendations related to fielding sexual harassment questions to the STEM population.

Implementation Plan

Leveraging the knowledge gathered from each of these three phases, we have designed an implementation plan that offers NCSES a nuanced understanding of how sexual harassment and related constructs function within the STEM workforce and education system. Our plan also includes survey options for measuring these topics, encompassing approaches that leverage existing NCSES surveys, and novel efforts. Additionally, the plan outlines the implications (e.g., strengths, weaknesses) and outcomes associated with these options.

This implementation plan discusses nine main considerations for measuring sexual harassment:

- Collaborator engagement
- Constructs of interest
- Coverage of populations of interest
- Sampling design
- Question development
- Fielding frequency and timing
- Mode of delivery
- Privacy concerns
- Communication material

With these considerations in mind, we identify three possible courses of action for NCSES to measure the incidence of sexual harassment in STEM.

Option 1: Field a New Survey

The first potential pathway is to develop an entirely new survey dedicated to understanding sexual harassment and related constructs in STEM. The central advantage of this option is the ability to design a methodology tailored specifically to meet NCSES's informational needs, including designing a questionnaire that covers a range of relevant constructs of interest. Additionally, fielding a new survey enables the optimization of the sample size and the implementation of oversampling where needed. Finally, a new survey can be conducted without disrupting existing survey efforts. However, practical constraints such as cost, time, and burden may be prohibitive.

Option 2: Field a Supplemental Survey

The second potential option is to field an off-cycle supplemental survey of sexual harassment using the National Training, Education, and Workforce Survey (NTEWS) and National Survey of College Graduates (NSCG). Although this option would be more costly than adding items to an existing survey (Option 3), it would require fewer resources than fielding an entirely new survey (Option 1). The added cost of using a supplemental survey comes with the considerable benefit of having the space to use a fully validated, behaviorally based scale of sexual harassment and explore multiple related constructs. Fielding a supplemental survey also allows for access to demographic information and covariates from the larger survey while reducing participant burden by not requiring respondents to answer duplicative questions. As with all the options provided, the addition of sensitive questions may impact respondents and affect response rates. Unlike Option 1, which can provide clear messaging that the survey will be about their personal experiences, respondents in Option 2 may be unprepared for the shift in subject matter. For example, since the host survey topics would be unrelated to sexual harassment and other constructs, the shift to these sensitive topics may cause participants to hesitate when responding to the survey more so than if it were on a stand-alone survey. However, leveraging a supplemental survey allows messaging to be crafted specifically to the supplemental questions to help prepare survey respondents for the potentially uncomfortable nature of the topics.

Option 3: Add Questions to an Existing Survey

The third potential avenue is to add questions that measure sexual harassment directly to an existing NCSES survey or surveys. Adding questions to an existing survey avoids some of the practical barriers associated with fielding a new survey; however, we anticipate costs related to data quality, such item nonresponse, with this course of action. Some techniques could mitigate these data-quality issues, such as placing items related to sexual harassment at the end of the survey and removing or moving questions that may induce context effects (e.g., a question related to military service near a question about an unwelcomed experience, such as hearing a sexual joke, might invoke the wrong context) earlier in the survey. Other data-quality considerations include ensuring that imputation, weighting, and item-nonresponse issues are assessed and planned for prior to including new items in a survey. Adding new questions could also cause disruptions to the selected established surveys. The limited space available in an existing survey may necessitate the selection of a fewer-than-optimal number of questions to evaluate sexual harassment (rather than a full validated measure) or limit examination to fewer constructs. We would recommend taking steps to ensure that the resulting survey with the additional questions maintains validity and that bias is diminished from the addition of the sexual harassment questions.

Based on considerations of our main objectives, the three potential options we identified, and knowledge gained from the three previous research phases, we provide three main recommendations in this implementation plan that are related to developing a measure of sexual harassment and related constructs in STEM:

1. Leverage behaviorally based validated scales.
2. Prioritize an intersectional approach to data gathering.
3. Field a new or supplemental survey.

These recommendations provide a strong foundation on which NCSES can build upon to understand the prevalence of sexual harassment within the STEM enterprise. Beyond these recommendations, we also describe several areas for continued research and exploration.

1 Chapter 1: Project Overview

Introduction

Central to the National Center for Science and Engineering Statistics' (NCSES) goals is the need to ensure a diverse and equitable workforce by identifying factors that contribute to or inhibit diversity and inclusion. A diverse STEM workforce is crucial to fostering innovation, broadening research agendas, engendering scientific breakthroughs, and increasing the public's benefit from STEM research (Cohen et al., 2002; Silver et al., 2019). Identifying barriers to entry and retention among underrepresented groups (e.g., women, sexual and gender minorities [SGM]) in STEM is crucial for ensuring these positive outcomes. Understanding these barriers and the extent to which certain groups are underrepresented requires high-quality data sources. Although NCSES currently collects valuable data used to understand the representation of women and racial minorities in STEM, these data do not measure the incidence of sexual harassment. Previous research has demonstrated that sexual harassment is not only a deterrent to women's entry into STEM careers, but also causes women to actively leave STEM (National Academies of Sciences, Engineering, and Medicine [NASEM], 2018; NASEM, 2020; Rodrigues & Clancy, 2020). In 2016, NASEM commissioned a study to examine the impact of sexual harassment within academia and found that sexual harassment occurs at all levels within academia and has negative, harmful impacts not only for the victim, but for those who witness the unwelcome behaviors as well (NASEM, 2018). The authors conclude that sexual harassment derails women's professional trajectories in myriad ways (e.g., increasing dissatisfaction at work, decreasing professional and educational opportunities, decreasing physical health and wellbeing; NASEM, 2018). Additionally, the authors found that as sexual harassment increases, so do women's intentions to leave the field (NASEM, 2018).

Further evidence for the exodus of women from STEM can be seen in the decline of the number of women graduating with STEM degrees compared to workforce participation. In 2020, women were awarded about half of the STEM degrees, yet they only comprise about 35% of the STEM workforce (NCSES, 2023), a phenomenon known as pipeline loss. Moreover, the extant literature indicates women with multiple marginalized identities (e.g., racial and ethnic minorities, SGMs) are at a greater risk for experiencing multiple forms of harassment, including sexual harassment (NASEM, 2018; NASEM, 2020). Despite efforts to diversify STEM, women, especially those who also hold other marginalized identities, remain underrepresented in the field relative to the general population (NCSES, 2023). Given the current labor shortages within STEM, NASEM notes that preventing sexual harassment within STEM is not only a moral and ethical imperative, but is also crucial to stop the continued loss of talented workers (NASEM, 2020).

Although there is significant evidence demonstrating the pernicious impacts of sexual harassment on women's participation in STEM, there are no high-quality, national data sets available to help further understand the extent and implications of sexual harassment in the STEM workforce (NASEM, 2018). Without these data, it is difficult to assess the extent to which sexual harassment exists within STEM, whether certain disciplines or professions in STEM are at particular risk, and to systematically examine the impacts of harassment on the STEM professional pipeline. As such, the collection of data to understand the incidence of sexual harassment is necessary to fully elucidate barriers to entry, advancement, and retention, and disparate career outcomes among women and other members from marginalized backgrounds (e.g., SGMs, racial and ethnic minorities).

Because sexual harassment is underreported through official pathways, prevalence estimates based on official reporting are not an accurate reflection of its occurrence (United States Government Accountability Office [GAO], 2020; Johnson et al., 2018; McCann et al., 2018; NASEM, 2020). A more reliable approach to measuring the incidence of sexual harassment is to leverage survey data. Currently, NCSES fields several surveys of STEM students and professionals, but does not currently include measures of sexual harassment. Additionally, membership in categories other than gender (e.g., race, gender identity, sexual orientation) is important to include to help understand which subgroups may be particularly vulnerable to

harassment. For example, research has found that women of color and LGBTQIA+ students in STEM face heightened levels of harassment, marginalization, and devaluation (Cech & Rothwell, 2018). As such, it is also important to consider measures of constructs beyond sexual harassment to fully understand how behaviors related to sexual harassment contribute to the exclusion of women and members of other marginalized groups in STEM.

To address this gap, we have conducted a three-phase research approach informing the fourth and final phase, the implementation plan, providing guidance to NCSES related to the measurement of sexual harassment and related constructs in STEM. The first and second phases consisted of a thorough literature review and a review of six surveys that NCSES currently fields.2 The third phase was qualitative message boards to understand the perspectives of STEM students and professionals and incorporate those perspectives into our recommendations for measuring sexual harassment. In this plan, we provide recommendations for developing survey measures of sexual harassment and related constructs in STEM.

Literature Review

We conducted a literature review, examining sexual harassment through both a legal and behavioral framework. The legal definition of sexual harassment focuses on sex-based discrimination, where an individual is subjected to unwanted behaviors based on their sex and/or their sexual orientation (Bostock v. Clayton County, 2020; US Equal Opportunity Commission, n.d.b.). Behavioral conceptualizations of sexual harassment center on observed behaviors regardless of their frequency (GAO, 2020). Although multiple behavioral definitions exist, the most frequently used definition is from Fitzgerald and colleagues (1988). This definition organizes sexually harassing behaviors into three categories: gender harassment, unwanted sexual attention, and sexual coercion. When compared to the legal definition, behaviors that constitute gender harassment and unwanted sexual attention are often grouped together as contributors to a hostile work environment, with sexual coercion mapping to quid pro quo. Often, researchers use the behavioral (versus legal) definition when measuring sexual harassment because it encompasses a larger spectrum of behaviors and more accurately estimates prevalence (Rospenda et al., 2009). Moreover, sexual harassment often co-occurs with other forms of inappropriate workplace behavior (e.g., incivility, sexually crude/offensive behavior; Konik & Cortina, 2008; Lee, 2018; Leskinen & Cortina, 2014), making it important to broaden our models and measures of harassment. The literature review identified behaviors that frequently co-occur with sexual harassment—such as workplace incivility, heterosexist harassment, and racialized sexual harassment—that NCSES may wish to measure in addition to sexual harassment to capture a more complete picture of the environment in which women and underrepresented groups in STEM work and study.

We also identified established, validated measures for sexual harassment (see Appendix A for summary of measures). Of the measures reviewed, the Sexual Experiences Questionnaire (SEQ; Fitzgerald et al., 1988), a behavioral measure, is the most commonly used measure and has been widely adapted to fit different workplace and academic environments. For example, the U.S. Department of Defense (DoD) worked with researchers to tailor the SEQ to the meet needs of the military workplace, resulting in the development of the SEQ-DoD (Fitzgerald et al., 1999). We provide the complete SEQ scale in Appendix A.

Survey Review

We assessed six NCSES surveys to determine the extent to which any may be appropriate vehicles for inclusion for sexual harassment questions. We first coordinated with NCSES to identify surveys for review. Surveys reviewed include both surveys of individuals and institutional surveys of the STEM workforce: National Survey Of College Graduates (NSCG), National Training, Education, and Workforce Survey (NTEWS), Survey of Earned Doctorates (SED), Survey of Graduate Students and Postdoctorates in

[2] Throughout this implementation plan, we refer to the earlier phases of this effort as the "literature review" and the "survey review."

Science and Engineering (GSS), and the Survey of Postdocs at Federally Funded Research and Development Centers.

The survey review drew from scientific literature on sensitive questions and survey methodology to establish the key consideration on which to evaluate NCSES surveys. Specifically, including sensitive questions on an existing survey can impact survey data quality. We also considered how these impacts can be mitigated through decisions related to survey administration. We then assessed NCSES surveys based on the considerations identified in the first part of the survey review for adding questions about sexual harassment. We reviewed the core components of each survey and offered recommendations for existing surveys that provide a potential fit for the addition of sexual harassment questions. Finally, we provided preliminary, high-level recommendations for measuring sexual harassment in STEM. Specifically, we identified two potential options: (1) commissioning a new survey; or (2) adding sexual harassment measures to one or more existing surveys.

Qualitative Message Boards

We conducted data through qualitative online message boards to inform topics for further exploration in the implementation plan, including perceptions of climate and culture, definitions of sexual harassment, and receptiveness to the inclusion of a proposed measure of sexual harassment in STEM on a survey.3 A third-party vendor recruited 177 message board participants who were organized based on their status (i.e., student or professional) and their sexual and gender identity (i.e., cisgender, straight women; cisgender, straight men; or sexual or gender minority). As shown in Table 1, we fielded 12 message boards over the course of three weeks in the fall of 2022. Participant demographics can be found in Tables B2–B4 in Appendix B.

Each board remained open for a week (Monday through Friday). Appendix B provides a description of the message board methodology, and Appendix C provides the protocol. Participants responded to prompts ranging from the climate and culture of their organization (e.g., “What are some positives that come to mind when you think of your primary academic department’s/workplace’s climate and culture?”), social identities (e.g., “How important are social identities in how people experience the culture in your work division or workplace?”), and respectful behaviors (e.g., “What does respect look like at your workplace or work division/primary academic department?”). They responded to questions related to sexual harassment, including what behaviors they consider to be sexual harassment and locations they consider to be high risk for harassment. Of particular importance for the implementation plan, participants shared their views on how they would feel about responding to questions regarding sexual harassment in STEM. The insights generated from participants informed the identification of additional constructs of interest, target populations, considerations related to the timing of survey fielding, and potential barriers to responding to questions about sexual harassment. The information gained from the message boards has been incorporated throughout this plan.

Table 1. Overview of the Message Board Sample Coverage

Women Boards	One Undergraduate STEM Students Board
	One Graduate STEM Students Board
	One Early-Career STEM Professionals Board (1–10 Years Post-Terminal Degree)
	One Mid-Career STEM Professionals Board (11–20 Years Post-Terminal Degree)
	One Late-Career STEM Professionals Board (20+ Years Post-Terminal Degree)
Men Boards	One Undergraduate STEM Students Board
	One Graduate STEM Students Board
	One Early-Career STEM Professionals Board (1–10 Years Post-Terminal Degree)
	One Mid-Career STEM Professionals Board (11–20 Years Post-Terminal Degree)

[3] We had originally planned to conduct online focus groups with STEM students and professionals as well as interviews with late-career professional women. However, due to circumstances that emerged over the course of this effort, it became necessary to pivot our qualitative data collection approach.

	One Late-Career STEM Professionals (20+ Years Post-Terminal Degree)
Sexual and Gender Minority Boards	One STEM Undergraduate and Graduate Board Students (Mixed Gender Groups)
	One STEM Early-, Mid- and Late-Career Board Professionals (Mixed Gender Groups)

In the next section, we provide an overview of the remainder of the implementation plan.

Chapter 2: Recommended Constructs of Interest

In this chapter, we define the constructs of interest to consider when measuring sexual harassment from the literature review and qualitative message boards. We describe the four core constructs to assess: sexual harassment, reporting of sexual harassment, heterosexist harassment, and racialized sexual harassment. We then describe an additional construct for consideration—technology-facilitated sexual violence.

Chapter 3: Key Considerations for All Survey Approaches

In this chapter, we describe our eight key considerations for NCSES. For each consideration, we provide our recommendations based on our previous phases' findings. These considerations are applicable to all three approaches and should be taken into consideration regardless of the option selected.

Chapter 4: Option 1 – Field a New Survey

In this chapter, we describe the first option, fielding a new survey, for measuring sexual harassment and related constructs in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations.

Chapter 5: Option 2 – Field a Supplemental Survey

In this chapter, we describe the second option, field a supplemental survey, for measuring sexual harassment and related constructs in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations.

Chapter 6: Option 3 – Add to an Existing Survey

In this chapter, we describe the third option, add measure to an existing survey, for measuring sexual harassment and related constructs in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations.

Chapter 7: Conclusion

In this chapter, we summarize our main recommendations for measuring sexual harassment in STEM and describe opportunities for future research aimed at further developing an approach to measuring workplace and gender relations within STEM.

2 Chapter 2: Recommended Constructs of Interest

In this chapter, we describe the main constructs that we propose measuring. Before designing a survey, a crucial first step is to identify what information needs to be obtained and then from whom that information needs to be gathered. Identifying the constructs to be measured will influence the selection of the population of interest, question development, the mode of survey delivery, and production of communication materials. As such, we have several recommendations related to constructs to measure, regardless of selected survey approach.

The proposed constructs are drawn from both the literature review phase and the qualitative message boards phase. Based on the literature review and the qualitative message boards, we identified four main constructs we would recommend measuring: (1) sexual harassment; (2) reporting of sexual harassment; (3) heterosexist harassment; and (4) racialized sexual harassment.4 We recommend the inclusion of reporting of sexual harassment to understand the general reporting climate and the potential impacts to reporting. We recommend heterosexist harassment and racialized sexual harassment in addition to sexual harassment since they often co-occur with sexual harassment and may provide additional context to the environments in which sexual harassment is prevalent (Konik & Cortina, 2008; Lee, 2018; Leskinen & Cortina, 2014). Last, we identified an additional construct for consideration through the qualitative message boards: technology-facilitated sexual violence. For each proposed construct, we provide a description of the construct and support for these constructs from our literature review and qualitative message boards.

Sexual Harassment

To understand the extent to which sexual harassment exists in STEM, we recommend measuring sexual harassment through sexually harassing behaviors. By sexually harassing behaviors, we mean those behaviors falling under the classifications of gender harassment, unwanted sexual attention, and sexual coercion (Fitzgerald et al., 1998). As highlighted in the literature review, it is important to use behaviorally based measures to capture a range of experiences and behaviors that qualify as sexual harassment even if a respondent would not identify them as such (Government Accountability Office [GAO], 2020; NASEM, 2020). In fact, scales like the Sexual Experiences Questionnaire (SEQ) are designed specifically to ensure that researchers are not relying on a respondent to identify a sexually harassing behavior as sexual harassment, but rather identifying their experiences with the behavior itself.

The need to understand sexual harassment through behaviorally based measures also guided the discussion prompts we developed for the qualitative message board data collection. As part of the qualitative message boards, we wanted to understand whether participants would identify additional behaviors they considered to be sexual harassment beyond the ones typically measured in existing, validated sexual harassment scales. To assess this, participants reviewed a list of sexually harassing and heterosexist behaviors that may occur in their workplace or primary academic department. The list of behaviors was drawn from our literature review and existing scales to include behaviors consistent with unwanted sexual attention, gender harassment, and sexual coercion (which is consistent with quid pro quo). Additionally, we drew behaviors consistent with heterosexist harassment (see Konik and Cortina, 2008) to be consistent with a broader understanding of workplace harassment (detailed information related to these definitions is provided in the literature review; we further describe findings related to heterosexist

[4] We considered gender discrimination as a construct of interest. However, researchers have not coalesced around an agreed-upon definition of gender discrimination, as is found with sexual harassment, leading to variations in measures of gender discrimination (de la Torre-Pérez et al., 2022).

harassment in an upcoming section). Table 2 provides the list of sexually harassing behaviors we provided to message board participants.

Table 2. List of Sexually Harassing and Heterosexist Behaviors

Telling sexual jokes
Touching someone
Making sexual gestures
Making comments based on gender
Sharing sexual pictures or videos of themselves
Sharing sexual pictures or videos of other people
Making sexual advances
Talking about their sexual activity
Asking about sexual activity or preferences
Making repeated attempts to establish a romantic or sexual relationship with the same person
Telling offensive jokes about sexual or gender minorities
Using homophobic names or slurs
Sharing homophobic materials in your office

After participants reviewed the list, they were then asked to identify any behaviors not included that they may consider sexual harassment. The goal of this exercise was to learn whether there were additional behaviors that emerged as salient to include in a measurement beyond the typical sexually harassing behaviors already considered. Responses showed that most message board participants found the list to be comprehensive. However, some participants identified staring and intrusions on personal space as additional behaviors. It is worth noting, these behaviors are accounted for in the SEQ's Gender Harassment: Sexual Hostility subscale (Appendix A provides the complete SEQ).[5] Thus, based on findings from both the literature review and the qualitative message boards, we recommend the National Center for Science and Engineering Statistics (NCSES) use a behaviorally based validated scale, such as the SEQ, for measuring sexual harassment on a survey. We discuss this recommendation in more detail in Chapter 3.

Reporting of Sexual Harassment

Beyond measuring experiences of harassment, it would be valuable to collect information related to reporting behaviors and knowledge of the reporting process. Climates in which sexual harassment thrives may also be indicative of environments in which the likelihood to report such experiences is suppressed and the likelihood of negative ramifications for reporting is (or is perceived to be) increased (Hart, 2019). Although most cases of sexual harassment go unreported (McCann et al., 2018), when people do report, they indicate experiencing negative consequences such as retaliation, hostility, and ostracism (Knapp et al., 1997; McCann et al., 2018). Further, reporting sexual harassment can result in diminished career opportunities (Hart, 2016; Hart, 2019). Because sexual harassment tends to occur in environments shared by the perpetrator and victim, who also often know one another, victims often try to ignore the situation and may not report their experiences to avoid confrontation, continued interactions, or conflict, particularly with the harasser (Knapp et al., 1997; Wasti & Cortina, 2002).

[5] The SEQ includes two subscales related to gender harassment—Gender Harassment: Sexist Hostility (i.e., discriminatory treatment because of one's sex/gender) and Gender Harassment: Sexual Hostility (i.e., unwanted sexual behavior targeting someone based on their sex/gender, but not to force sexual compliance; Fitzgerald et al., 1999). The unwanted behavior described by message board participants fits within the Sexual Hostility subscale and would already be captured if using a scale like the SEQ.

Assessing the disclosure of sexual harassment in the professional realm often relies on reporting behaviors to the individual company or organization, which limits the availability of measures and data related to knowledge of the reporting process, barriers to reporting sexual harassment, and the rate of disclosure. Although organization-wide surveys on reporting sexual harassment exist within specific contexts, such as assessments within the U.S. Department of Defense (DoD), these measures may be situationally limited.[6] Campus climate surveys also tend to include items assessing whether victims of sexual harassment (or sexual misconduct more broadly) reported their experiences, to whom they reported them, and how the university is perceived to react to such reports (e.g., Cantor et al., 2020; Krebs et al., 2016). The Association of American Universities (AAU) includes questions assessing students' knowledge of university resources available to those experiencing sexual assault and sexual misconduct, as well as questions related to reporting sexual misconduct, including sexual harassment (Appendix A provides item wording and response options; Cantor et al, 2020).[7] Although the examples here are limited to the academic environment, understanding working professionals' experiences with reporting may also be crucial for identifying the impact of sexual harassment on women's and underrepresented minorities' career trajectories in STEM. Thus, we recommend that items assessing reporting behaviors and experiences be included on NCSES's measure of sexual harassment for both student and professional populations. We recognize that, given the potential circumstance-specific nature of reporting questions, specific questions that measure reporting may need to be adapted from existing surveys or developed for certain contexts.

Heterosexist Harassment

To further understand all experiences related to sexual harassment in STEM, we also recommend measuring heterosexist behaviors in the survey. Heterosexist behaviors are hostile actions or comments that reinforce traditional gender norms and degrade sexual and gender minorities (SGM) based on their *perceived* sexual orientation (Konik & Cortina, 2008; Silverschanz et al., 2008). In their three-factor model of workplace harassment, Konick and Cortina (2008) conceptualize heterosexist harassment as a separate—but related—construct to sexualized harassment and gender harassment (see the literature review for a more detailed description of their framework). Similar to gender harassment, heterosexist behaviors are rooted in ideas of traditional gender norms. Heterosexist behaviors can target any individual who appears to be deviating from the perceived societal norms for their gender (e.g., male nurses) (Konik & Cortina, 2008; Silverschanz et al., 2008). These behaviors can be direct (e.g., use of a homophobic slur) or ambient (e.g., overhearing an anti-gay joke at work) and provide insight into the climate and organization (Konik & Cortina, 2008; Silverschanz et al., 2008; Waldo, 1999). Measures such as the Workplace Heterosexist Experiences Questionnaire (Waldo, 1999) capture harmful behaviors (e.g., being denied career advancement opportunities due to your sexual orientation) that measures of sexually harassing behaviors do not include. As SGMs are at an increased risk for sexual harassment, it is important to capture their unique experiences (Richey et al., 2019; NASEM, 2020).

The qualitative message boards phase specifically explored heterosexist behaviors in addition to the sexually harassing behaviors included on the behaviors list (Table 2). For example, SGM participants mentioned invalidating a person based on their sexual orientation and/or gender, such as not using the correct pronouns, "outing" a person to others (e.g., peers, coworkers), and joking or making comments about a person's sexuality as additional sexually harassing and heterosexist behaviors. It is noteworthy that it was predominately SGM participants who made these observations

[6] The *Workplace and Gender Relations Survey of Military Members* assesses if sexual harassment was disclosed, along with the type of disclosure which occurred (e.g., disclosed to DoD authority, disclosed unofficially to leadership).

[7] Although measures assessing the reporting of sexual harassment are more limited, measures assessing sexual assault reporting are more prevalent (Krebes et al., 2016; Follingstad et al., 2020; Cantor et al., 2020) and may be useful for adapting measures of reporting sexual harassment.

related to additional behaviors beyond those included in the list we provided. This suggests that SGM STEM students and professionals do in fact have a unique experience with regard to harassing behaviors that is not necessarily captured by existing scales of sexual harassment. Additionally, although SGM message board participants were quick to identify heterosexist harassment as a form of sexual harassment, cisgender and heterosexual participants often did not. Most often, participants in the cisgender and heterosexual boards did not independently mention any heterosexist behaviors. Whether participants labeled heterosexist harassment as harassment depended—at least to some degree—on the identity of the participant, and we might expect similar patterns in response to survey questions.

There is very little research on the prevalence of heterosexist harassment in STEM, making it difficult to define its broader impact on the field (Marín-Spiotta et al., 2023; Richey et al, 2019). Because heterosexist behaviors are not immediately recognized as harassment—particularly by cisgender, heterosexual students and professionals—it is important to use behaviorally based measures that do not require respondents to consciously label a behavior as harassment to effectively evaluate the prevalence of heterosexist behaviors. Because heterosexist harassment is situated within the broader workplace harassment framework and there is a need to further understand the intersections between sexual and heterosexist harassment in STEM, we recommend including behavioral measures of heterosexist harassment on future surveys.

Racialized Sexual Harassment

Given the underrepresentation of women and racial and ethnic minorities in STEM, we also suggest exploring how gender and race/ethnicity intersect through the measurement of racialized sexual harassment, as both constructs influence experiences of sexual harassment (Buchanan, 2005a; Buchanan et al., 2018; Cho, 1997). Racialized sexual harassment provides a framework to discuss the interplay of gender and race/ethnicity in experiences of victimization and perpetration (Buchanan, 2005a; Buchanan et al., 2018; Cho, 1997). As discussed in the literature review, one such measurement capturing this intersection is the Racialized Sexual Harassment Scale (RSHS), a behaviorally based measure capturing sexual harassment, racial harassment, and racialized harassment (Buchanan, 2016). Given that women from racially and ethnically diverse backgrounds are at greater risk of experiencing sexual harassment, measuring the occurrence of racial and sexual harassment will provide important context to the experience of people from such backgrounds (Beal, 2008; Bowleg et al., 2003; Buchanan, 2016).

Based on the findings from the qualitative message boards, individuals may not typically label experiences of racialized sexual harassment as such. Only two participants specifically listed examples of racialized sexual harassment, such as racist comments/jokes about a person's body and sexualizing people based on their race/ethnicity. However, when discussing issues of workplace climate and respect more broadly, nine message board participants reported experiencing or witnessing differences in treatment based on race/ethnicity.

As there are no current estimates for the occurrence of racial and ethnic harassment or racialized sexual harassment in STEM, including a measure of these behaviors in NCSES's chosen approach to measuring sexual harassment would be useful in understanding their prevalence. Due to the known co-occurrence of sexual and racial/ethnic harassment in the general population (Buchanan et al., 2018; Buchanan & Fitzgerald, 2008), the inclusion of racialized sexual harassment measures will allow for a more nuanced understanding of the experiences of all individuals within the STEM community and will potentially capture actionable behaviors that impact the retention of women in STEM. Developing this understanding will be central to the attraction and retention of diverse talent within STEM. As such, we recommend including a measure of racialized sexual harassment alongside measures of sexual harassment among STEM students and professionals.

Technology-Facilitated Sexual Violence (TFSV)

Social media and the online environment were consistently listed as high risk for sexually harassing behaviors during our qualitative message boards. As such, including questions on harassment that occurs online may provide more information about the types of harassing experiences students and professionals in STEM experience, especially given the shift to the remote environment. For example, young adults (i.e., 18–24 years old), are at a greater risk for TFSV victimization, and more understanding of its prevalence and impact on STEM students could indicate where interventions are needed (Powell & Henry, 2019). Given that both students and professionals across professional experience levels identified TFSV as being of particular concern, it may be an important construct to consider when measuring sexual harassment in STEM.

The literature on online harassment is disparate, with a variety of terms, definitions, and measures (e.g., cyber sexual harassment, digital dating abuse, cyberviolence; Backe et al., 2018). We elected to use the term TFSV because it is a broad term that encompasses experiences of harassment or abuse that occurs online (e.g., on social media) or through the use of technology (e.g., smartphones; Powell & Henry, 2019). These experiences include digital sexual harassment (i.e., use of electronics to engage in unwanted and/or unwelcomed sexual behavior), image-based sexual abuse (i.e., the nonconsensual sharing of sexual media), sexual aggression and coercion, and gender and sexuality-based harassment (i.e., derogatory comments based on a person's gender and/or sexuality). The harassing behavior may be public (e.g., derogatory comments in a public, online space) or private (e.g., constant texting after work hours). Although these actions can occur in person, the anonymity and disinhibition in an online environment may encourage these behaviors, a phenomenon documented in the literature and reported by message board participants (Zhong, Kebbell, & Webster, 2020).

The anonymity provided by the internet may contribute to its perception as a high-risk environment. Across all message boards, participants listed social media and the online environment more broadly as a high-risk location. For example, one late-career male STEM professional said, "The location I consider the highest risk for these [sexually harassing] behaviors is social media because people feel like they can do anything on social media without consequences." Message board participants cited anonymity and a lack of oversight and consequences as reasons why they believed the online environment is high risk.

A survey seeking to understand sexual harassment should consider including questions to specifically understand how sexual harassment manifests—both in person and online. Assessing TFSV allows for broader consideration of the environments and mode thorough which harassment occurs. Given the ubiquity of technology in modern life, particularly the professional and academic environments, understanding these nuances will provide important and actionable information.

Summary

Although we recommend capturing a wide range of related constructs to understand a more complete picture of the sexual harassment landscape impacting participating in the STEM field, particularly among women and members of marginalized groups, we understand that capturing this breadth of information would require the use of a significant number of survey questions. Depending on the number of constructs to be measured, we recommend leveraging modules where there are a core set of questions (e.g., related to experiences of sexual harassment and reporting) and then assigning other constructs to separate modules. To reduce survey burden and ensure no one respondent is receiving an excessive number of sensitive questions, we recommend that no one participant receive more than two modules.

Beyond considerations related to selection of constructs to measure, NCSES will decide what questions to ask to measure these constructs. We recommend leveraging behaviorally based,

validated scales. However, we also recognize that survey questions may need to be updated to reflect the needs of a specific data collection effort or target population. In Chapter 3, we further discuss considerations related to question development.

In the next chapter, we discuss several potential approaches NCSES could leverage to measure sexual harassment and related constructs among STEM students and professionals. Our recommendations related to survey options are—in part—driven by our proposed constructs and measurements of interest (i.e., behaviorally based scales).

3 Chapter 3: Considerations for Survey Measures of Sexual Harassment and Related Constructs

This chapter provides an overview of nine key considerations on which we evaluate the three survey approaches in Chapter 4. The nine considerations are interconnected, and therefore, some information across these approaches may overlap or contain similar information. Each of these considerations, and ultimately the decisions implemented for measuring sexual harassment in STEM, will impact survey and item completion and, in turn, overall data quality. In the next section, we describe the nine considerations and provide our overarching recommendation for the approach. Table 3 describes each consideration, provides a brief summary of each consideration, and details the overarching approaches we recommend that should be applied regardless of the survey option selected.

Table 3. Key Considerations Overview

Consideration	Summary	Overarching Approach
Collaborator Engagement	Describes potential opportunities to engage with key collaborators on the development of the survey to engender buy-in from relevant entities and raise the profile of the survey to increase participation	• Identify organizations and entities that would yield meaningful partnerships in the final development of a survey or survey measures related to sexual harassment and related constructs in STEM. • Identify strategic points through survey development at which to engage them and gather their feedback. • Leverage these partnerships to proactively address potential concerns or criticisms of the developing approach.
Constructs of Interest	Describes considerations related to identification of sexual harassment constructs of interest as well as related constructs to measure	• Prioritize an intersectional approach when assessing sexual harassment and the reporting of sexual harassment by including measures assessing heterosexist harassment, racialized sexual harassment, and the reporting of these experiences.
Coverage of Populations of Interest	Captures information related to the target population for the survey for which we are interested in assessing	• Target both STEM students and professionals at various stages (i.e., undergraduate, graduate, early-, mid-, late-career professionals) • Sampling considerations should also account for targeting women, sexual and gender minorities (SGM), and racial and ethnic minorities to ensure appropriate coverage of underrepresented groups in STEM.
Sampling Design	Provides information related to the process for sample creation	• The sufficient sample size needed for analyses should be determined prior to survey administration to ensure that there is sufficient power to detect effects and achieve desired precision while avoiding costly oversampling.

Question Development	Describes considerations related to measurement and scale selection and adaption, recall period, response levels, question ordering and modular design, and assessment (cognitive interviews and pilot testing)	• Use established and validated behaviorally based measures of sexual harassment and related constructs. • Recall periods for experiences of interest should be determined based on the fielding frequency. • Constructs should be assessed in separate modules, and survey respondents should receive no more than two modules. • Consideration of order effects is crucial when including sensitive questions in the survey. • Pilot testing and cognitive interviews should be employed to ensure accuracy and reliability of newly fielded survey questions.
Fielding Frequency and Timing	Captures information on how often data should be collected via the survey and the frequency of survey fielding	• Surveys should be designed to support longitudinal data collection, if possible. • The fielding frequency should consider how quickly experiences are expected to change within the population.
Mode of Delivery	Provides considerations for the method of survey administration	• Web- or paper-based surveys should be used due to the sensitive nature of questions. Phone-assisted surveys should *not* be used. • The survey should be sent by an independent party and not by the school or employer.
Privacy Concerns	Captures information on privacy concerns, privacy policies and materials related to the survey, and considerations for the sampling plan and public release data	• The privacy policy should be developed early in the process and be clearly communicated to all staff working on the survey. • The survey materials provided to participants should contain a plain-language explanation of the safeguards that are in place and the rights of the participants. • Any sensitive questions should only be assessed via self-administered surveys. • Key subgroups should be well represented in the sample to reduce the chances of small cell sizes or unique demographic combinations that often increase disclosure risk.
Communication Materials	Describes information related to general communication strategy, such as pre-survey outreach materials and supplemental materials, developing language about the survey itself related not only to privacy, but also to purpose and impact; partnering with other entities	• Incentives may be considered to improve survey response. • Clear messaging about the purpose, impact, and use of survey as well as privacy concerns. • Consider partnerships with related organizations to raise profile of survey and understand and mitigate any potential concerns

Key Approach Considerations

Collaborator Engagement

Early engagement at organizations or events dedicated to preventing sexual harassment, such as the National Academies of Science, Engineering, and Medicine (NASEM) annual public summit or the public summit on preventing sexual harassment in higher education, could assist with identifying potential collaborators who can be consulted on the development of the proposed survey, as well as help uplift this overall effort. Identifying other collaborators (e.g., Title IV offices, like-minded organizations) and working with them throughout the process of survey development is crucial to the overall success of the project (Mahoney et al., 2022; Merhill et al., 2021), as the can provide valuable feedback and insight on survey development, results, and communication strategies, and can increase visibility for the survey. Additionally, proactively seeking out the feedback of those who may have a vested interest in this data collection can allow the National Center for Science and Engineering Statistics (NCSES) to address and mitigate any concerns they may have about the data-collection effort itself or the data collected. Furthermore, it can ensure that a broad range of perspectives are actively incorporated into survey development. As such, regardless of the final survey approach selected, we recommend conducting an environmental scan of potential collaborators who may be able to serve as vocal supporters of this effort, and then establish specific time points at which they will be engaged in the survey development process (e.g., after construct selection, to provide feedback on proposed questions).

Collaborators could also provide quotes about the importance and relevancy of the survey that could be used in a press release. For this effort, we have identified three potential groups of collaborators:

> 1) Organizations (non-profits, government etc.) that work on issues with sexual harassment and gender/sexuality in STEM at a national, state, or international level (e.g., National Academies Committee on Women in Science, Engineering, and Medicine; Association for Women in Science; American Association for University Women; Society of Women Engineers);
>
> 2) University administrators, especially those who handle issues regarding gender and harassment (e.g., university Title IX offices); and
>
> 3) Students, faculty, and professionals in STEM fields who are the target population of the survey. This final group can also include academic researchers who study topics related to sexual harassment in STEM and therefore might have particular interest in the methods NCSES is using.

A comprehensive environmental scan of potential collaborators would allow for the identification of additional groups of collaborators as well as specific organizations or entities within the three groups above.

We believe it could be beneficial to begin outreach to these collaborators prior to final survey creation to determine common goals across parties. This could mean including organizations such as the National Academies Committee on Women in Science, Engineering, and Medicine in discussions about measure selection, as well as seeking feedback on questions and planned fielding methods. Collaborators could also be involved during the pilot testing and cognitive interviewing phase of survey development to get more detailed feedback on items and/or measures NCSES plans to include and to potentially help with recruitment for participants. Finally, collaborators can also help with tailoring messaging about the survey and raising its profile among target populations. Buy-in from these organizations (and mitigating potential criticism from them) can help bolster the survey's credibility, particularly to groups who may be initially hesitant to take it.

RECOMMENDATIONS

- Identify organizations and entities that would yield meaningful partnerships in the final development of a survey or survey measures related to sexual harassment and related constructs in STEM.
- Identify strategic points through survey development at which to engage them and gather their feedback.
- Leverage these partnerships to proactively address potential concerns or criticisms of the developing approach.

Constructs of Interest

The primary consideration in survey design involves identifying constructs to assess sexual harassment within the population of interest. As discussed in detail in Chapter 2, we recommend assessing four constructs: sexual harassment, reporting of sexual harassment, heterosexist harassment, and racialized sexual harassment. We recommend the inclusion of sexual harassment reporting to understand the reporting climate across the STEM enterprise and help identify potential barriers to reporting. We recommend heterosexist harassment and racialized sexual harassment in addition to sexual harassment because they often co-occur with sexual harassment (Konik & Cortina, 2008; Lee, 2018; Leskinen & Cortina, 2014). Given that experiences with sexual harassment vary across social identities (Beal, 2008; Bowleg et al., 2003), we recommend including measures that examine harassment within and across minority populations that may be at higher risk of experiencing sexual harassment. By examining sexual harassment through an intersectional lens, we can attempt to understand individuals in STEM who may have unique and nuanced experiences due to their social identities. If of interest to NCSES, the qualitative message boards phase also identified technology-facilitated sexual violence as an additional construct to measure. We provide additional context for each of the four recommended constructs and an additional construct in Chapter 2.

RECOMMENDATIONS

- Prioritize an intersectional approach when assessing sexual harassment and the reporting of sexual harassment by including measures assessing heterosexist harassment, racialized sexual harassment, and the reporting of these experiences.

Coverage of Populations of Interests

Once constructs to measure are identified, it is crucial to select the population in which to assess these problematic behaviors. When selecting the population of interest, it is important to examine the current survey coverage of survey populations (i.e., current undergraduate and graduate students pursuing a STEM degree and STEM professionals) and identify gaps that may exist. To fully assess the extent to which sexual harassment exists within STEM, it's important to have coverage of both the STEM students who are actively seeking degrees and the professionals in the field.

Given that victimization rates and types have been found to vary across subgroups such as SGMs and racial minorities (see the Theories Related to Harassment and the Impact of Sexual Harassment sections of the literature review), it is crucial to have a sampling design that allows for the experiences of a diversity of subgroups to be represented in the final results. The absence of respondents in specific subgroups can result in an inability to estimate prevalence rates for these groups with statistical precision, ultimately limiting the extent to which sexual harassment can be identified, understood, and prevented equitably (e.g., Krebs et al., 2016). Depending on the subgroups of interest to NCSES, such as SGMs, racial minorities, undergraduate versus graduate student statuses, or career stages, careful consideration in the construction of the sampling frame for the sexual harassment instrument and expanded demographic data (e.g., to capture gender identities and sexual orientations more comprehensively and accurately; Suen et al., 2020) may be necessary.

The coverage we propose would allow for a comprehensive understanding of the prevalence and nature of these behaviors and would provide a more accurate picture of the extent of the problem and its impact on individuals and the field as a whole. Assessing individuals across all levels of the workforce and potential workforce (i.e., students) and across various demographic groups can help to identify the barriers that prevent individuals from reporting these behaviors and potentially shed light on pipeline loss due to problematic behaviors. It also enables the identification of patterns of behavior, such as whether certain types of behavior are more common at certain levels of the workforce, within specific STEM fields, or within specific demographic groups (i.e., behaviors targeted at women or SGMs).

RECOMMENDATIONS

- Target both STEM students and professionals at various stages (i.e., undergraduate, graduate, early-, mid-, late-career professionals).
- Sampling considerations should also address the risks women, SGMs, and racial and ethnic minorities face to ensure appropriate coverage of underrepresented groups in STEM.

Sampling Design

When developing the sampling design, it is important to consider sample size. If the sample is too small, the survey data may lack the power necessary for inference, rendering it less useful. Conversely, an excessively large sample would expend unnecessary resources without providing benefits beyond those allotted by using an adequate sample size. Within each of the survey options, it will be important to consider the necessary sample size to reliably assess the constructs of interest. Prior to data collection, determining a sufficient sample size can be informed by factors such as how the data will be used (e.g., planned comparisons, cross-sectional versus longitudinal analyses), the survey design selected (e.g., constructs assessed via modules), the desired level of oversampling of targeted subgroups, expected rates of survey nonresponse, and expected rates of item nonresponse (i.e., item missingness).

As with all surveys, selecting an appropriate sample size is crucial for success. Standard approaches to optimizing sample sizes based on desired precision for subgroup estimates and power for multivariate models will serve as a helpful starting point in sample design. However, given the sensitive nature of the survey questions, the design will need to account for the likelihood that item nonresponse will be higher than usual (Saewyc et al., 2004).

Sample sizes should be increased in line with expected item nonresponse rates to maintain the utility of the survey data for these questions. As discussed in the survey review, sensitive questions tend to have higher item nonresponse rates than questions that are not perceived to be about sensitive topics. Most of the items on the proposed survey will measure constructs that are sensitive to at least some portion of the sample. Different individuals will find different items to be sensitive depending on their background and experiences, making it unlikely that missing data will be completely at random and the underlying mechanisms that lead to item nonresponse will have some degree of variation across items. This makes it crucial to develop a comprehensive plan to address and minimize the biases that could arise on this survey, likely including the use of sophisticated imputation methods, as it will be more important here than on a survey with fewer sensitive items.

Due to the expected non-random nature of the item nonresponse, imputation will likely be necessary to reduce the bias of survey estimates. The imputation methods will increase standard errors, and these increases will need to be accounted for in any power analyses or sample optimization analyses. Additionally, the sampling plan will need to account for any modularity of the questionnaire.

RECOMMENDATION

- The sufficient sample size needed for analyses should be determined prior to survey administration to ensure that there is sufficient power to detect effects and have desired precision for key survey estimates while avoiding costly oversampling.

Question Development

Once the constructs of interests and target population have been selected, it is important to select survey items, scales, and measures to be included within the proposed surveys. There are several considerations related to selection and development of survey questions measuring sexual harassment and other related constructs.

Behaviorally based questions. To assess the prevalence of sexual harassment within the STEM field, we recommend the use of an established and validated behaviorally based measure of sexual harassment, such as the Sexual Experiences Questionnaire (SEQ), which is aligns with the recommendation from NASEM (Merhill et al., 2021). This recommendation is essential to establishing the reliability and validity of the survey results. Relying on a behaviorally based measure does not require the respondent to know the definition of sexual harassment, nor does it require them to make a subjective judgment regarding the alignment of their experiences with such a definition. This approach reduces bias in participants' responses and tends to increase estimated prevalence rates (Government Accountability Office [GAO], 2020), which may be more representative of the actual occurrence of sexually harassing behaviors (Johnson et al., 2018). However, items that ask respondents to label their experiences of sexual harassment (e.g., ask them if they have experienced sexual harassment; Fitzgerald et al., 1997) may be useful *in combination* with behaviorally based items to understand how the victim's perception of their experience relates to their experienced outcomes.

Most current measures of sexual harassment are behaviorally based. Complete questionnaires, such as the SEQ, have undergone testing to ensure that survey responses are measuring the intended construct in a reliable way across subgroups. If a behaviorally based measure is not used, any questions added to the survey will need to be scrutinized for measurement invariance and undergo standard measurement development practices to ensure reliability and validity.

In Chapter 2, we also suggest additional measures to assess the reporting of sexual harassment, heterosexist harassment, and racialized sexual harassment. Although we generally suggest not deviating from identified measures of sexual harassment (e.g., not adding or removing questions that may disrupt the overall validity or reliability of the measure), responses on the message boards and literature review suggested several additional constructs to measure beyond sexual harassment. For example, message board participants indicated that it may be of use to include a measure assessing technology-facilitated sexual violence due to the high reported rates of sexual harassment on social media platforms and within online environments. Therefore, it may be of interest to seek out subscales that assess these additional constructs of interest. Appendix A details some existing scales of sexual harassment as well as these related constructs.

Gendered language. Although sexual harassment, and in particular gender harassment, are concepts that are contextualized by social identity dynamics, explicitly referring to one gender identity in a sexual harassment measure (e.g., "Been treated as if you are more vulnerable than men" or "Denied a raise… because you are a woman"; Klonoff & Landrine, 1995; Kuchynka et al., 2018) may not accurately capture the specific dynamics at play depending on the respondents' own gender identity. Such language may suppress the number of experiences captured by the measure and may be exclusive of respondents' actual gender identities. Gender-neutral items may be recommended over those that apply gendered language to respondents' experiences. Although we recommend leveraging existing, validated scales, in some cases, updates to question wording will be necessary to ensure questions are inclusive to all potential respondents.

Legal criteria. It may be necessary to identify cases that meet certain legal criteria; however, leveraging measures that align solely with cases that meet an established legal criterion can result in a lower prevalence rate that may not reflect the frequency with which sexually harassing behaviors tend to occur. If the ability to identify experiences that align with a legal definition of sexual harassment is of interest, we recommend using follow-up items representing legal criteria with a measure that aligns with a behavioral definition (e.g., Cantor et al., 2020; Breslin et al., 2019).

Recall period. The measures reviewed in the literature review and recommended here have been used to examine sexual harassment experiences within a variety of recall periods, from as recent as the previous 12 months to as broad as "ever" (e.g., Gutek et al., 2004). Generally, assessing the previous 12 months compared to a longer recall period may improve respondents' ability to accurately remember and report their experiences (Krebs et al., 2016) and may best allow for trending over time. However, the recommended recall period may be dependent on how frequently the measure is administered, when it is administered, and if trend analyses are of interest. As such, decisions related to the recall period may also be influenced by the fielding frequency (i.e., annually, biannually) of the survey on which sexual harassment and related questions are placed.

Response options. Leveraging existing scales where possible will provide responses options accompanying existing questions. However, if adapting questions, designing response options will have important impacts on the quality of data collected. The benefits of options that represent the frequency of the behavior (e.g., never, three or four times per month) versus binary "yes" and "no" options may depend on the specific goal of the measure. However, an important practice to consider is avoiding "select all that apply" response options. A "yes/no" grid is recommended to better ensure that the measure comprehensively captures all potential experiences for all respondents by providing a clear indicator of which behaviors they had experienced and which they had not (Krebs et al., 2016). Further, open-ended response options may be useful when gathering exploratory information but may not be recommended when developing a measure of prevalence of behaviors given potential inconsistencies in defining and applying criteria across participant responses.

Question ordering. Following selection of measurements and scales, question ordering will need to be decided. The decision for question ordering differs based on the approach selected (see Chapters 4–6). However, given the sensitive nature of these questions, it is important to consider how question ordering may impact item completion regardless of the option selected. The order of questions on a questionnaire can influence the lens through which someone interprets future questions. It is important to be mindful of how sensitive questions on a questionnaire can have downstream effects, and potentially disrupt the data collection of not only the measure of interest for this effort, but potentially other measures that may be included in the survey.

Additionally, the order in which the behaviors of interest are assessed may be important for accurately capturing and categorizing the types of sexual harassment experienced by students and working professionals (e.g., Abbey et al., 2005). For example, Krebs et al. (2016) found that participants who had experienced sexual harassment responded affirmatively to items assessing unwanted sexual contact (which led to the sexual assault battery) when those items were presented first, resulting in a potential mis-categorization of their experiences. These measures were reordered so that the sexual harassment measure appeared first, ensuring that experiences with sexual harassment were not included in the measure of sexual assault.

In general, we recommend keeping existing scale measurements in their original ordering and providing clear language explaining the purpose of the question set. Additionally, we suggest that each measure (i.e., sexual harassment, heterosexist harassment, and racialized sexual harassment) be in separate modules and that participants receive no more than two modules to prevent respondent burden and maintain response quality. Using this design, paired with a large enough sample size to support the division of constructs across respondents, inferences for each construct and the relationship between constructs can be made. Pilot testing and cognitive interviewing methods will be helpful in identifying instances of bias from priming and/or order effects as well as to understand respondents' perceptions of the sexual harassment measure.

If NCSES chooses to place sexual harassment questions directly on an existing survey, a particular concern is the proximity of these questions to non-sexual harassment questions. The sexual harassment measures reviewed in the literature review tend to be included on relatively shorter surveys within studies focused on understanding the antecedents or outcomes of sexual harassment (e.g., Fitzgerald et al., 1997) or on larger campus climate surveys of which the focal content is conceptually related to sexual harassment (e.g., Cantor et al., 2020). Given the sensitive nature of sexual harassment, measures of such experiences included on surveys collecting less sensitive, but still crucial information for NCSES's purposes may impact item response rates on such surveys depending on their proximity to one another. It may also impact the overall survey response rates—if the content of the survey must be disclosed to potential respondents, the inclusion of a sexual harassment measure on a survey with otherwise less sensitive content (e.g., demographics, education) may decrease individuals' willingness to participate. Thus, if placing questions directly on an existing survey, comprehensive testing is recommended to assess the extent of such impacts and inform decisions surrounding whether NCSES's sexual harassment measure and other related measures are best integrated with a current survey instrument or developed as their own survey.

Cognitive interview and pilot testing. Moreover, it is essential to conduct cognitive interviews, pilot tests, and pretesting on any new questions, modified scales, and on general question ordering prior to inclusion in the final survey. This step is crucial for the identification of potential problems with the questions, such as confusing wording or response options, and can also give an indication of whether the questions and/or the survey are measuring what they intend to measure. Given the sensitivity of the constructs of interest, it is also important to ensure that questions within the surveys are clear and respectful and that they do not cause distress or discomfort to survey respondents. If new questions or questions outside of a validated measure are to be used, cognitive interviews are recommended. Cognitive interviews are especially useful in identifying any potential cultural barriers due to the language used in the questions, and therefore, it is an important step in any survey implementation to ensure that respondents are understanding the potential questions for inclusion in the survey. This process can help identify any issues with comprehension, interpretation, or response accuracy. It is recommended to conduct cognitive interviews with a large enough sample size to ensure that any potential problems are identified and addressed.

RECOMMENDATIONS

- Use established and validated behaviorally based measures of sexual harassment and related constructs.
- Recall periods for experiences of interest should be determined based on the informational needs of NCSES and the survey fielding frequency.
- Constructs should be assessed in separate modules, and survey respondents should receive no more than two modules.
- Consideration of order effects is crucial when including sensitive questions in the survey.
- Pilot testing and cognitive interviews should be employed to ensure accuracy and reliability of newly fielded survey questions.

Fielding Frequency and Timing

After survey items have been assessed and order has been determined, the next step in survey design relates to the frequency of survey fielding over time. This consideration includes identifying a survey approach to allow for tracking over time, how often the survey should be fielded to facilitate data collection, and the time frame of interest for understanding these experiences.

First, when evaluating survey options, it is important to consider longitudinal aspects of the analysis that will follow. The survey should be administered in a way that facilitates longitudinal tracking of the experiences of both the overall population and key subgroups. This includes asking questions that are consistent over time and that capture the same types of experiences, behaviors, and attitudes.

By doing so, it becomes possible to identify changes in the incidence or nature of experiences of harassment. Ideally, all important subgroups will be identified prior to the first survey. However, if analysis of the first wave identifies some groups that warrant further study and larger sample sizes, the sample can be increased for that group in subsequent years, if possible, under the survey option. Thus, we recommend implementing either cross-sectional and longitudinal or panel designs to track changes in the STEM student and professional population over time.

A panel approach would allow for sampling of the same respondents over time, which can enable the study of causal relationships. However, there are some drawbacks to panel surveys, such as the likelihood that bias increases over time due to panel attrition and the costs of administration. If attrition rates are expected to be high or the primary concern for analysis is point-in-time survey estimates, then a cross-sectional data collection effort may be preferable. Due to the sensitivity of the survey topic, attrition rates may be higher than other NCSES panel surveys, making this of greater concern than usual. Incentives paid to panel members to encourage continued participation are a common tool to combat panel attrition, but they can significantly increase the cost of data collection. To further address attrition problems, longitudinal and cross-sectional design approaches can be blended, such that a panel is maintained for longitudinal data collection and new participants are added to the sample pool each time the survey is fielded to replenish the panel. This also ensures that newer population members, such as those who aged into the population of interest since the previous iteration, are captured in survey estimates.

Second, the frequency at which the survey is fielded should be determined (e.g., annually, biannually). The fielding frequency should be determined by balancing cost considerations against two other concerns: how quickly experiences are expected to change within the population, and the needs of policy makers and other stakeholders. Other NCSES surveys like the National Survey of College Graduates (NSCG) and the National Training, Education, and Workforce Survey (NTEWS) field biennially. That schedule could be appropriate for the proposed survey if changes in professionals' and students' experiences are expected to be gradual. However, it is preferable to maintain some degree of flexibility in the fielding schedule if possible. If events occur that could significantly change the experiences of STEM students or professionals, such as the implementation of a new, relevant policy, it may be important to adjust fielding timelines to provide up-to-date information for stakeholders and policy makers (e.g., fielding a survey annually).

Similarly, the timing of other surveys which respondents might participate in should be taken into consideration to ensure that the fielding of multiple surveys supported by NCSES do not overlap. Message board participants' opinions about the ideal time frame for fielding the proposed survey diverged. Neither students nor professionals identified a single time of year or time in the semester that would be preferred. Participants on the professional boards noted that it is important to avoid holidays and the summer because people tend to be busier or on vacation during those timeframes.

Last, as described in the Question Development section of this chapter, fielding frequency will likely influence the recall period assessed in the measure.

RECOMMENDATIONS

- Surveys should be designed to support longitudinal data collection, if possible.
- The fielding frequency should consider how quickly experiences are expected to change within the population.

Mode of Delivery

In selecting a survey option, NCSES should assess how the survey will be collected and how respondents will be contacted. Selection of survey mode will impact data quality in a number of ways and impact respondents' perspectives on the level of privacy afforded to them in responding to this type of survey.

Currently, NCSES uses a trimodal data-collection approach with both the NSCG and NTEWS: a self-administered web survey; a self-administered mailed, paper survey; and a computer-assisted telephone interview. Given the sensitive nature of these questions, we recommend implementing these surveys only via self-administered web or paper surveys. We *do not recommend* computer-assisted telephone interviews, as this may compromise feelings of anonymity and discourage participation.[8] The ability to distribute self-administered surveys over online platforms provides an additional layer of privacy for respondents that may be crucial when examining sensitive topics such as sexual harassment. For example, cognitive interview participants in the development of Krebs and colleagues' (2016) campus climate survey reported that asking victims to provide a description of their experience may be upsetting and may cause respondents to drop off. In addition to limiting the number of follow-up items that are asked about each experience (Krebs et al., 2016) and providing relevant support resources for respondents who have experienced sexual misconduct at the end of the survey (Cantor et al., 2020; Krebs et al., 2016; Swartout et al., 2019), allowing participants to complete the survey in the location of their choosing using a web-based instrument may further reduce any discomfort or distress that could arise.

In fielding a self-administered web or paper survey, NCSES should also consider the method for contacting participants for taking the survey. Message board participants were asked how they would like to be contacted to take a survey that assesses sensitive questions, and most student and professional message board participants indicated that email outreach would be the best way to contact them. In terms of whether to send email outreach to personal or official (work, school) email accounts, there was less agreement. Participants considered their work or school accounts to be trusted sources, and seeing a survey arrive to those accounts may be seen with more legitimacy. Several participants indicated that sending the survey to their personal accounts may increase the likelihood that they would delete it, ignore it, or dismiss it as spam. Other message board participants listed privacy concerns as the primary reason why they would prefer the proposed survey be sent to a personal email address. This finding may indicate where additional, clear messaging is needed regarding privacy policies governing data collection and use. Additionally, concerns regarding privacy appear to be especially salient for SGM individuals, as the majority of participants on both the SGM students' and professionals' boards stated a preference to receive the proposed survey at a personal email address. This observation may suggest that additional research is needed to more fully consider the best forms of outreach to SGM participants to ensure their comfort in participating in a survey of this nature.

Regardless of where the survey link is delivered, the survey link itself should be provided by an independent third party such as a government agency. It is important that recipients of the survey know that the survey is not being provided by their school or employer. Providing the link from an independent third party will highlight the independent nature of the data collection and emphasize confidentiality.

RECOMMENDATIONS

- Web- or paper-based surveys should be used due to the sensitive nature of questions. Phone-assisted surveys should not be used.
- The survey should be sent by an independent party and not by the school or employer.

[8] Findings from our qualitative research indicated that although some participants were open to receiving phone calls on this topic, not all were. Many participants stated it would be stressful for them to be asked such questions over the phone. In particular, SGM message board participants were more likely in comparison to non-SGM participants to express hesitance and discomfort at the idea of responding to questions related to sexual experiences and sexuality over the phone.

Privacy Concerns

Prior to fielding, privacy policies and infrastructure should be created, updated, or assessed to better reflect the collection of sensitive information. Due to the extremely sensitive nature of these survey questions, privacy considerations should receive extra focus. Specific attention should be given to the privacy policies, the reporting of results, and the release of survey data sets. It is important to develop a practical plan for how data and results can be shared with interested groups in a way that maintains the privacy and confidence of respondents. This plan should be developed early in the process and clearly communicated to all staff working on the survey.

Issues of privacy were mentioned in response to three separate questions about hesitancy to take a survey, trustworthy sponsors for such a survey, and how comfortable respondents felt answering questions of a sensitive nature. This highlights the need to carefully develop privacy policies and clearly communicate those policies to survey participants. Whenever sensitive questions are asked in a survey, researchers have a specific responsibility to ensure that information cannot be traced back to participants to ensure their safety. If potential participants perceive this as even a faint possibility, trust in the survey will be reduced. Concerns that a third party might be able to trace participants' responses back to them will be particularly salient for those who have experienced sexual harassment, causing them to feel less comfortable answering such questions and skip these questions at higher rates than other participants. Therefore, survey materials provided to participants should contain a plain language explanation of the safeguards that are in place and the rights of the participants. As discussed in the survey review, any questions on these topics should only be asked on self-administered surveys. Self-administered surveys have better perceptions of privacy by sample members. This enhanced sense of privacy will lead to fewer refusals and more honest responses, which will yield better data quality (McNeeley, 2012).

Another consideration is the distribution of public release data and whether aggregate or individual-level data should undergo heightened scrutiny. Given the sensitive nature of responses to questions about sexual harassment, the balance of utility and privacy should tilt more strongly toward privacy. Some of these concerns can be ameliorated with a well-designed sampling plan. Ensuring that key subgroups are well represented in the sample can reduce the chances of small cell sizes or unique demographic combinations that often increase disclosure risk. Given the sensitive nature of these questions and potential impact on survey respondents and their responses, we recommend that NCSES consider using this opportunity to review their data-sharing policies to identify any areas that could potentially be updated and improved. To further bolster NCSES's already robust system of data management, it may be beneficial to conduct an environmental scan of how other government agencies (e.g., the U.S. Department of Defense [DoD]) manage similarly sensitive data and handle messaging regarding data management and sharing policies to potential respondents. These data management practices from other organizations could then potentially be incorporated into NCSES's strategies for these types of sensitive questions. It would also be beneficial to clearly communicate data-sharing policies in messaging to potential survey respondents to increase response rates overall and reduce nonresponse bias. A clear, accessible explanation of the standard legal protections and authorizations, such as Title V or Title XIII, will be an important component of this effort.

As such, methods to safeguard participant privacy should be described in survey outreach materials when fielding the survey. For example, NCSES should consider providing plain language describing how individual responses will not be linked directly to their identity in a single data source (e.g., mailing address will not be stored in the same data file as survey responses), the limited number of people who will have direct access to raw survey responses, and other strategies that will be employed to ensure that responses will not be directly traced back to respondents. Additionally, language should clearly describe that academic institutions and employers will not directly receive these data, and therefore, students and employees should not worry that their employers will know their responses to the survey. Finally, language should describe the disclosure avoidance strategies that will be implemented before any data are released to ensure that data cannot be triangulated back to an individual based on survey responses and sample frame information.

RECOMMENDATIONS

- The privacy policy should be developed early in the process and be clearly communicated to all staff working on the survey.
- The survey materials provided to participants should contain a plain-language explanation of the safeguards that are in place and the rights of the participants.
- Any sensitive questions should only be assessed via self-administered surveys.
- Key subgroups should be well represented in the sample to reduce the chances of small cell sizes or unique demographic combinations that often increase disclosure risk.

Communication Materials

Given the sensitive nature of measuring sexual harassment and related constructs, we recommend the development of a robust communication plan for the survey and how results will be used to help encourage participation. When message board participants were asked about factors that would influence the likelihood of response to questions about sexual harassment in STEM, one common sentiment that emerged was a desire to know how the survey results would be used. In fact, a motivator many participants identified, regardless of educational or career stage, for wanting to participate in a survey measuring sexual harassment was the positive impact they believed it could have on understanding and improving outcomes of sexual harassment. This repeated importance placed on the impact of such a survey suggests that proactively considering and articulating how results may be used could positively impact survey participation among the target population. Given that NCSES is a policy-neutral agency, collaborating with offices and researchers who may leverage these data could be important to anticipate and address questions regarding how survey results will be used to positively impact the STEM field.

Additionally, certain groups of potential respondents may benefit from enhanced outreach to encourage participation. Krebs and colleagues (2016) suggest that men's participation in their campus climate survey may have been lower compared to women's participation due to their feelings that a survey of experiences with sexual misconduct may not be relevant to them. Given the more strongly established history of sexual harassment research targeted at women's experiences versus those of men and individuals who identify as non-binary, gender fluid, and/or trans, recruitment materials that are targeted at different subgroups and that emphasize the importance of all individuals' participation may be recommended (Krebs et al., 2016). Further, the stated purpose of a survey and the organization who recruits for and administers a survey with sexual harassment items may impact individuals' responses (e.g., surveys framed as more general and less specific to sexual harassment may result in lower recall of relevant experiences; Galesic & Tourangeau, 2007).

We recommend conducting additional research in support of developing a strategic communications plan to both raise the profile of the survey measuring sexual harassment and to proactively address questions or concerns potential respondents may have. Conducting additional qualitative research is one way to inform the development of this plan. Additionally, engaging with collaborators (as identified in the "Collaborator Engagement" section of this chapter) could be instrumental in developing a messaging campaign for the survey measures. Preliminarily, we would recommend avoiding any framing that indicates that the survey is focused on sexual harassment or detailed descriptions of other unwanted workplace behavior that may be measured, as that could impact individuals' willingness to take the survey. Messaging for the survey should adequately prepare potential respondents for the sensitive nature of the questions but should avoid directly labeling the survey as a sexual harassment survey.

We recommend that NCSES develop a communications strategy to both enhance survey and/or item response rates, but also to anticipate any potential media (including social media) attention the survey, survey items, or its results may engender. Based on our experience with fielding surveys of

a sensitive and highly visible nature, we recommend exploring the following options, outlined in Table 4, as part of a communication strategy plan.

Table 4. Outline of Communication Strategy Plan

General communication strategy	• Have a cohesive communication strategy across the organization to prevent misunderstandings or misinformation as the survey is fielded. • Provide talking points to leaders or employees within NCSES who will likely be making public statements, giving talks at conferences, or taking media interviews about the survey. • Talking points can be centered on: motivation for the survey, sample population, fielding/recruitment strategies, who the data will be available to and when, privacy protections, and details about how data and results of the survey will be used and communicated. • Develop target population-specific materials to encourage participation and address concerns specific to that population (e.g., students, professionals, men, women, etc. may have unique concerns or questions related to survey participation).
FAQs	• Create a FAQ webpage, which is accessible to a wide audience with anticipated questions and their responses. • Anticipate questions about survey items that are outside of the realm of NCSES purview (e.g., how results will inform policy) and prepare appropriate ways to address these questions or point people to additional resources. • Establish a primary contact who can respond to additional questions from potential respondents and provide contact information or a contact form along with the FAQ.
Social Media	• Develop social media toolkits with useful graphics to communicate information about the survey. • Provide recommended language to use when posting about the survey to disseminate to anyone who might be interested in posting about the survey online.
Press	• Prepare press releases ahead of survey release. • Monitor current news stories and ongoing policy changes or debates in the area of sexual harassment and related constructs in STEM, which might bring additional attention to the survey fielding and results.

These items are not an exhaustive list of communication strategy elements, but rather are intended to provide initial ideas for NCSES to consider as they begin to plan survey development. Further research to inform the communications strategy is necessary as well as potentially leveraging collaborator input. For example, collaborators can help identify areas of concern that need to be addressed and questions that should be included in the FAQ list. Collaborators could also provide quotes about the importance and relevancy of the survey that could be used in a press release. Using clear communication of results and intended next steps will be important steps in negating concerns that the data is not being used to effect change and in combating any impressions that issues related to sexual harassment in STEM are being ignored.

RECOMMENDATIONS

- Conduct additional research to inform and develop a strategic communications plan ahead of fielding a survey with items measuring sexual harassment and related constructs.
- Partner with collaborators to inform the communications plan and engage in strategic communications for the survey, its purpose, and how gathered data will be used.

Summary

This chapter has described the main considerations on which we assess each of the following three possible avenues to measure sexual harassment and related constructs in STEM. Each of these considerations is described broadly in this chapter, but specific considerations are described in the following chapters. The structure of each of the following chapters largely mirrors the order of the considerations described in this chapter. The next chapter describes considerations related to fielding a new survey.

4 Chapter 4: Option 1 – Field a New Survey

In this chapter, we describe the first option for measuring sexual harassment and related constructs in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations. The structure of this chapter mirrors that of Chapter 3. Chapter 3 provides the overarching considerations for measuring sexual harassment and related constructs on a survey, and this chapter describes these same considerations as they relate specifically to fielding a new survey. As described in Chapter 3, each of the considerations are interconnected and, in many cases, decisions related to one consideration will impact decisions related to other considerations.

Overview

For this approach, the National Center for Science and Engineering Statistics (NCSES) would commission an entirely new survey designed specifically to measure the experiences of sexual harassment and related constructs among STEM students and professionals. Although a new survey is expected to be time intensive and costly, it is also the option with the greatest amount of control over the constructs measured and data collection procedures, leading to a methodology tailored to the specific needs of the data collection effort. There was also a clear preference for a new survey from message board participants, with about 68% (*n* = 109) of message board participants indicating that they would prefer a stand-alone survey over responding to sexual harassment measures as part of a larger survey. Message board participants commonly reported that a separate survey dedicated to measuring sexual harassment would emphasize the importance of the issue and prevent survey fatigue. Table 5 provides an overview of the specific approaches for Option 1.

Table 5. Option 1 Overview

Consideration	Approach
Collaborator Engagement	• Identify organizations and entities that would yield meaningful partnerships in the final development of a survey related to sexual harassment and related constructs in STEM. • Identify strategic points through survey development at which to engage them and gather their feedback. • Leverage these partnerships to proactively address potential concerns or criticisms of the developing approach.
Constructs of Interest	• Prioritize an intersectional approach when assessing sexual harassment and the reporting of sexual harassment by including measures assessing heterosexist harassment, racialized sexual harassment, and the reporting of these experiences.
Coverage of Populations of Interests	• Target both STEM students and professionals. • Oversample members of underrepresented groups (women and racial and ethnic minorities). ○ Currently, there is no mechanism to directly oversample sexual and gender minority members.
Sampling Design	• Leverage existing sampling frames to aid in the sampling design, such as the American Community Survey (ACS). • Consider over- or under-sampling subgroups as needed.
Question Development	• Use established and validated behaviorally based measures of sexual harassment and related constructs.

	• Recall periods for experiences of interest should be determined based on the fielding frequency. • Constructs should be assessed in separate modules, and survey respondents should receive no more than two modules. • Pilot testing and cognitive interviews should be employed to ensure accuracy and reliability of newly fielded survey questions.
Fielding Frequency and Timing	• Frequency of fielding and timing of administration should be determined based on informational needs. • The new survey should not overlap with other NCSES surveys. • Surveys should be designed to support longitudinal data collection, if possible. • A panel approach would help increase sample size for cross-sectional analyses.
Mode of Delivery	• Surveys should be self-administered web- or paper-based surveys. • Surveys should not be administered via telephone. • Data collection could mirror NCSES's current web- and paper-based survey approach.
Privacy Concerns	• Develop clear privacy policies for survey data collection and data holding. ○ This could be based upon existing NCSES privacy policies. • Clear communication on privacy measures should be included in the survey and outreach materials. • Evaluation of existing data infrastructure may be helpful to ensure security of data.
Communication Materials	• Convey clear messaging about the purpose, impact, and use of survey before, during, and following survey administration. • Partner with organizations to raise awareness of the survey and potentially improve trust among survey respondents. • Develop new communication materials specifically in support of the new survey effort.

Approach Considerations

Collaborator Engagement

In developing an entirely new survey to measure sensitive workplace experiences, we recommend conducting additional research, such as an environmental scan, to identify potential collaborators to coordinate with on survey development—particularly to ensure that key constructs and target populations are included and to develop outreach materials to bolster participation. Additionally, a survey measuring sexual harassment and related constructs in STEM will likely garner widespread attention. Collaborating with interested parties will help ensure the attention is largely positive and that potential concerns or criticisms are addressed in the development of the survey.

We recommend coordinating with identified collaborators through survey development to incorporate their input into key decisions. NCSES could develop a strategy to determine where their input is most needed and engage them at crucial junctures in survey development. Incorporating their feedback will demonstrate NCSES's dedication to collecting data to combat sexual harassment in STEM. In some circumstances, their feedback will not be able to be incorporated. Having engaged them early in the process, however, allows for NCSES to proactively communicate practical limitations to incorporating certain pieces of feedback and to potentially troubleshoot solutions.

Implementation Recommendation: As a first step in standing up a new survey, we recommend identifying and engaging key collaborators and developing a plan to strategically incorporate their feedback into survey development.

Constructs of Interest

For this approach, the entire survey can be devoted to measuring sexual harassment and related constructs. Designing a new survey offers a great deal of flexibility and the space for inclusion of several constructs of interest. As such, all four recommended constructs (as described in Chapter 2 and in the Constructs of Interest section of Chapter 3), could be included on the new survey. Further, if space is available, additional measures of interest, such as technology-facilitated sexual violence (TFSV) could be included. We recommend the same approach as outlined in Chapter 2: prioritizing an intersectional approach when assessing sexual harassment by including measures assessing heterosexist harassment and racialized sexual harassment as well as TFSV.

Implementation Recommendation: In coordination with key collaborators, we recommend that NCSES first identify the constructs of interests that should be measured on a stand-alone survey. As described above, we recommend measuring and reporting sexual harassment and several additional behaviors that have been shown to co-occur with sexual harassment. Given the number of constructs recommended, it may be necessary to prioritize certain constructs, yet to fully understand the interconnected nature of these experiences and how they disproportionately impact underrepresented groups in STEM, collecting a wider range of harassing experiences will be central to future data analyses to understand these workplace cultures and climates and their impact on the STEM workforce.

Coverage of Populations of Interest

Although current NCSES surveys broadly capture the recommended STEM population (i.e., STEM students and STEM professionals), there may be gaps in specific, vulnerable populations that a new survey fielding could mitigate. For example, existing surveys, such as the National Survey of College Graduates (NSCG) and the National Training, Education, and Workforce Survey (NTEWS), would be good candidates for covering the population of interest (see Survey Options 2 and 3). However, the NSCG does not oversample members of underrepresented groups, such as women and racial and ethnic minorities, and the extent of oversampling of these groups on the NTEWS has not yet been published. As women and racial and ethnic minorities are already underrepresented in the STEM field, oversampling these groups may be necessary to better capture the experiences of these groups—which is especially important because members of these groups are more likely to experience sexual harassment (Beal, 2008; Bowleg et al., 2003). By fielding a new survey, the degree of needed oversampling of these vulnerable populations can be considered prior to survey administration, offering better representation of the targeted population. Additionally, NCSES surveys do not currently ask about sexual or gender minority status, thus this approach allows for the opportunity to include these questions in the survey.

Implementation Recommendation: We recommend identifying the population of interest from which information should be gathered. Specifically, we recommend measuring experiences of sexual harassment and related constructs among STEM students and professionals to fully understand career progression and pipeline loss. For professionals, we recommend sampling from all career levels and from a wide variety of backgrounds (e.g., skilled technical workforce). For students, we recommend including current and recently graduated undergraduate and graduate levels. Collecting data from students is particularly important to understand pipeline loss and potentially identify crucial time points for intervention to help retain talent in the STEM workforce. Additionally, we recommend specifically focusing on measuring these experiences among underrepresented groups such as sexual and gender minorities (SGM) and racial and ethnic minorities. Determining which populations are of most interest to gather data will guide many of the subsequent survey development steps. If possible, questions related to sexual orientation and gender identity should be added to understand the prevalence of these unwanted behaviors within this population.

Sampling Design

As with all surveys, it is expected that sample size determinations will be made dependent on the sampling design selected. A new survey has the advantage of optimizing sample size requirements, given the selected sampling design of the survey. This approach could leverage existing sampling

frames to aid in the sampling design. For example, some NCSES surveys of STEM students and professionals source their sample frames from the ACS. The ACS, or another large Census survey, would be an appropriate frame source for a new survey fielding as well. To be successful, the frame must enable accurate identification of members of the target population and allow for over- or under-sampling of different subgroups as needed. Important factors for over- or under-sampling may include gender, race or ethnicity, age, field of study, level of education, and stage of career. If these are not measured in the surveys used to generate the sample frame, we recommend NCSES discuss the possibility of adding these items with the U.S. Census Bureau. This adds the benefit of controlling the level of oversampling of subgroups of interest, which can provide greater precision in survey estimates.

Implementation Recommendation: We recommend focusing on sampling strategies that will allow for oversampling under-represented populations in STEM. As a baseline for sampling design, existing sampling frames can be used, such as the ACS.

Question Development

Similar to the key considerations discussed for Question Development in Chapter 3, this approach would use existing and validated behaviorally based measures of sexual harassment and related constructs. This approach affords the opportunity to include multiple scales of interest on a single survey, which will enable the examination of relationships between key constructs, extend the utility of the data for modelling, and enhance the understanding of STEM students' and professionals' experiences. Specifically, a new survey would have the space to include multiple validated scales without removing or modifying any items, which may compromise scale validity.[9] This helps ensure the quality of data collected is reliable and enables the examination of relationships between key constructs, extend the utility of the data for modelling, and enhance the understanding of STEM students' and professionals' experiences.

Although the entire survey can be devoted to measuring sexual harassment and related constructs, that does not suggest that all constructs of interest could or should be placed into a single survey. Rather, we recommend using a modular design. In a modular questionnaire design, the data would be collected for all scales of interest. However, individual sample members would only receive a subset of the scales. The subsets of scales assigned to participants are determined through random assignment. This means that different participants receive different combinations of scales, covering all possible permutations. This design ensures that no individual is overwhelmed with too many sensitive items or an excessively long questionnaire while facilitating the study of the interrelationships between different scales. The degree to which individuals may feel overwhelmed by the sensitive nature of the survey can be assessed using cognitive interviewing and pilot testing.

Implementation Recommendation: We recommend leveraging validated behaviorally based scales for each of the identified constructs. We have provided several examples of behaviorally based scales from previous research in Appendix A. The scales provided are not an exhaustive list of available scales, but rather focus on the scales discussed in the literature review. Therefore, as a next step, we recommend conducting an environmental scan of existing validated scales for each of the identified constructs of interest to measure and then selecting scales based on NCSES's informational needs and the target populations. This will allow for the selection of scales to measure constructs of interest.

After establishing which scales will serve as the primary foundation for survey questions, it may be necessary to adapt these questions to some degree to ensure they are appropriate for the intended audience. For example, some questions may need wording updates to reflect professional or student

[9] Minor wording modifications to tailor survey questions to the target population is potentially advisable. For example, making minor edits to ensure that question wording is gender neutral and is inclusive of gender non-conforming populations would be potentially appropriate. As Merhill et al. (2021) notes, in cases where the target population is not captured in an item's original wording, retaining the original wording may harm item or scale validity more than making minor wording edits (Merhill et al., 2021).

status. We recommend making minimal wording changes to existing scales (as opposed to sweeping changes to question wording, question order, or the number of questions for each scale), documenting those changes from the original scale, and providing a justification for the change (Merhill et al., 2021). Changes in scales may not be uniform and may need to be adapted slightly differently depending on target population (e.g., students, professionals).

As an additional step to survey instrument development, item ordering will need to be decided, keeping in mind considerations related to item response. As outlined in the Question Development section of Chapter 3, the order of questions on a questionnaire can influence the lens in which someone interprets future questions. Therefore, given the sensitive nature of the topics described in the stand-alone survey, it is important to be mindful of which constructs are presented first and how sections are prefaced for respondents. With the aim to alleviate potential respondent burden and maintain response quality, measures should be implemented as modules where participants receive no more than two modules to prevent respondent burden and maintain response quality.

Once the survey instrument is developed, conducting cognitive interviews and pilot and pretesting will be necessary. Even using existing, validated scales, it will be necessary to conduct cognitive interviews to further adapt question wording in ways to ensure that it resonates with the target populations as the order of constructs and collection of demographic variables which may influence survey responses. Additionally, changes to question wording from the original scales may impact how respondents interpret questions, and cognitive interviews will help NCSES more fully understand the potential impacts of question wording on responses. Since this survey would include multiple constructs of interest, cognitive interviews can help determine the most appropriate order for respondents to see questions or receive modules. Similarly, questions will need to be pilot tested to examine the validity of each question.

To ensure that order effects and the assessment of important demographic covariates do not bias survey responses, sufficient sample sizes should be used to conduct both the pilot testing and the cognitive interviews. The sample size for the pilot testing and cognitive interviewing will need to be large enough to make inferences about the target population and should include subpopulations of interest. Additionally, consideration should be given to the representation of women, racial and ethnic minorities, and SGMs during pilot testing and cognitive interviews to ensure that these populations are sufficiently covered with respect to the planned sampling of these populations within the general survey design. As these groups are the most vulnerable to sexual harassment and are underrepresented in STEM, deliberate efforts (e.g., oversampling) should be given to ensure they are sufficiently represented during the testing phase. This will ensure subpopulations of interest are sufficiently covered with respect to the planned sampling within the general survey design. The final task in this step would be to review findings from the cognitive interviews and the pilot and pretesting and to update the survey accordingly.

Fielding Frequency and Timing

Similar to the key consideration for Fielding Frequency and Timing in Chapter 3, the frequency of fielding and the timing of the survey should be determined based on NCSES and collaboration needs. Additionally, the new survey should not overlap with other surveys and should be administered during a time that is beneficial for both students and professionals, such as in the spring or fall (but avoiding holidays).

Establishing a new survey ensures it will measure sexual harassment and related constructs in consistent, reliable ways over time. It is important to consider longitudinal aspects of the analysis that will follow in survey design (for more information on this consideration, please refer to the Fielding Frequency and Timing section of Chapter 3). In addition to the considerations outlined in Chapter 3, it is important to consider the expected attrition rates of the survey to determine if sample sizes should be increased due to expected attrition. Therefore, in designing a new survey, it is important to consider whether the survey should be implemented over time with a panel design to increase sample size for cross-sectional analysis. Said differently, to increase sample sizes for measurements of sexually harassing behaviors, it may be useful to implement a panel design where

new individuals could be surveyed at each administration, allowing for increased power in cross-sectional analyses. This approach would be similar to the NSCG and NTEWS.

Implementation Recommendation: The frequency of fielding (e.g., annual, biannual) and timing of fielding should be determined, largely based on the overall goals of NCSES. However, we do recommend the survey fielding to not overlap with the fielding of existing NCSES surveys. If interested in trending experiences of sexual harassment over time, we recommend incorporating a panel approach for data collection.

Mode of Delivery

This option's approach mirrors the key consideration approach described in Chapter 3. Given the sensitive nature of the questions, we recommend implementing the new survey only via self-administered web or paper surveys. However, we do not recommend computer-assisted telephone interviews, as this may compromise feelings of anonymity and discourage participation. The survey materials, including the link for the survey, should be emailed directly from NCSES, or another independent party, and not by the respondents' school or employer.

Implementation Recommendation: Given the sensitive nature of the stand-alone survey, we recommend implementing the new survey via self-administered web or paper surveys and not through computer-assisted telephone interviews.

Privacy Concerns and Communication Material

Given that this is a new survey, new documentation related to privacy practices will need to be created. Unlike an existing survey, which may have established protocols, a new survey covering sensitive topics will need clear documentation to ensure confidence in the respondents. Privacy materials will need to describe the purpose of the survey, why it was created as a stand-alone survey, how data will be handled to ensure anonymity, and outline guidelines related to the release of data. In addition to clear communication within the survey and during outreach, it is important to test the infrastructure where the data will be collected and housed to ensure there are no security breaches or potential leaks. It is worth noting that the new survey's privacy materials could be based upon existing surveys' privacy policies and can be housed in a similar location, but due to the sensitive nature, extra precautions should be taken.

In addition to the key considerations surrounding Communication Material outlined in Chapter 3, communications surrounding a new survey will be crucial. Since this is a new survey, materials explaining the purpose and impact of this survey distributed prior to, during, and following survey fielding have the potential to bolster response rates and interest in the survey. Moreover, coordination with collaborators will be essential in raising awareness and potentially improving trust among survey respondents.

Implementation Recommendation: Communication about the survey effort should be clearly developed both to support its fielding (e.g., outreach materials when fielding the survey) and to share with the target population ahead of fielding the survey. This communication should clearly articulate the safeguards in place to protect the survey data once collected and ensure that there are no negative repercussions for survey respondents. Given the sensitive nature of the topic, a robust privacy policy will be necessary to ensure STEM students and professionals are comfortable participating in the survey and answering questions honestly about their experiences.

Additionally, a robust communications plan will help encourage STEM students and professionals to participate in the survey. We recommend leveraging the qualitative data gathered as part of this effort to begin developing a communications strategy based on the concerns and questions raised by message board respondents. This information can also be used to develop outreach materials. Additionally, we recommend NCSES consider sharing information related to the survey purpose with collaborators or through their website well ahead of fielding the survey to raise their awareness of the upcoming survey and to garner interest. A proactive communication and outreach campaign would be intended to assuage concerns about data privacy, to alert STEM students and professionals about the impending survey, and to potentially increase the likelihood of survey

participation. Based on qualitative message boards, we recommend highlighting the potential impact that a survey of this nature would have as well as assuring respondents that their responses will not be shared with their home institution or employer and that responses will not be linked to them. A robust communications campaign can be further informed by collecting additional qualitative data to speak directly to both potential respondents' and interested parties' (e.g., university Title IX offices) concerns.

Summary of Advantages and Limitations

Developing and fielding a new survey to measure experiences of sexual harassment has several advantages. The central advantage is the ability to design a methodology tailored to meet the needs of this specific effort, starting with designing a questionnaire that is fully inclusive of the constructs of interest and allows for incorporation of fully validated scales without compromising validity or reliability. Additionally, fielding a new survey allows for the optimization of the sample size and the implementation of oversampling where needed while avoiding collecting unnecessary extra data from groups that are already adequately represented. This approach can provide greater precision and reduce the margin of error for survey estimates, leading to more accurate and reliable data. However, this requires the use of a robust sampling frame, both in terms of number of records and covariates of interest. With an appropriate sample frame, fielding a new survey provides the best opportunity to achieve the study's goals for representation of the total population and key subgroups of interest, including women, racial and ethnic minorities, and SGMs. Furthermore, a new survey can be conducted without disrupting other survey efforts. As part of our discussion in Option 3—leveraging an existing survey—we will discuss some of the risks involved with leveraging existing surveys to collect data on experiences of sexual harassment. Those risks will be fully mitigated if a new survey effort is launched instead.

Although fielding a new survey has many advantages, there are some limitations to this approach. First, this would be the most expensive and time-intensive option. Fielding a new survey requires additional measures to ensure accuracy and reliability but does allow NCSES to meet their goals in measuring the prevalence of sexual harassment and other related behaviors in STEM while gathering more nuanced information from a shortened metric of sexual harassment. Second, as with all options, the survey will contain sensitive questions about potentially troubling experiences, which may cause hesitation on part of the respondent to answer these questions. This may impact response rates. However, this approach allows us to craft specific materials for survey respondents regarding the potentially uncomfortable nature of the topics while highlighting the importance of capturing this information.

In summary, fielding a new survey solely dedicated to assessing sexual harassment and related constructs has the benefit of including only a necessary number of items and reducing the chance of respondent burnout. However, as noted, the sensitive nature of the items may lead to a greater likelihood of attrition. Of the options under consideration, fielding a full new survey would be the most expensive and time intensive, yet it allows the greatest flexibility in assessing prevalence of sexual harassment. Moreover, our qualitative message boards identified a preference for a stand-alone survey, and participants highlighted that a stand-alone survey would underscore the importance and serious nature of the subject matter, which may increase response rates.

5 Chapter 5: Option 2 – Field a Supplemental Survey

In this chapter, we describe the second survey option for measuring sexual harassment in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations. The structure of this chapter mirrors that of Chapter 3. Chapter 3 provides the overarching considerations for measuring sexual harassment and related constructs on a survey, and this chapter describes these same considerations as they relate specifically to fielding a supplemental survey. As described in Chapter 3, each of the considerations are interconnected and, in many cases, decisions related to one consideration will impact decisions related to other considerations.

Overview

The second approach under consideration leverages existing surveys to measure experiences of sexual harassment—in this case, by fielding off-cycle supplemental surveys that use either or both the National Survey of College Graduates (NSCG) and the National Training, Education, and Workforce Survey (NTEWS) panels to obtain a sample. This approach represents a middle ground between fielding an entirely new survey and adding items to an existing survey (Survey Option 3). In this section, we describe the specific considerations and identified strengths and limitations of administering a supplemental survey. Table 6 provides an overview of the specific approaches for Option 2.

Table 6. Option 2 Overview

Consideration	Approach
Collaborator Engagement	• Identify and engage with collaborators in the selection of constructs to measure sexual harassment and related constructs. • Proactively communicate with collaborators about potential constraints and practical limitations to measuring sexual harassment and related constructs specific to this survey option. • Leverage these partnerships to proactively address potential concerns or criticisms of the developing approach.
Constructs of Interest	• Prioritize an intersectional approach when assessing sexual harassment and the reporting of sexual harassment by including measures assessing heterosexist harassment, racialized sexual harassment, and the reporting of these experiences.
Coverage of Populations of Interests	• Add a supplemental survey to both the NSCG or NTEWS. • If possible, consider oversampling subpopulations of interest and adding sexual orientation and gender identity questions to help cover populations of interest.
Sampling Design	• Use a subset of the current NSCG and NTEWS sample to receive the supplemental survey. • Potentially bolster sample by sourcing from the American Community Survey (ACS).
Question Development	• Use established and validated behaviorally based measures of sexual harassment and related constructs. • Implement a modular format where select participants are randomized to receive a maximum of two modules assessing different constructs.

	• Assess the utility of administering a supplemental survey for both surveys.
Fielding Frequency and Timing	• Field surveys at a time when it will be minimally disruptive to the main surveys. • Leave a gap between the conclusion of the host survey's previous wave and the start of the supplemental survey fielding.
Mode of Delivery	• If possible, only use self-administration web and paper surveys for the supplemental survey. • Review existing survey outreach policies, and update to include information on the supplemental survey (i.e., contact participants via email for participation, if possible).
Privacy Concerns	• The existing NSCG and NTEWS privacy policies should be reviewed to ensure protections cover sensitive data. If needed, privacy policies should be updated.
Communication Materials	• Clearly communicate the purpose, impact, and use of the survey before, during, and following survey administration. • Partner with organizations to raise awareness of the survey and potentially improve trust among survey respondents.

Approach Considerations

Collaborator Engagement

In developing a supplemental survey to measure sexual harassment and related constructs, as with the other survey options, it is important to consider identifying and engaging potential collaborators. For designing a supplemental survey, collaborators could be instrumental in selecting and prioritizing question selection and constructs to measure. With this option, there may be constraints in terms of the number of constructs that can be measured and limitations related to survey fielding frequency and coverage of populations of interest (e.g., sexual and gender minorities [SGM] students and professionals). As such, having direct lines of communication with individuals with an interest in the data from this survey will afford National Center for Science and Engineering Statistics (NCSES) the opportunity to proactively communicate these limitations and constraints instead of responding to related criticisms after the fact.

Implementation Recommendation: Like with Option 1, the first step to engage potential collaborators is to identify them and then develop a plan to strategically incorporate their input into survey development. We also recommend that NCSES coordinate with collaborators to proactively help them understand the limitations of this approach (described further throughout this chapter).

Constructs of Interest

Similar to Option 1, the administration of a supplemental survey offers the ability to include all four constructs of interest (i.e., sexual harassment, reporting of sexual harassment, heterosexist harassment, and racialized sexual harassment). If of interest to NCSES, the supplemental survey could also capture other facets of sexual harassment (i.e., technology-facilitated sexual violence). Chapter 2 and the Constructs of Interest section in Chapter 3 provide more information on constructs that are recommended for inclusion.

Implementation Recommendation: Like Option 1, we recommend that NCSES first identify the constructs of interests that should be measured on a supplemental survey in coordination with collaborators. We recommend measuring sexual harassment and several additional behaviors that have been shown to co-occur with sexual harassment. Given the number of constructs recommended, it may be necessary to prioritize certain constructs. To fully understand the interconnected nature of these experiences and how they disproportionately impact underrepresented groups in STEM, ultimately collecting a wider range of harassing experiences will

be central to future data analyses to understand these workplace cultures and climates and their impact on the STEM workforce.

Coverage of Populations of Interests

If fielding a supplemental survey, we recommend using both the NSCG and NTEWS as host surveys. To understand the prevalence of sexually harassing behaviors in STEM, STEM students and STEM professionals should be assessed. Both of these groups of interest are at least partially represented in the sample of the NSCG and NTEWS. The NSCG covers the select workforce population with at least a bachelor's degree and who are younger than 76. The NTEWS covers individuals ages 16–75[10] across all levels of the workforce (e.g., entry level up to management and upper-leadership positions) and includes individuals who are currently pursuing an undergraduate or graduate STEM degree. Both surveys also contain individuals who are outside of the population of interest (neither STEM students nor STEM professionals). Additionally, we note that leveraging these surveys may result in limitations in terms of coverage of underrepresented groups such as women, SGMs, and racial and ethnic minority students and professionals (described further in the Sampling Design question). Moreover, because neither the NTEWS nor NSCG ask about SGM status in the demographic questions, the experiences of these groups cannot be fully accounted for with this approach without adding questions measuring SGM status to the surveys.

Implementation Recommendation: If fielding a supplemental survey, we recommend leveraging both the NSCG and NTEWS. We also recommend using a subset of the current NSCG and NTEWS sampling frame to field the survey, which may allow for targeting of specific subgroups of interest (e.g., women, racial and ethnic minorities). If possible, NCSES may consider oversampling the subpopulations of interest from the existing sampling frames for the supplemental survey only. Another consideration would be to use the ACS to bolster sample sizes for key subpopulations.

Sampling Design

For this option, we recommend that only a subset of the current NSCG and NTEWS sample should receive the supplemental survey. While collecting data for this broad population, both surveys will present opportunities to focus sampling for key subgroups of interest, potentially sampling them at higher rates. The subgroups could be based on a combination of factors, including gender, race or ethnicity, age, field of study, level of education, and stage of career. The sampling plan will need to consider the desired overall level of precision and the needed levels of precision for each subgroup of interest. The overall sample size and sample allocation should be minimized to meet these constraints, as it needs to be large enough for inference, but small enough to avoid incurring unnecessary bias in the legacy data collection if response rates or response patterns change. If needed, additional fresh samples can be sourced from the ACS to bolster sample sizes for smaller subgroups. This sampling plan will need to be paired with a customized weighting plan that accounts for these additional design features.

This approach does have limitations related to the ability to oversample underrepresented groups and the assessment of gender or sexual minority status. Currently, the NSCG does not oversample members of underrepresented groups, such as women and racial/ethnic minorities, and although the NTEWS does stratify based on racial/ethnic identity and sex, the degree of oversampling and the final composition of the eligible respondents has not been published. As women and racial and ethnic minorities are already underrepresented in the STEM field, oversampling these groups would better capture their experiences—which is especially important because members of these groups are more likely to experience sexual harassment (Beal, 2008; Bowleg et al., 2003). Data from any oversampling efforts for these groups from the NTEWS should inform the final sample design of a potential supplemental survey. If this option is selected, we may recommend oversampling specific subpopulations on future administrations of these surveys. Moreover, NCSES surveys do not currently ask about sexual or gender minority status, so the experiences of this group cannot be

[10] Further discussion of Institutional Review Board (IRB), Office of Management and Budget (OMB), and legal considerations relate to the inclusion of individuals under 18 years old is provided in Appendix D.

explored within the proposed context. Adding these questions to future NCSES surveys could address this limitation. Similarly, the supplemental survey could include these demographic questions, which would allow for the exploration of the relationships between sexual or gender minority status and sexually harassing behaviors. However, care would need to be taken to avoid bias or priming in responses. The location of these additional demographic variables would need to undergo pilot testing and cognitive interviewing to ensure that survey error is minimized. With limited survey space, adding additional demographic variables to a supplemental survey would likely be inadvisable.

Specific to Option 2, we identified *two* potential surveys to add the supplemental survey. Although it may be simpler to use a single survey rather than two surveys, spreading the needed sample between two surveys would reduce the impact on any single existing panel. Feasibility for using both surveys depends on the ability to find a period to field that is off cycle for both surveys. If that window does not exist, then only a single panel should be used. As mentioned previously, the supplemental survey items should only be included in self-administered surveys. Also, operational simplicity should be considered. If it is much simpler or less error prone to use one survey's panel instead of another or both, then that survey would likely be the better option for implementation.

Implementation Recommendation: We recommend leveraging a subset of the current NSCG and NTEWS sample to receive the supplemental survey. We also recommend a fresh look at sample design informed by the recent NTEWS administration to ensure that the sample has adequate coverage of key populations of interest. Finally, we suggest potentially adding questions related to SGM status.

Question Development

Many of the Question Development considerations from Option 1 apply to fielding a supplemental survey. Similar to Option 1, using a supplemental survey dedicated specifically to the assessment of sexual harassment and related constructs offers a great deal of flexibility with respect to survey design. It also allows for the inclusion of existing and validated behaviorally based measures of sexual harassment and related constructs. Additionally, using a modular questionnaire design as described in Option 1 is recommended for fielding a supplemental survey so that multiple constructs of interest can be assessed while avoiding unnecessary error due to external factors, such as respondent burden. The supplemental should implement a modular format where select participants are randomized to receive a maximum of two modules assessing different constructs. For more information on the modular design, please refer to the Question Development section of Chapter 3 and Option 1 (Chapter 4).

Implementation Recommendation: Question development recommendations mirror those in Option 1 to include leveraging behaviorally based scales and conducting an environmental scan to select scales based on NCSES's informational needs. As with the first option, we recommend making minimal wording changes to existing scales (as opposed to sweeping changes to question wording, question order, or the number of questions for each scale) and then documenting those changes from the original scale along with a justification for the change (Merhill et al., 2021). Changes in scales may not be uniform and may need to be adapted slightly differently depending on target population (e.g., students, professionals).

Following item selection, we recommend deciding which order the questions will be asked. As outlined in the Question Development section of Chapter 3, the order of questions on a questionnaire can influence the lens in which someone interprets future questions. Because this is a supplemental survey with information unrelated to the host survey, it is crucial to be mindful of which constructs are presented first and how modules are prefaced for respondents.

Once the supplemental survey instrument is developed, conducting cognitive interviews and pilot and pretesting will be necessary. As with Option 1, even when using existing, validated scales, it will be necessary to conduct cognitive interviews to adapt question wording in ways to ensure that it resonates with the target populations. Additionally, changes to question wording from the original

scales may impact how respondents interpret questions, and cognitive interviews will help NCSES more fully understand the potential impacts of question wording on responses. Since this survey would include multiple constructs of interest we advise conducting cognitive interviews to determine the most appropriate order for respondents to see questions or receive modules. Similarly, questions will need to be pilot tested to examine the validity of each question. When conducting cognitive interviews and pilot testing, the same sample size considerations apply as with Option 1.

Fielding Frequency and Timing

Since both the NSCG and the NTEWS include longitudinal data collection, these surveys will allow for sexual harassment data to be collected over time and use the same panel approach as those surveys. As for the timing of the supplemental survey, it should be fielded at a time when it will be minimally disruptive to the main surveys. If only one of the NSCG or NTEWS panels will be used, then fielding should take place near the middle of the period between survey administrations. Because the NSCG and NTEWS are fielded on an alternate biannual schedule, there would be ample time between surveys for panel members to participate, reducing the potential for any respondent fatigue. If the NSCG and NTEWS are both used as sample sources, then finding a fielding window is more complex, since the surveys have different "off years." In this scenario, the primary focus should be on minimizing impact on existing survey efforts. Supplemental survey fielding should stop before the host surveys are fielded—potentially around six months before host survey fielding, but the time frame will be partially dependent on the length of fieldwork for the supplemental survey. Similarly, there needs to be a gap between the conclusion of the host survey's previous wave and the start of supplemental survey fielding to avoid affecting the response rate of the supplemental survey. If it is not feasible to avoid overlapping fielding with both of the host surveys, then only a single survey panel should be used as a potential sample source.

Implementation Recommendation: Because this is a supplement to the NSCG or NTEWS, it is recommended to field the supplemental survey during a time when it is minimally disruptive to the main surveys.

Mode of Delivery

The supplemental survey will ultimately use the same mode of delivery as the NSCG and NTEWS. However, we recommend limiting the survey administration for the supplemental survey to only self-administration to avoid biases in participation from subgroups that would be more hesitant to participate due to the sensitive nature of the survey. For example, on the NSCG, very few sample members opt for the phone survey, and we would expect the same to be true on the NTEWS, which means the bias incurred by only asking these questions on self-administered surveys would be small. However, if significant portions of the NTEWS sample responds by telephone, the risk of biasing survey estimates of experiences of sexual harassment increases. If that occurs, then the NSCG sample should be leaned on more heavily, as it will likely be less biased for the portion of the population it covers.

Additionally, it may be beneficial to review the existing outreach and dissemination of the NSCG and NTEWS (e.g., contacting participants to complete the survey via email) to reflect the addition of the supplemental survey. Specifically, language would be added to address the purpose of the supplemental survey and when it would field in the outreach phase of the NSCG and NTEWS to increase visibility of the survey.

Implementation Recommendation: We recommend limiting the survey administration for the supplemental survey to self-administration via web or paper surveys.

Privacy Concerns and Communication Material

The infrastructure for protecting respondent data is largely already in place for the NSCG and the NTEWS. However, given the sensitive nature of the survey topic and questions, it will be important to develop and communicate a robust privacy policy to ensure respondents are aware of the data

privacy protections in place. Moreover, privacy policies should be reviewed and updated to incorporate considerations surrounding sensitive topics and included on the NSCG and NTEWS.

As with Option 1, we recommend NCSES consider sharing information related to the supplemental survey purpose with collaborators or through their website well ahead of fielding to raise their awareness of the upcoming survey and garner interest. Additionally, communication about the supplemental survey should be developed and incorporated into the NSCG and the NTEWS as well as on the supplemental survey itself. We recommend including language that focuses on the purpose and expected fielding of the supplemental survey during the outreach and fielding phases of the NSCG and NTEWS.

Implementation Recommendation: We recommend survey-specific privacy policies to ensure the required protections are in place to handle the sensitive data. If additional protections are required, the survey privacy policies should be updated. Additionally, language related to privacy considerations should be shared during outreach as well as updating language directly on the survey to describe the purpose of the supplemental survey, how data will be handled to ensure anonymity, and outline guidelines related to the release of data. As with Option 1, we recommend developing a robust communications campaign for the supplemental survey.

Summary of Advantages and Limitations

Fielding off-cycle supplemental surveys of sexual harassment using either the NSCG and NTEWS is an efficient method that strikes a compromise between fielding a new survey (Option 1) and adding questions directly to an existing survey (Option 3) and offers some clear advantages. One of the key advantages of fielding a supplemental survey is its efficiency without a compromise in data quality. Both the NSCG and the NTEWS have a panel of respondents that could be used to field a survey. The data from previous administrations of both surveys can be leveraged to identify which panel members are in the target population and which are outside of the population of interest. Having a ready-made sampling frame with rich demographic data will significantly reduce the time and cost associated with getting an entirely new survey into the field. Moreover, a key feature of this option allows for the opportunity to have a supplemental survey fully devoted to measuring experiences of sexual harassment and other related constructs using fully validated scales. The ability to leverage validated scales (as opposed to selecting a subset of questions or drafting entirely new questions) will significantly improve confidence in the quality of the data collected without the constraint of limited survey space.

Although this option has the potential to provide high-quality data, there are several limitations. First, as with each of the options, the new sexual harassment measures are sensitive topics that may impact response rates. A feature of the supplemental survey is that messaging can be crafted specific to the supplemental questions to help prepare survey respondents for the potentially uncomfortable nature of the topics. Nevertheless, the host survey topic is unrelated to questions about sexual harassment and other related constructs, and the shift to these sensitive topics may cause hesitation on part of the respondents. The sensitive nature of the subject matter and the potential perception that the supplemental survey questions are unrelated to the core nature of the host survey pose the potential risk of panel attrition. This attrition could be active, where they notify the fieldwork agency directly and ask to be removed from the panel, or passive, where they simply do not respond to future NSCGs or NTEWSs. Attrition is a known issue for all panel surveys and can be partially addressed with survey weights, but some increases in both the bias and the variance of the host survey estimates are inevitable. Lower response rates will yield less data and wider confidence intervals. Additionally, the panel members who attrit will not be a random subset, and although attrition rates may be higher in some identifiable demographic subgroups than others, traditional survey weighting adjustments can only partially address the increased bias.

Another limitation includes the host surveys' limited sample pool, which limits options in terms of sample composition. As discussed previously, the sampling plan will need to consider the desired overall level of precision as well as the needed levels of precision for each subgroup of interest. There may not be enough participants in the NSCG or NTEWS sample pools to achieve an adequate sample size for all subgroups. The ACS sampling frame could be used to supplement the NSCG or NTEWS frames, but if this option is not available, analytic power may be insufficient for some subgroups of interest.

Of the options under consideration, fielding a supplemental survey is more costly than adding items directly to an existing survey, but less expensive than fielding an entirely new survey. This option strikes a balance between Options 1 and 3 by providing the flexibility to assess sexual harassment and related constructs while using an existing survey's sample.

‘

6 Chapter 6: Option 3 – Add to an Existing Survey

In this chapter, we describe the third survey option for measuring sexual harassment in STEM. We provide an overview of the approach, discuss specific considerations (based on the key considerations in Chapter 3), describe implementation recommendations, and provide a summary of the advantages and limitations. The structure of this chapter mirrors that of Chapter 3. Chapter 3 provides the overarching considerations for measuring sexual harassment and related constructs on a survey, and this chapter describes these same considerations as they relate specifically to adding measures to an existing survey. As described in Chapter 3, each of the considerations are interconnected and, in many cases, decisions related to one consideration will impact decisions related to other considerations.

Overview

This section describes the third potential option to gather information related to sexual harassment in STEM by adding questions to an existing National Center for Science and Engineering Statistics (NCSES) survey. This option poses the greatest risk to obtaining high quality data and to disrupting the existing NCSES surveys that would contain these new questions. Specifically, due to survey length, NCSES would be limited to the measurement of one construct, as adding multiple constructs would likely increase survey length beyond what is acceptable. Even with the inclusion of one construct, leveraging behaviorally based measures would add significantly to survey length. Due to the inability to fully assess sexual harassment and related constructs and potential disruption to the existing surveys, we do not recommend this option. Table 7 provides an overview of the specific approaches for Option 3.

Table 7. Option 3 Overview

Consideration	Approach
Collaborator engagement	• Identify and engage with collaborators in the selection of constructs to measure sexual harassment and related constructs. • Proactively communicate with collaborators about potential constraints and practical limitations to measuring sexual harassment and related constructs specific to this survey option. • Leverage these partnerships to proactively address potential concerns or criticisms of the developing approach.
Constructs of interest	• Prioritize the most important construct of interest—sexual harassment—to include on an existing survey.
Coverage of populations of interests	• Add items to the National Survey of College Graduates (NSCG) and/or the National Training, Education, and Workforce Survey (NTEWS) to capture both STEM students and professionals.
Sampling design	• Use existing sampling approach as the NSCG and NTEWS.
Question development	• Select items from valid, established measures of the constructs of interest to include on the NSCG and NTEWS. • Recall period should be determined based on needs of NCSES. • If measures are modified, pretesting and cognitive interviewing are required to ensure that reliability and validity of the selected items remains.

	• Examine the effects of using only a subset of items from a valid measure by pilot testing and using cognitive interviewing to ensure that reliability and validity of the selected items remains.
Fielding frequency and timing	• Use existing fielding frequency and timing of administration as the NSCG and NTEWS.
Mode of delivery.	• Use the existing NSCG and NTEWS data collection approach. • If a significant portion of the NTEWS sample responds by telephone, the sample may be biased, and therefore, the NSCG sample may result in more reliable results. • Limit sexual harassment survey items to self-administered web and paper surveys only.
Privacy Concerns	• Use the existing NSCG and NTEWS privacy policies. Update policies as needed to ensure the safeguarding of sensitive information. • Update privacy language within the survey and outreach materials to directly address the addition of the sexual harassment survey items.
Communication materials	• Convey clear messaging about the purpose, impact, and use of sexual harassment survey items before, during, and following survey administration. • Use clear messaging in the survey itself explaining the purpose of a sexual harassment measurement. • Partner with organizations to raise awareness of the survey and potentially improve trust among survey respondents.

Approach Considerations

Collaborator Engagement

As with other survey options, it is important to identify and engage potential collaborators. Because this option presents the most challenges in terms of the range of constructs that can be measured, it will be important to coordinate with collaborators to prioritize the construct to measure and questions to be used. Additionally, because this approach limits the number of constructs that can be considered, some groups or individuals who are committed to combatting sexual harassment may not see constructs they view as important included in the survey. As a result, they may criticize the survey approach and attempt to undermine its credibility. Developing a proactive communication strategy with those with a vested interest in these data may help mitigate the likelihood that they openly criticize the approach. Collaboration with potential detractors could bolster their investment in the survey's success, even if the new measures do not meet their data needs in the immediate term. As with previous options, establishing direct lines of communication with individuals with an interest in the data from this survey will afford NCSES the opportunity to proactively communicate practical limitations and constraints instead of responding to criticisms related to them after the fact.

Implementation Recommendation: Like with Options 1 and 2, the first step to engage potential collaborators is to identify them and then develop a plan to strategically incorporate their input into survey development. We also recommend NCSES coordinate with collaborators to proactively help them understand the limitations of this approach (described further throughout this chapter).

Constructs of Interest

Due to limited survey space and the need to leverage behaviorally based scales, we recommend only identifying and measuring one construct if using this approach. We would recommend measuring sexual harassment. Ultimately, the ability to only measure one construct would limit the extent to which analyses can be used to understand experiences associated with sexual harassment, and the data obtained would be far less comprehensive.

Implementation Recommendation: Prioritize measuring sexual harassment.

Coverage of Populations of Interests

To capture responses from both STEM students and professionals, it would be necessary to add items to both the NSCG and the NTEWS. Information on survey coverage of the NSCG and NTEWS is provided in the Population of Interest Section in Option 2. As described in Option 2, adding new items to existing NCSES surveys limits the ability to oversample underrepresented groups and to assess experiences of sexual and gender minorities (SGM). Currently, the NSCG does not oversample members of underrepresented groups, such as women and racial/ethnic minorities, and the extent of any oversampling for NTEWS has not been published. Similar to Option 2, we recommend fielding these sensitive questions to only a subset of the NSCG and NTEWS respondents, which may allow for targeting of specific subgroups of interest (e.g., women, racial and ethnic minorities) while minimizing overall impact on host surveys.

Implementation Recommendation: Add sexual harassment questions to the NSCG and NTEWS. If possible, only field to a subset of the NSCG and NTEWS samples.

Sampling Design

Given that we recommend questions to be added to the NSCG and the NTEWS, this option will use their sampling approaches. As mentioned in the Coverage of Populations of Interest section of this chapter and similar to Option 2, we recommend fielding these sensitive questions to only a subset of the NSCG and NTEWS respondents. Specifically, we recommend oversampling specific subpopulations within the subset of the sample, if possible, to ensure adequate coverage of marginalized subgroups. By only fielding these questions to a subset of the population, disruptions to the NSCG and NTEWs may be minimized. Additional information on recommendations related to oversampling subpopulations of interest can be found in the Coverage of Populations of Interest sections in Options 1 and 2. If this approach is implemented, the overall sample size and sample allocation should be minimized to meet these constraints, as it needs to be large enough for inference, but as small as possible to avoid incurring unnecessary bias in the legacy data collection if response rates or response patterns change. Moreover, the existing weighting plan will need to be updated to account for these additional design features.

Implementation Recommendation: We recommend fielding sensitive questions to only a subset of the NSCG and NTEWS sample. To ensure adequate coverage of specific subpopulations (e.g., women and racial/ethnic minorities), we recommend adjusting the sampling approach to oversample specific subgroups. This would involve additional steps to update the sampling approach and require adjustment to the existing weighting plan.

Question Development

When adding questions to an existing survey, space is prioritized, and NCSES should plan to include only sexual harassment, which is the construct of the highest interest. As with Options 1 and 2, it is recommended to use an existing, validated behaviorally based measure of sexual harassment. However, it may not be feasible to use the entire sexual harassment scale, and as a result, NCSES may want to modify (i.e., shorten) an existing scale to consolidate the number of questions asked.[11] To shorten an existing scale, variable selection methods are often used to select a subset of items from a larger scale. These variable selection methods seek to identify items that offer the most information within the context of the overall scale; however, these methods require a high level of effort and often only eliminate a small number of items. For instance, Item Response Theory (IRT) is a well-established and respected method for creating shortened measures since it identifies items that contribute the most information to the overall scale. Prior research has evaluated the ability to shorten the 23-item Sexual Experiences Questionnaire – DoD (SEQ-DoD) measure using IRT methods. The researchers were able to create the SEQ-DoD-s, which is a 16-item shortened measure of the SEQ-DoD (Stark et al., 2002; see Appendix A for items on the SEQ-DoD

[11] It is not advisable to remove items from a construct without extensive testing to ensure that items are still valid without being positioned within the greater measure.

and the SEQ-DoD-s). Although seven items were removed from the scale, many steps had to be taken to eliminate those items. The steps used to eliminate the seven items are listed below (Stark et al., 2002).

1. Data collection: IRT requires large sample sizes to produce stable parameter estimates, so a large sample was obtained that assessed the full SEQ-DoD to evaluate items for deletion.
2. Evaluation of item response option coverage: IRT relies on coverage across all spectrums of the response options, meaning that each item and their response options need to be evaluated to determine if items need to be eliminated prior to conducting analyses or to determine if response options need to be collapsed.[12]
3. Identification of items to delete: The full scale can be evaluated for item deletion by examining each item individually and the contribution of that item to the full scale. An item which contributes little to the overall scale can be evaluated for deletion. Although items can be identified for deletion, it is important to maintain the general structure of the original scale, meaning that constructs within the SEQ (i.e., Sexist Hostility, sexist behavior; Sexist Hostility, crude or offensive behavior; Unwanted Sexual Attention; and Sexual Coercion) stay separated within the shortened form.
4. Evaluation of model fit: The fit of the model without the item is evaluated compared to the fit of the model with the item. If the overall model is not significantly weakened by the deletion of the item, the item can be removed from the scale.
5. Repeat steps 3 and 4 until no more items can be removed. It should be noted that a minimum of three items per construct is required for mathematical estimation within analyses (Watkins, 2018). Therefore, to create a shorted assessment of sexual harassment using the SEQ-DoD, a minimum of 12 items (i.e., three items per construct) is required.

After identifying the shortened scale, it is important to validate the new scale using an independent sample in which the shortened scale is administered and then compared to the full scale. Given the level of effort associated with this approach and the potential to eliminate only a small number of questions, we do not recommend pursuing the development of a short form of an existing measure of sexual harassment to attach to an existing NCSES survey.

If methods are to be taken to modify an existing measure, it is crucial to include a measure of gender harassment in the shortened measure. Failing to account for gender harassment can lead to an underestimation of the prevalence of sexual harassment (Merhill et al., 2021). NCSES should meet with collaborators to determine scale reduction, recall periods, and response scales of the modified measure. Additional validation testing will be needed to ensure reliability and validity of the scale. It should be noted, creating a measure of sexual harassment based on a subset of items from an existing scale is not recommended and would jeopardize the reliability and validity of data collected (Merhill et al., 2021). The steps to modify an existing scale are outlined in the Implementation Recommendation section of this chapter.

Once a measure is selected, the placement of the sexual harassment items within the greater survey will need to be examined. Survey context can influence the way respondents understand survey questions (see *Survey Review* for larger discussion). One concern would be the placement of sexual harassment questions amidst unrelated topics (e.g., college STEM coursework or educational expectations) on the NSCG and NTEWS. Placing sexual harassment questions with entirely unrelated questions may lead to confusion or hesitation among respondents. To minimize impact on existing survey questions, we recommend that questions asking about experiences of sexual harassment be placed at the end of the survey, with an introduction to the questions to better prepare respondents.

[12] Collapsing response options is not recommended without thorough investigation into how reliability and overall validity of the scale is affected.

Additionally, we recommend that questions on topics such as military service and functional ability[13] either be removed or placed at or near the beginning of the demographics section. These types of questions could unintentionally lead to potential context effects when respondents answer questions about experiences of sexual harassment. For example, military service questions may prompt respondents who have served to focus on their time in the military when answering questions about sexual harassment. Additionally, some individuals may have long-term functional disabilities resulting from their experience of sexual harassment, and asking about functional limitations shortly before asking about experiences of sexual harassment may impact their responses to the sexual harassment items. These questions may also engender greater emotional duress from respondents. Moving these questions to earlier placement in the questionnaire or removing them all together may help avoid these potential issues. Cognitive interviews and pilot testing can help identify instances of these disruptions to data reliability and validity prior to survey administration, and therefore, is strongly recommended.

Due to potential limitations in available survey space, a modified or new measure may be implemented. Cognitive interviews and pilot testing will be needed. To ensure that modified or new survey items are measuring the intended construct, the items should undergo pretesting and cognitive interviews within the context of the larger survey. Item nonresponse is a threat to validity that cognitive interviewing and pilot testing can attempt to identify and mitigate. As we discussed earlier and in Option 2, in some cases, the addition of these sensitive items to surveys will depress survey response rates, which will create nonresponse bias for measurements of sexual harassment. Low item response rates are common for sensitive questions, potentially increasing the standard error and widening confidence intervals for survey estimates. Item nonresponse also increases risk of bias, as there is likely to be a relationship between the constructs themselves and who chooses to respond. Cognitive interviewing and pilot testing can identify instances of nonrandom item nonresponse, and cognitive interviewing can be used to probe when participants would prefer not to respond to an item.

Cognitive interviewing can also give insight into unintended ordering effects by probing participants on their thoughts when prompted with the selected sexual harassment questions within the context of the greater survey. Pilot testing can ensure that items maintain reliability and validity, especially if only a selected number of participants are administered the full measure for comparative purposes. More specifically, one subgroup could be administered the full scale within the context of the full survey, and another subgroup could be administered only the selected items within the context of the full survey. Reliability measures and validation techniques (e.g., examinations of relationships with known correlates) can then be examined to ensure that reliability and validity hold for the subset of items.

Additional items being added to a survey can potentially create complications for existing imputation and weighting methods. As discussed in the survey review, sensitive items may follow different item-nonresponse patterns compared to non-sensitive items on the same survey, which may require fresh assessment of imputation approaches for longitudinal surveys. Secondly, adding sensitive items that are administered to a subset of the sample may require separate nonresponse modelling from the rest of the sample. Not only will these items need different weights than items asked of the full sample, but these items may induce survey nonresponse differentially across demographic subgroups in a different fashion than the surveys with less sensitive survey items. In such a case, this would require re-evaluation of current survey weighting approaches to ensure they are still appropriate for the legacy data-collection efforts.

Implementation Recommendation: We recommend leveraging behaviorally based questions drawn from existing, validated scales. As with the other options, we recommend conducting an environmental scan of existing validated scales measuring sexual harassment. This process may allow for identification of a shorter validated measurement of sexual harassment to include in the

[13] Both the NSCG and NTEWS ask questions regarding military service and functional ability.

survey. However, if no measures exist that fit the needs of the existing survey (i.e., are too long), it may be necessary to modify an existing measure or create a new measure. It should be noted that it is not advisable to remove items from a construct without extensive testing to ensure that items are still valid without being positioned within the greater measure. In the event that NCSES decides to modify an existing scale, we recommend that they coordinate with the originators of that scale to discuss potential modifications and fully examine modifications that had been previously made to that scale. We then recommend NCSES conduct cognitive interviews and pilot testing on the modified scale. Because sexual harassment questions would be incorporated into an existing survey of unrelated content, we recommend the sexual harassment measure be placed at the end of the survey. Additionally, it may be beneficial to move questions related, such as military service and functional limitations, to earlier in the survey to reduce potential context effects. Following the placement into the existing surveys, conducting additional cognitive interviews and pilot and pretesting will be necessary to assess item ordering.

Fielding Frequency and Timing

The fielding frequency and timing of survey administration is already in place for the NSCG and the NTEWS. However, the recall period (as mentioned in Question Development) for the sexual harassment measure should be determined based on the needs of NCSES and collaborators.

Mode of Delivery

The infrastructure for data collection and processing and survey weight development is already in place for the NTWES and the NSCG. Therefore, data collection concerns are limited to those surrounding the dissemination of the selected surveys. Adding a few items would not require significant additional resources. However, the method in which the survey will be administered is worth considering. As mentioned in Option 2, survey questions assessing sexually harassing behaviors and related constructs can induce unequal levels of stress among subgroups of interest. As in Option 2, it is recommended that these items should only be included in self-administered surveys (see Survey Feilding for Option 2, as the same considerations apply). Limit sexual harassment survey items to self-administered web and paper surveys only.

Implementation Recommendation: We recommend only measuring sexual harassment by way of self-administered web or paper surveys.

Privacy Concerns and Communication Material

The infrastructure for protecting respondent data is already in place for the NTWES and the NSCG. Therefore, privacy concerns are limited to respondents within the selected existing surveys. It may be beneficial to review the existing privacy policies in place to ensure they meet the requirements to safeguard sensitive data. Additionally, we do recommend updating privacy language to directly address concerns surrounding the addition of sensitive questions on survey materials.

Similar to Option 1 and 2, communication materials surrounding the addition of the sensitive questions should be created. The same considerations surrounding the need for a communication strategy and the importance of working with collaborators also apply. Unique to Option 3 is the need to clearly communicate within the survey the purpose of these items and focus messaging on why they are included in an existing survey and not measured separately. This may alleviate potential concern or surprise as to why those items are added to the NSCG and NTEWS. Therefore, we recommend creating clear messaging prior to survey administration to state that the construct was added as well as the purpose of the addition. Specifically, language within the survey will need to be updated to reflect the addition of the construct and preface the purpose of the measurement prior to the survey items. Messaging surrounding the use and reporting of results will also be important to include to encourage item response.

Implementation Recommendation: Update privacy language to include considerations related to measuring sensitive topics, and review existing privacy policies and protections to ensure they are appropriate for the collection of sensitive data. We also recommend creating communication materials intended to explain the addition of sensitive questions on an existing NCSES survey.

Further, we recommend developing a methodological brief to proactively address concerns or questions related to survey item development and methodological rigor associated with the data collection. This would be particularly important if existing scales undergo significant modifications.

Summary of Advantages and Limitations

Taking advantage of the existing survey infrastructures of the NSCG and NTEWS presents the possibility that this option is less costly than either of the first two options. This option would limit NCSES to measuring only one new construct—sexual harassment—as opposed to sexual harassment in addition to several related constructs. Although there is a potential cost savings to be had by adding questions to an existing survey, this approach would require adjustments to existing survey questions and adapting the questionnaire for the inclusion of sensitive content.

Additionally, the feature of leveraging the infrastructure of an existing survey also poses significant limitations. Even if only adding a few additional survey items, due to their sensitive nature, their effects may be felt in several key areas creating a degree of disruption to existing survey efforts. The first potential limitation is the shift of survey items to the questionnaire. As noted above, at least two items may need to be moved to earlier in the survey or removed altogether from the survey to include sexual harassment questions. Depending on NCSES priorities, the questions we recommend removing may be required for inclusion on those surveys and therefore cannot be removed. Removing a question from the survey or shifting survey items would disrupt data collection for that item. To determine the impact of moving a survey item to a different location, it is important to consider both the nature of the item and the potential for contextual effects through an evaluation of the updated survey. This evaluation should ensure that the items that come before the moved question in its new location do not influence its interpretation. The ideal final placement would be one that minimizes the sum of these effects. Related to this concern is that of relevance. Although the questions on sexual harassment need to be placed at the end of the survey, they run the risk of feeling “tacked on.” Most of the other items in the survey are unrelated to the topic of sexual harassment, which may cause confusion or hesitation on part of the respondents when they encounter these items.

The second main impact could be on response rates and reduced representativeness of the data. As discussed in the survey review and in Options 1 and 2, the sensitivity of questions regarding sexual harassment may lead to nonresponse or underreporting due to concerns of embarrassment or invasion of privacy, as well as the risk of disclosure of sensitive information. Nonresponse can lead to reduced sample size, reduced statistical power, and may create bias if individuals who have experienced sexual harassment are less likely to respond. Additionally, rates of experiencing sexual harassment vary across subpopulations, and the impact will vary by group. This will require detailed analysis after data collection and comparisons to prior administrations to gauge the extent to which bias may have been introduced into the longitudinal data collection process. This is particularly crucial for the longitudinal NSCG and NTEWS, as any induced nonresponse bias may be baked into the panel on future surveys. Long term, there is also the potential that adding such sensitive questions may induce panel attrition.

Of the options under consideration, this option is potentially the least expensive way to collect this data compared to Options 1 and 2. However, we do not recommend this option. Although the immediate cost may be lower, the quality of data obtained will likely not meet NCSES’s goals.

7 Chapter 7: Summary and Future Research Opportunities

Chapter Overview

This chapter summarizes the three survey options described in Chapters 4–6, as well as the three core recommendations for the National Center for Science and Engineering Statistics (NCSES) in their development of survey measures to better understand sexual harassment in STEM. We conclude the chapter with proposed lines of future research both for continued development of a sexual harassment survey measure and the creation of messaging approaches as well as looking toward potential analyses to conduct with sexual harassment and related construct data once it is gathered.

Survey Design Recommendations Summary

We recommend Options 1 or 2 to best address NCSES's data collection needs: fielding a new survey or fielding a supplemental survey. Given the strengths and limitations of each approach, either Options 1 or 2 would provide NCSES with high-quality, valid, and reliable data related to the incidence of sexual harassment and other related constructs in STEM. Although Option 3 has the potential to be the least costly, there are significant considerations related to survey design and limitations associated with data that can be collected. As such, the costs associated with Option 3 still likely would not result in data of such a quality that would reliably provide a rate of sexual harassment in STEM.

Survey Option	Recommendation
Option 1: Field a New Survey	✓ Recommended
Option 2: Field a Supplemental Survey	✓ Recommended
Option 3: Add Items to Existing Survey	⚠ Not recommended

Based on the findings from our literature review, survey review, and qualitative message boards, we identified three core recommendations for the development of a measurement capturing sexual harassment in STEM:

Recommendation 1: Prioritize an Intersectional Approach
To assess the prevalence of sexual harassment in STEM and understand the impact across the entire STEM enterprise, we recommend not only measuring sexual harassment, but also constructs that often co-occur, specifically heterosexist behaviors and racialized sexual harassment (Konik & Cortina, 2008; Lee, 2018; Leskinen & Cortina, 2014). We also recommend assessing reporting behaviors associated with experiences of harassment. It is essential to include measures that examine harassment from a broader social lens to understand how social identities intersect and influence experiences of harassment (Buchanan, 2005b; Buchanan et al., 2018; Cho, 1997; Williams, 2014). Given that research has found individuals with multiple, marginalized identities are at increased risk for sexual harassment (Beal, 2008; Bowleg et al., 2003; Schuyler et al., 2020), understanding the impact of sexual harassment within these communities in STEM is crucial, particularly with goals of eliminating barriers to individuals from diverse backgrounds fully participating in the STEM enterprise. We expect that these three distinct—but related—constructs

will capture a wider range of behaviors and provide better insight into the experiences of STEM students and professionals.

Recommendation 2: Leverage Validated Behaviorally Based Measures

Regardless of the constructs selected for inclusion, we strongly recommend the use of behaviorally based measures of sexual harassment and/or related constructs to assess its prevalence within the STEM enterprise—preferably an established, validated measure, such as the Sexual Experiences Questionnaire (SEQ). This is consistent with recommendations from the National Academies of Sciences, Engineering, and Medicine (NASEM) and best practices recommended by the U.S. Government Accountability Office (GAO) for government agencies seeking to measure sexual harassment (GAO, 2020; NASEM, 2020). As recommended by NASEM, any selected measure should include gender harassment, as it is the most common form of sexual harassment (NASEM, 2018). Additionally, as mentioned in Recommendation 1, to understand the full extent of sexual harassment in the STEM enterprise, heterosexist behaviors, racialized sexual harassment, reporting behaviors, and potentially technology-facilitated sexual violence should also be included and assessed using validated scales. Appendix A provides several examples of existing, validated scales for sexual harassment and related constructs, including the SEQ, which is the most commonly used measure for sexual harassment (NASEM, 2018).

Recommendation 3: Field a New or Supplemental Survey

To capture a complete picture of sexual harassment among students and professionals in STEM, we recommend the creation of a new or supplemental survey (fielding off-cycle from the main survey). Because of the robust nature of these measures, it may not be feasible to add a few questions onto an existing survey without sacrificing data quality and validity. Therefore, a stand-alone or supplemental survey is recommended, as it would allow for the inclusion of all constructs of interest and the use of validated measures (which often include numerous survey items) without the need for modification. Additionally, fielding a new or supplemental survey would negate serious concerns regarding the potential disruption to existing surveys (e.g., panel attrition) while emphasizing the seriousness of the subject matter to respondents. This recommendation aligns with findings from our qualitative message boards: the majority of participants indicated their preference for a stand-alone survey on sexual harassment as opposed to adding questions to an existing survey.

Ultimately, decisions about how to collect data on the prevalence of sexual harassment in STEM will be shaped by the realities of available resources. As such, adding questions to an existing survey may be perceived as an expediate approach. If NCSES does elect to add questions to an existing survey, additional steps to mitigate potential risks (i.e., unvalidated measures, shift in survey topics, item nonresponse, disruption to existing surveys) should be taken. Specifically, if possible, selecting a validated measure or conducting extensive validation testing, thoughtful item placement of sensitive topics, and sufficient preface to sensitive topics may help mitigate risks. These risks may nullify the perceived benefits of such a course of action—for more information on this approach, refer to Option 3 (Chapter 7). Ultimately, the risks associated with adding questions directly to an existing survey are such that we recommend against this option.

Future Research Opportunities

In addition to the recommendations described in this report, we have identified several lines of further exploration that could aid in measuring sexual harassment in STEM. As NCSES expands its research into workplace and gender relations within STEM, there are additional areas that warrant additional investigation. Specifically, additional research on the messaging and delivery of this measure may be needed to help alleviate participant concerns. We also identify potential research that leverages administrative data and qualitative research to potentially bolster the survey on future administrations.

Collaborator Research

Given the centrality of identifying and engaging with collaborators in the development and messaging regarding measures of sexual harassment and related constructs, we recommend conducting an environmental scan to identify parties who may use these data to create policy changes and understand the climate and culture within STEM disciplines, and organizations committed to combatting sexual harassment in the workplace, academia, and more specifically, in STEM fields. After identifying potential collaborators, we recommend developing a plan for their involvement in survey development and identifying crucial junctures in the development phase at which to engage them for feedback. Engaging potential collaborators will allow for them to feel involved in the process and will engender greater investment from them in the success of the survey. Moreover, leveraging these collaborators to both structure communication materials and to promote the survey can help increase response rates. By partnering with well-known and trustworthy collaborators, respondents may be more likely to take the survey. For example, if partnering with Title IX offices, student participants, especially those at greater risk for these experiences, may view the survey as more trustworthy and be more aware of the survey and its intended implications, which may increase their likelihood to complete the survey.

Survey Messaging

We also recommend considering conducting research to inform the communications strategy for a future survey assessing experiences of sexual harassment to this specific population. Findings from this research could be used to shape the communication and marketing plan once the survey is ready to field. Specifically, research is needed to understand how different subpopulations perceive the selected measures of sexual harassment and related constructs, especially individuals from marginalized backgrounds (e.g., women, racial/ethnic minorities, sexual and gender minorities [SGM]). Gathering this information will help NCSES understand motivations to answering these items, and more importantly, potential barriers to survey completion. Although our qualitative message boards gave us a glimpse into these potential motivators (making an impact on policy) and barriers (privacy concerns, lack of impact), additional research is needed when an option (new, supplemental, or existing survey) is selected, as motivators and barriers are likely to differ based on the option. Once a communication strategy is created, additional qualitative research could engender an understanding of how the messaging campaign resonates with potential participants.

Leverage Administrative Data

Leveraging administrative data can be a powerful and cost-effective tool to both understand patterns in career progression and pipeline loss as well as to inform survey development. For example, information about gender and racial demographics is already available in the administrative data collected by universities and can be used to help understand the general demographic makeup of STEM departments in academic institutions and track demographic trends over time.[14] Administrative data could be used to better understand underrepresentation in academic institutions (e.g., specific roles or departments within a university) and identify where and when pipeline loss most often occurs. This information may point to crucial time periods that STEM students and professionals need to be surveyed and potentially inform survey fielding frequency and recall periods for experiences of interest.

Additionally, NCSES could leverage administrative data to examine demographic trends in STEM, which would align with NASEM (2020) recommendations aimed at combatting the underrepresentation of women in STEM. Administrative data could be used to understand the gender, racial, and ethnic compositions of STEM academic departments including students, trainees, staff, and faculty. Understanding baseline patterns using administrative data may also

[14] Information about gender identity (other than male/female) or sexual orientation are not currently collected in administrative data. Therefore, patterns related to these factors currently cannot be explored using administrative data alone. If these demographic factors are of interest, the addition of survey questions about sexual orientation and gender identity should be considered.

inform future research opportunities that pair administrative and survey data to further understand patterns of underrepresentation.

Administrative data's utility is also further extended when paired with survey data. In developing a survey, it could be useful to consider what additional research questions could be answered using survey data and potentially pairing it with administrative data. Although administrative data are not collected for research purposes, survey data often are. As such, considerations related to future analyses can be built into survey development. Moreover, understanding where to pair administrative data with survey data after it has been collected can open further possibilities for analyses that can be conducted by combining the two. For example, NCSES could conduct an analysis to understand pipeline loss between students who have and have not experienced sexual harassment in STEM by leveraging administrative records to detect when STEM students drop out of STEM fields. As such, we recommend NSCSES consider developing a research agenda aimed at further understanding the impacts of sexual harassment in STEM as well as risk and protective factors associated with sexual harassment (and other forms of harassment) to better understand the environments in which these experiences occur. The development of a longer-term research agenda can also help inform needs for data collection efforts. For example, in the longer term, if NCSES will want to understand barriers to reporting sexual harassment and patterns associated with reporting/non-reporting behavior, it would be important for NCSES to include measures assessing reporting barriers on their survey. We also recommend conducting an environmental scan of available administrative data sources capturing key information about STEM students and professionals, although we recognize that administrative records on STEM professionals may be sparser than those on STEM students.

Future Qualitative Research

Another potential avenue for future research is additional qualitative research to help illuminate the different ways harassment and discrimination manifest within STEM. This is of particular importance since there is a lack of consistency in the currently available gender discrimination measures (de la Torre-Pérez et al., 2022). First, qualitative research could be conducted to explicitly investigate more subtle forms of sexual harassment and discrimination in STEM. For example, message boards identified parental bias or maternal wall bias in the workplace as behaviors not currently assessed in measures of gender discrimination. The term maternal wall bias refers to barriers in the workplace related to past, current, or future pregnancies or maternity leaves (Williams, 2014). This phenomenon can manifest during the hiring process, interactions with colleagues, and during performance reviews. Given that women are already underrepresented in STEM, further qualitative research exploring experiences of gender discrimination, in particular maternal wall bias, could provide additional understanding of the experiences of women in STEM. Further qualitative research related to maternal wall bias could inform future measures of sexual harassment.

Second, additional qualitative research may identify field-specific factors that might influence sexually harassing behaviors in STEM. For example, in message board discussions about where harassment behaviors are most prevalent, offsite field work was repeatedly mentioned. Field work might present a unique and potentially vulnerable circumstance for sexual harassment and discrimination to occur. Additionally, as mentioned in the literature review, underrepresentation of women is variable across sub-fields within STEM, which results in varying sexual harassment rates across STEM fields. Together, these two field-specific factors (i.e., field work and variable representation) could have implications for the sexual harassment measurement. Therefore, conducting career field-specific focus groups (e.g., physics, chemistry, civil engineering) may identify additional behaviors or demographic factors that may provide additional context that is important for understanding quantitative survey data once it is collected. For instance, there may be differences in the types of experiences in a field work setting as compared to a non-field setting; therefore, including a question on where these behaviors occurred may help provide additional context to understand the extent of sexual harassment in STEM.

Conclusion

Our main goal throughout this plan is to provide NCSES with the primary considerations to incorporate into the development of a measure of sexual harassment in STEM with the goal of understanding the incidence of sexual harassment among STEM students and professionals. Each of the components of this plan were drawn directly from the three prior phases of research: the literature review, survey review, and qualitative message board data collection. Stemming from these phases, we also identified several related constructs that, if possible, will be important to measure alongside sexual harassment. As such, this implementation plan outlined several overarching considerations for measuring sexual harassment in STEM. Each of these considerations were incorporated into a discussion of three potential paths forward for measuring sexual harassment among STEM students and professionals.

Although we have provided these recommendations based on our research, we acknowledge that many of the decisions regarding the fielding of these survey questions will be based on NCSES's continued research and availability of resources. This effort serves as the foundation on which NCSES can build to ultimately develop a robust, reliable measure of sexual harassment in STEM. Having such a measure will allow for NCSES to understand the persistence of sexual harassment and related constructs in STEM, which can ultimately provide important information related to the underrepresentation of women, SGMs, and racial and ethnic minorities in STEM. Future research and the development of policies and programs to combat sexual harassment can be conducted to further understand the incidence of sexual harassment in STEM only after there is a foundational understanding of the pervasiveness of the issue.

References

Abbey, A., Parkhill, M. R., & Koss, M. P. (2005). The effects of frame of reference on responses to questions about sexual assault victimization and perpetration. *Psychology of Women Quarterly, 29*(4), 364–373. https://doi.org/10.1111/j.1471-6402.2005.00236.x

Administrator-Researcher Campus Climate Collaborative. (2015, September 1). *ARC3 Campus Climate Survey*. https://campusclimate.gsu.edu/arc3-campus-climate-survey/

Backe, E.L., Lilleston, P., & McCleary-Sills,, J. (2018). Networked individuals, gendered violence: A literature review of cyberviolence. *Violence & Gender*, *5*(3), 135–146. http://doi.org/10.1089/vio.2017.0056

Beal, F. M. (2008). Double jeopardy: To be Black and female. *Meridians, 8*(2), 166–176. 10.1300/J155v07n04_06

Bostock v. Clayton County, 590 U.S. __ (2020)

Bowleg, L., Huang, J., Brooks, K., Black, A., & Burkholder, G. (2003). Triple jeopardy and beyond: Multiple minority stress and resilience among Black lesbians. *Journal of Lesbian Studies, 7*(4), 87–108. https://doi.org/10.1300/J155v07n04_06

Breslin, R., Klahr, A., Hylton, K., Petusky, M., White, A., & Techa, J. (2019). *2018 Workplace and Gender Relations Survey of the Active Duty Military. Results and Trends* (Report No. 2019-024). Office of People Analytics

Buchanan, N. T. (2005a). The nexus of race and gender domination: The racialized sexual harassment of African American women. In P. Morgan & J. Gruber (Eds.), *In the company of men: Re-discovering the links between sexual harassment and male domination (pp. 294–320).* Boston, MA: Northeastern University Press.

Buchanan, N. T. (2005b). Incorporating race and gender in sexual harassment research: The Racialized Sexual Harassment Scale (RSHS). In N. T. Buchanan (Chair), *Expanding sexual harassment research to include diverse populations and intersecting forms of harassment*. Symposium presented at the meeting of the International Coalition Against Sexual Harassment, Philadelphia, PA.

Buchanan, N. (2016). *Racialized Sexual Harassment Scale: Items, citations, and terms of use*. http://hdl.handle.net/20.500.11990/1389

Buchanan, N. T., & Fitzgerald, L. F. (2008). Effects of racial and sexual harassment on work and the psychological well-being of African American women. *Journal of Occupational Health Psychology*, *13*(2), 137–151. https://doi.org/10.1037/1076-8998.13.2.137

Buchanan, N. T., Settles, I. H., Wu, I. H., & Hayashino, D. S. (2018). Sexual harassment, racial harassment, and well-being among Asian American women: An intersectional approach. *Women & Therapy, 41*(3-4), 261–280. https://doi.org/10.1080/02703149.2018.1425030

Campbell, J. (2014). *Johns Hopkins University campus sexual assault project*. Johns Hopkins University.

Cantor, D., Fisher, B., Chibnall, S., Harps, S., Townsend, R., Thomas, G., Lee, H., Kranz, V., Herbison, R., & Madden, K. (2020). *Report on the AAU Campus Climate Survey on sexual assault and misconduct*. https://www.aau.edu/sites/default/files/AAU-Files/Key-Issues/Campus-Safety/Revised%20Aggregate%20report%20%20and%20appendices%201-7_(01-16-2020_FINAL).pdf

Cech, E.A. & Rothwell, W.R. (2018). LGBTQ inequality in engineering education. *Journal of Engineering Education, 107*(4), 583–610. https://doi.org/10.1002/jee.20239

Cho, S. K. (1997). Converging stereotypes in racialized sexual harassment: Where the model minority meets Suzie Wong. *The Journal of Gender, Race, and Justice*, 1, 177–211.

Cohen, J. J., Gabriel, B. A., & Terrell, C. (2002). The case for diversity in the health care workforce. *Health Affairs, 21*(5), 90–102. https://doi.org/10.1377/hlthaff.21.5.90

Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D. (2001). Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology, 6*(1), 64–80. https://doi.org/10.1037/1076-8998.6.1.64

Fitzgerald, L.F., Drasgow, F., Hulin, C.L., Gelfand, M.J., & Magley, V.J. (1997). Antecedents and consequences of sexual harassment in organizations: A test of an integrated model. *Journal of Applied Psychology, 82*(4), 578–589. https://doi.org/10.1037/0021-9010.82.4.578

Fitzgerald, L. F., Gelfand, M. J., & Drasgow, F. (1995). Measuring sexual harassment: Theoretical and psychometric advances. *Basic and Applied Social Psychology, 17*(4), 425–445. https://doi.org/10.1207/s15324834basp1704_2

Fitzgerald, L.F., Magley, V.J., Drasgow, F., & Waldo, C.R. (1999). Measuring sexual harassment in the military: The Sexual Experiences Questionnaire (SEQ-DoD). *Military Psychology*, *11*(3), 243–263. https://doi.org/10.1207/s15327876mp1103_3

Fitzgerald, L.F., Shullman, S.L., Bailey, N., Richards, M., Swecker, J., Gold, Y., …, Weitzman, L. (1988). The incidence and dimensions of sexual harassment in academia and the workplace. *Journal of Vocational Behavior, 32*(2), 152–175. http://dx.doi.org/10.1016/0001-8791(88)90012-7

Follingstad, D. R., Chahal, J. K., Bush, H. M., Coker, A. L., Li, C. R., Wu, X., …, & Carlson, C. R. (2020). A campus climate/violence survey's psychometrics and findings. *Violence Against Women*, *27*(14), 2735–2767. https://DOI.org/10.1177/1077801220969870

Galesic, M., & Tourangeau, R. (2007). What is sexual harassment? It depends on who asks! Framing effects on survey responses. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *21*(2), 189–202. DOI: 10.1002/acp.1336

Government Accountability Office (GAO). (2020). *Workplace sexual harassment: Experts suggest expanding data collection to improved understanding of prevalence and cost.* https://www.gao.gov/assets/710/709862.pdf

Gutek, B. A., Murphy, R. O., & Douma, B. (2004). A review and critique of the Sexual Experiences Questionnaire (SEQ). *Law and Human Behavior, 28*(4), 457–482. https://doi.org/10.1023/B:LAHU.0000039335.96042.26

Hart, C.G. (2019). The penalties for self-reporting sexual harassment. *Gender & Society, 33*(4), 534–559. https://DOI.org/10.1177/0891243219842147

Hart, J. (2016). Dissecting a gendered organization: Implications for career trajectories for mid-career faculty women in STEM. *The journal of Higher Education, 87*(5), 605–634. https://doi.org/10.1080/00221546.2016.11777416

Hill, C., & Kearl, H. (2011). *Crossing the Line: Sexual Harassment at School*. Washington, DC: American Association of University Women. https://www.aauw.org/app/uploads/2020/03/Crossing-the-Line-Sexual-Harassment-at-School.pdf

Johnson, P. A., Widnall, S. E., & Benya, F. F. (Eds.). (2018). *Sexual harassment of women: climate, culture, and consequences in academic sciences, engineering, and medicine*. Washington, DC: National Academies Press. https://www.nap.edu/catalog/24994/sexual-harassment-of-women-climate-culture-and-consequences-in-academic

Klonoff, E. A., & Landrine, H. (1995). The schedule of sexist events: A measure of lifetime and recent sexist discrimination in women's lives. *Psychology of Women Quarterly, 19*(4), 439–470. http://dx.doi.org/10.1111/j.1471-6402.1995.tb00086.x

Knapp, D.E., Faley, R.H., Ekeberg, S.E., Dubois, C.I.Z. (1997). Determinants of target responses to sexual harassment: A conceptual framework. *Academy of Management Review*, *22*(3), 687–729. https://doi.org/10.5465/amr.1997.9708210723

Krebs, C., Lindquist, C., Berzofsky, M., Shook-Sa, B., Peterson, K., Planty, M., & Stroop, J. (2016). *Campus climate survey validation study: Final technical report*. Washington, DC: BJS, Office of Justice Programs.

Konik, J., & Cortina, L. M. (2008). Policing gender at work: Intersections of harassment based on sex and sexuality. *Social Justice Research, 21*(3), 313–337. https://doi.org/10.1007/s11211-008-0074-z

Kuchynka, S. L., Salomon, K., Bosson, J. K., El-Hout, M., Kiebel, E., Cooperman, C., & Toomey, R. (2018). Hostile and benevolent sexism and college women's STEM outcomes. *Psychology of Women Quarterly, 42*(1), 72–87. http://dx.doi.org/10.1177/0361684317741889

Lee, J. (2018). Passive leadership and sexual harassment: Roles of observed hostility and workplace gender ratio. *Personnel Review*, *47*(3), 594–612. https://doi.org/10.1108/PR-07-2016-0169

Leskinen, E.A. & Cortina, L.M. (2014). Dimensions of disrespect: Mapping and measuring gender harassment in organizations. *Psychology of Women Quarterly, 38*(1), 107–123. https://doi.org/10.1177/0361684313496549

Lytle, A., & Shin, J. E. (2020). Incremental beliefs, STEM efficacy and STEM interest among first-year undergraduate students. *Journal of Science Education and Technology*, *29*(2), 272–281. https://doi.org/10.1007/s10956-020-09813-z

Mahoney, P., LaCure, M., Erdice, S., & Decker, M.R. (2022). *Environmental and Situational Strategies for Sexual Violence Prevention: A Practitioners' Guide for Leveraging Evidence for Impact on College Campuses*. Baltimore, Maryland: Maryland Coalition Against Sexual Assault and John Hopkins Bloomberg School of Public Health.

Marín-Spiotta, E., Díaz-Vallejo, E.J., Barnes, R.T., Mattheis, A., Schneider, B., Berhe, A.A., Hastings, M.G., Williams, B.M., & Magley, V.J. (2023). Exclusionary behaviors reinforce historical biases and contribute to loss of talent in the earth sciences. *Earth's Future, 11*, 1–14. https://doi.org/10.1029/2022EF002912

McCann, C. Tomaskociv-Devey, D. & Badgett, M.V.L. (2018). *Employer's responses to sexual harassment.* Center for Employment Equity. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3407960

McNeeley, S. (2012). Sensitive issues in surveys: Reducing refusals while increasing reliability and quality of response to sensitive survey items. In L. Gideon (ed.). *Handbook of survey methodology for the social sciences* (pp. 377–396). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-3876-2_22

Merhill, N.H., Bonner, K.A., & Baker, A.L. (Eds.). (2021). *Guidance for measuring sexual harassment prevalence using campus climate surveys*. Washington, DC: National Academies of Services, Engineering, and Medicine. https://doi.org/10.7226/26346

Murdoch, M. & McGovern, P.G. (1998). Measuring sexual harassment: Development and validation of the sexual harassment inventory. *Violence and Victims, 13*(3), 203–216. doi: 10.1891/0886-6708.13.3.203

National Academies of Sciences, Engineering, and Medicine. (2018). *Sexual harassment of women: Climate, culture, and consequences in academic sciences, engineering, and medicine*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24994.

National Academies of Sciences, Engineering, and Medicine. (2020). *Promising practices for addressing the underrepresentation of women in science, engineering, and medicine: Opening doors*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25585

National Center for Science and Engineering Statistics (NCSES). 2023. *Diversity and STEM: Women, minorities, and persons with disabilities 2023*. Special Report NSF 23-315. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/wmpd.

Powell, A., & Henry, N. (2019). Technology-facilitated sexual violence victimization: Results from an online survey of Australian adults. *Journal of Interpersonal Violence*. *34*(17), 3637–3665. https://doi.org/10.1177/0886260516672055

Richey, C.R., Lee, K.M., Rodgers, E.M., & Clancy, K.B. (2019). Gender and sexual minorities in astronomy and planetary science face increased risks of harassment and assault. *Bulletin of the American Astronomical Society, 51(4)*, 1–11. http://doi.org/10.3847/25C2CFEB.C985281E

Rodrigues, M.A., & Clancy, K.B.H. (2020). *A comparative examination of research on why women are more underrepresented in some STEMM disciplines compared to others, with a particular focus on computer science, engineering, physics, mathematics, medicine, chemistry, and*

biology. https://nap.nationalacademies.org/resource/25585/Commissioned_Paper_Rodrigues.pdf

Rospenda**,** K. M., Richman, J. A., & Shannon, C. A. (2009). Prevalence and mental health correlates of harassment and discrimination in the workplace: results from a national study. *Journal of Interpersonal Violence*, *24*(5), 819–843. https://doi.org/10.1177/0886260508317182

Saewyc, E. M., Bauer, G. R., Skay, C. L., Bearinger, L. H., Resnick, M. D., Reis, E., & Murphy, A. (2004). Measuring sexual orientation in adolescent health surveys: Evaluation of eight school-based surveys. *Journal of Adolescent Health*, *35*(4), 345.e1–345.e15. https://doi.org/10.1016/j.jadohealth.2004.06.002

Schuyler, A.C., Klemmer, C., Mamey, M.R., Schrager, S.M., Goldbach, J.T., Holloway, I.W., & Castro, C.A. (2020). Experiences of sexual harassment, stalking, and sexual assault during military service among LGBT and non-LGBT Service members. *Journal of Traumatic Stress*, *33*(3), 257–266. https://doi.org/1002/jts.22506

Silver, J. K., Bean, A. C., Slocum, C., Poorman, J. A., Tenforde, A., Blauwet, C. A., Kirch, R. A., Parekh, R., Amonoo, H. L., Zafonte, R., & Osterbur, D. (2019). Physician workforce disparities and patient care: A narrative review. *Health Equity, 3*(1), 360–377. https://doi.org/10.1089/heq.2019.0040

Silverschanz, P., Cortina, L. M., Konik, J., & Magley, V. J. (2008). Slurs, snubs, and queer jokes: Incidence and impact of heterosexist harassment in academia. *Sex Roles: A Journal of Research, 58*(3-4), 179–191. https://doi.org/10.1007/s11199-007-9329-7

Suen, L.W., Lunn, M.R., Katuzny, K., Finn, S., Duncan, L., Sevelius, J., Flentje, A., Capriotti, M.R., Lubensky, M.E., Hunt, C., Weber, S., Bibbins-Domingo, K., & Obedin-Maliver, J. (2020). What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: A qualitative study. *Archives of Sexual Behavior, 49*, 2301–2318. https://doi.org/10.1007/s10508-020-01810-y

Swartout, K. M., Flack, Jr., W. F., Cook, S. L., Olson, L. N., Smith, P. H., & White, J. W., (2019). Measuring campus sexual misconduct and its context: The administrator-researcher campus climate consortium (ARC3) survey. *Psychological Trauma: Theory, Research, Practice, and Policy, 11*(5), 495–504. http://dx.doi.org/10.1037/tra0000395

Tilley, D. S., Wang, W., Kolodetsky, A., & Yeatts, P. (2020). Factor Analysis of the Administrator-Research Campus Climate Collaborative (ARC3) Survey. *Health Education & Behavior, 47*(1_suppl), 54S-69S.

de la Torre-Pérez, L., Oliver-Parra, A., Torres, X., & Bertan, M.J. (2022). How do we measure gender discrimination? Proposing a construct of gender discrimination through a systematic scoping review. *International Journal for Equity in Health*, *21*(1). https://doi.org/10.1186/s12939-021-01581-5

U.S. Equal Employment Opportunity Commission (n.d.b). *Sexual harassment.* https://www.eeoc.gov/sexual-harassment

Waldo, C. R. (1999). Working in a majority context: A structural model of heterosexism as minority stress in the workplace. *Journal of Counseling Psychology, 46*(2), 218–232. https://doi.org/10.1037/0022-0167.46.2.218

Wasti, S. A., & Cortina, L. M. (2002). Coping in context: sociocultural determinants of responses to sexual harassment. *Journal of personality and social psychology*, *83*(2), 394. doi: 10.1037//0022-3514.83.2.394.

Williams, J.C. (2014). Double jeopardy? An empirical study with implications for the debates over implicit bias and intersectionality. *Harvard Journal of Law & Gender*, *37*, 186–242. https://repository.uchastings.edu/faculty_scholarship/1278

Zhong, L.R., Kebbell, M.R., & Webster, J.L. An exploratory study of Technology-Facilitated Sexual Violence in online romantic interactions: Can the Internet's toxic disinhibition exacerbate sexual aggression? *Computers in Human Behavior*, 108, 106314. https://doi.org/10.1016/j.chb.2020.106314

A Appendix A: Sexual Harassment Measures

Appendix A provides an overview of the existing scales measuring sexual harassment and other constructs of interest. Although this is not an exhaustive overview of existing scales, we have provided a table of measurements that were discussed in the literature review. For each measure, we highlight the construct assessed, if it is behaviorally based, the number of items and the response scale, recall period, and any considerations. As highlighted in Table A1, the majority of scales are behaviorally based and include multiple questions intended to measure sexual harassment and related constructs.

Table A1. Summary of Measures

Scale/Measurement	Construct(s)	Behaviorally Based	Number of Items	Response Scale	Recall Period	Considerations
Sexual Experiences Questionnaire (SEQ; Fitzgerald et al., 1988)	Sexual harassment: · Unwanted sexual attention · Sexual coercion · Gender harassment	Yes	18–24	5-point scale: 1 = Never 5 = Most of the time/Many times/Very often	· Ever · Past year · Past 2 years	· Has been widely used and adapted to various populations studying sexual harassment. · Lead-in to the scale, specific items, and lack of follow-up items do not allow for behaviors to be consistently assessed as unwanted, unreciprocated, or unwelcome.
Gender experiences questionnaire (GEQ; Leskinen & Cortina, 2014)	Gender harassment: · Sexist remarks · Sexually crude/offensive behavior · Infantilization · Work/family policing · Gender policing	Yes	20	5-point scale: 1 = Never 5 = Many times	· Past year	· Can be a stand-alone tool or used in combination with the unwanted sexual attention and sexual coercion subscales of the SEQ.
Sexual Harassment Inventory (SHI; Murdoch & McGovern, 1998)	Sexual harassment: · Hostile work environment · Quid pro quo · Criminal sexual misconduct	Yes	20	Binary response: Yes/No	· Past year	· Found participants to be confused by measurement of lifetime and past-year prevalence in pilot studies, recommend separate administrations. · Operationalizes sexual harassment according to the legal definition.
Schedule of Sexist Events (SSE; Klonoff & Landrine, 1995)	Sexist discrimination: · Sexist degradation · Sexist discrimination in distant relationships · Sexism in close relationships · Sexist discrimination in the workplace	No*	20	6-point scale: 1 = Never happened 6 = Happened almost all the time	· Ever · Past year	· Overlap with concept of sexual harassment is limited but has been used to develop measures of other similar constructs, such as ambivalent sexism (Kuchynka, et al., 2018). · *Frequent use of “sexism” throughout the instrument, requiring respondents to label their experiences as such.
Campus Climate Surveys						

Campus Climate Survey Validation Study (CCSVS; Krebs et al., 2016)	Sexual harassment: · Unwanted sexual attention	Yes	5	Binary response: Yes/No	· Since beginning of the current academic year	· Explicitly includes experiences that occurred in-person or over other media (e.g., by phone, on social media). · Assesses respondents' reporting experiences with respect to sexual assault, but not sexual harassment.
ARC3 Campus Climate Survey (ARC3, 2015)	Sexual harassment: · Physical harassment (similar to unwanted sexual attention) · Non-physical harassment (similar to gender harassment)	Yes	19–23 per module	5-point scale: 1 = Never 5 = Many times	· Not provided	· Includes separate modules for assessing behaviors carried out by students and by faculty/staff. · Includes modules assessing perceptions of institutional responses to reports of sexual misconduct and knowledge of sexual misconduct resources.
Follingstad et al. (2020) Campus Climate Survey	Sexual harassment	Yes	6–8	4-point scale: 0 = Never 3 = Often	· Past year	· Not developed to uniquely capture gender harassment, unwanted sexual attention, and/or sexual coercion. · Includes items specifically assessing sexual harassment by an authority figure. · Assesses perpetrator's relationship to the university, but only once rather than for each behavior experienced. · Assesses respondents' reporting experiences with respect to sexual assault, but not sexual harassment. · Includes assessment of perceptions of institutional responses to reports of sexual assault, knowledge of sexual misconduct resources, and accuracy of that knowledge.
AAU Campus Climate Survey (Cantor et al., 2020)	Sexual harassment	Yes	5	Binary response: Yes/No	· Since enrolling in school	· Specifically developed to measure the legal definition of sexual harassment.

						· Assesses experiences with reporting sexual harassment. · Assesses perpetrator's relationship to the university for each behavior experienced.
Surveys with Related Constructs						
Workplace Incivility Scale (Cortina, Magley, Williams, & Langhout, 2001)	Interpersonal mistreatment	Yes	7	5-point scale: 0 = Never 4 = Most of the time	· Past 5 years	· Most common measure of workplace insecurity · Can be adapted to assess incivility perpetration by asking respondents how frequently they have engaged in such behaviors toward others.
Workplace Heterosexist Experiences Questionnaire (WHEQ; Waldo, 1999)	Direct and indirect heterosexism: · Derogatory slurs · Assumptions of sexuality	Yes	25	5-point scale: 0 = Never 4 = Most of the time	· Past year	· Can be modified to address specific sexual and gender minority subgroups
Racialized Sexual Harassment Scale (RSHS; Buchanan, 2005b)	Racial harassment, sexual harassment, and racialized sexual harassment.	Yes	21	5-point scale: 0 = Never 4 = Many times	· Past 12 months	· The racialized sexual harassment subscale can be used on its own · Can be modified to update language to reflect university or workplace environment
Technology-Facilitated Sexual Violence Victimization Scale (Powell & Henry, 2019)	Digital sexual harassment, image-based sexual abuse, sexual aggression and/or coercion, and gender/sexuality-based harassment	Yes	21	Binary response: Yes/No	· Lifetime	· Specifically developed to assess multiple dimensions of technology-facilitated sexual violence

Sexual Experiences Questionnaire

The Sexual Experiences Questionnaire (SEQ; Fitzgerald et al., 1988) uses behavioral based questions to assess for experiences of sexual harassment. In the tables below, the questions from the various SEQs are aligned with their specific construct. More information about the SEQ and specific constructs is available in the literature review. As noted in the summary table above, the recall period for the SEQ may vary, but researchers typically ask about the following timeframes: lifetime (i.e., "ever"), past year, or past 2 years. Additionally, respondents are provided either a 3-point response scale (never, once, more than once) or a 5-point response scale ranging from 1 indicating they have never encountered the behavior to 5 indicating they experience the behavior most of the time/many times/very often. Below we provide various versions of the SEQ to include the revised SEQ, often referred to as the SEQ-Workplace or SEQ-W (Table A2) and the Department of Defense (DoD) SEQ (Table A3). Table A3 also denotes items which were removed from the SEQ-DoD to create the SEQ-DoD-s.

Table A2. Survey Items from the SEQ-W (Fitzgerald et al., 1995)

Construct	Questions
Have you ever been in a situation where a supervisor or coworker (gave you/attempted to/made)...	
Gender Harassment	…told suggestive stories
	…made crude sexual remarks
	…made offensive remarks
	…displayed offensive materials
	…sexist comments
Unwanted Sexual Attention	…attempted to discuss sex
	…unwanted sexual attention
	…staring, leering at you
	…attempts to establish a sexual relationship
	…repeated requests for drinks, dinner, despite rejection
	…touching in a way that made you feel uncomfortable
	…attempts to stroke or fondle
Sexual Coercion	…subtly bribed you
	…subtly threatened you
	…made it necessary to cooperate to be well treated
	…made you afraid of poor treatment if you didn't cooperate
	…experienced consequences for refusing

Table A3. Survey Items from the SEQ-DoD and SEQ-DoD-s (Fitzgerald et al., 1999; Stark et al., 2002)

Construct	Questions
Gender Harassment, Sexist Hostility	Treated you "differently" because of your sex (for example, mistreated, slighted, or ignored you)?
	Displayed, used, or distributed sexist or suggestive materials (for example, pictures, stories, or pornography which you found offensive)?
	Made offensive sexist remarks (for example, suggesting that people of your sex are not suited for the kind of work you do?)
	Put you down or was condescending to you because of your sex?
Gender Harassment, Sexual Hostility	Repeatedly told sexual stories or jokes that were offensive to you?
	♦ Whistled, called, or hooted at you in a sexual way?
	♦ Made unwelcomed attempts to draw you into a discussion of sexual matters (for example, attempted to discuss or comment on your sex life)?
	♦ Made crude and offensive sexual remarks, either publicly (for example, in your workplace) or to you privately?
	Made offensive remarks about your appearance, body, or sexual activities?
	Made gestures or used body language of a sexual nature which embarrassed or offended you?
	♦ Stared, leered, or ogled you in a way that made you feel uncomfortable?
	♦ Exposed themselves physically (for example, "mooned" you) in a way that embarrassed you or made you feel uncomfortable?
Unwanted Sexual Attention	♦ Displayed, used, or distributed sexist or suggestive materials (for example, pictures, stories, or pornography which you found offensive)?
	Made unwanted attempts to establish a romantic sexual relationship with you despite your efforts to discourage it?
	Continued to ask you for dates, drinks, dinner, etc., even though you said "No"?
	Touched you in a way that made you feel uncomfortable?
	Made unwanted attempts to stroke, fondle, or kiss you?
	♦ Attempted to have sex with you without your consent or against your will, but was unsuccessful?
	♦ Had sex with you without your consent or against your will?
Sexual Coercion	Made you feel like you were being bribed with some sort of reward or special treatment to engage in sexual behavior?
	Made you feel threatened with some sort of retaliation for not being sexually cooperative (for example, by mentioning an upcoming review)?
	Treated you badly for refusing to have sex?
	Implied faster promotions or better treatment if you were sexually cooperative?
	♦ Made you afraid you would be treated poorly if you didn't cooperate sexually?

Note. ♦ Denotes items which were removed from the SEQ-DoD to create the shortened, SEQ-DoD-s

Gender Experiences Questionnaire (GEQ; Leskinen & Cortina, 2014)

The Gender Experiences Questionnaire (GEQ; Leskinen & Cortina, 2014) is a specific measure for gender harassment. The stem for all of the items in the table below reads: ‘‘During the PAST YEAR, has anyone associated with your WORK (e.g., supervisors, coworkers, clients/customers, collaborators at other companies) done any of the following behaviors?’’ The five response options ranged from (1) never, once or twice, sometimes, often, to (5) many times. More information about the GEQ is available in the literature review.

Table A4. Survey Items from the Gender Experiences Questionnaire (GEQ; Leskinen & Cortina, 2014)

Construct	Question
During the PAST YEAR, has anyone associated with your WORK (e.g., supervisors, coworkers, clients/customers, collaborators at other companies) done any of the following behaviors?	
Sexist Remarks	Made sexist remarks about people of your gender
	Referred to people of your gender in insulting or offensive terms
	Made sexist remarks or jokes about women in your presence
	Made sexist jokes in your presence
Sexually Crude/Offensive Behavior	Said crude or gross sexual things in front of others or to you alone
	E-mailed, texted, or instant messaged offensive sexual jokes to you
	Made unwanted attempts to draw you into discussion of sexual matters
	Told you stories of their sexual exploits when you did not want to hear them
	Displayed or distributed dirty pictures or stories (e.g., nude pictures)
Infantilization	Talked to you as if you were a small child instead of speaking to you like an adult
	Treated you as if you were stupid or incompetent
	Publicly addressed you as if you were a child (e.g., dear, kid, etc.)
Work/Family Policing	Suggested women are better suited for raising children than being in the workplace
	Suggested women belong at home, not in the workplace
	Said employees who are mothers are less productive than other employees
	Said employees who are mothers are less dependable than other employees
Gender Policing	Referred to the workplace as a ‘‘man’s space’’ (e.g., women do not belong here)
	Made you feel like you were less of a woman because you had traditionally masculine interests
	Criticized you for not behaving ‘‘like a woman should’’
	Treated you negatively because you were not ‘‘feminine enough’’

Sexual Harassment Inventory (SHI; Murdoch & McGovern, 1998)

The Sexual Harassment Inventory (SHI; Murdoch & McGovern, 1998) operationalizes the legal definition of sexual harassment, focusing on hostile environment, quid pro quo, and criminal sexual misconduct. Because of its focus on experiences that align with the legal definition of sexual harassment, items related to gender harassment are not included in this measure. Respondents answer "yes," or "no," to a series of questions. The final question "Were there other things of a sexual nature that happened to you while you were in [specific context]?" allows for write-in answers. The military version of the SHI is presented below in Table A5.

Table A5. Survey Items from the Sexual Harassment Inventory -- Military Version (SHI; Murdoch & McGovern, 1998)

Construct	Question
Hostile Environment	People with whom I worked made sexual jokes that made me feel uncomfortable.
	I was touched by a coworker or supervisor in ways that made me feel uncomfortable.
	A coworker frequently asked me out for dates, even though I had asked him/her to stop.
	Coworkers made sexual comments about my body.
	My supervisor or superior officer made sexual comments about my body.
	My coworkers made demeaning comments to me because I am a woman/man.
	I was given the most unpleasant, difficult assignments because I was a woman/man.
	I was given the most unpleasant, difficult assignments because I was a woman/man.
	The people I worked with put up posters of women/men in provocative poses.
	My supervisor or superior officer attempted to have sex with me without my consent.
	Some of the people I worked with leered at me in a sexual way.
	Some of the people I worked with made catcalls or sexual remarks when I walked by.
Quid Pro Quo	A supervisor or superior officer asked me out for dates, even though I had asked him/her to stop.
	A supervisor or superior officer threatened to block my promotion unless I agreed to have sex with him/her.
	A supervisor or superior officer threatened to block a favorable transfer unless I agreed to have sex with him/her.
	I was offered favorable assignments in exchange for sex with my supervisor or commanding officer.
	I was offered promotions in exchange for having sex with my supervisor or commanding officer.
	Were you ever prevented from getting a promotion, favorable assignment, or transfer because you refused to have sex with someone?
Criminal Sexual Misconduct	My coworkers or superior officer exposed themselves to me in a sexual way.
	A coworker or coworkers attempted to have sex with me without my consent.
	I was forced by a coworker or supervisor to have sex without my consent.

Other (Open text response)	Were there other things of a sexual nature that happened to you while you were in the military? (Please list)

Schedule of Sexist Events (SSE; Klonoff & Landrine, 1995)

The Schedule of Sexist Events (SSE; Klonoff & Landrine, 1995) measures experiences of sexist discrimination. Because it uses the term "sexist" throughout, the respondent would need to label that experience as sexist themselves, a potential limitation of the measure. Respondents answer each question twice: once for lifetime (i.e., "from when you were a child to now") and again for past year. Response choices include 1 ("If the event has NEVER happened to you,") 2, ("If the event happened ONCE IN A WHILE [less than 10% of the time]"), 3, ("If the event happened SOMETIMES [10–25% of the time]), 4, ("If the event happened A LOT [26–49% of the time]), 5, (If the event happened MOST OF THE TIME [50–70% of the time]), and 6 ("If the event happened ALMOST ALL OF THE TIME [more than 70% of the time]").

Table A6. Survey Items from Schedule of Sexist Events (SSE; Klonoff & Landrine, 1995)

Construct	Question
Sexist Degradation and Its Consequences	How many times have you want to tell someone off for being sexist?
	How many times have you heard people making sexist jokes, or degrading sexual jokes?
	How many times have you been called sexist names like bitch, cunt, chick, or other names?
	How many times have you been really angry about something sexist that was done to you?
	How many times have people made inappropriate or unwanted sexual advances to you because you are a woman?
	How many times have you gotten into an argument or a fight about something sexist that was done or said to you or done to somebody else?
	How many times have people failed to show you the respect that you deserve because you are a woman?
	How many times have you been made fun of, picked on, pushed, shoved, hit, or threatened with harm because you are a woman?
Sexist Discrimination in Distant Relationships	How many times have you been treated unfairly by people in service jobs (by store clerks, waiters, bartenders, waitresses, bank tellers, mechanics and others) because you are a woman?
	How many times have you been treated unfairly by strangers because you are a woman?
	How many times have you been treated unfairly by people in helping jobs (by doctors, nurses, psychiatrists, case workers, dentists, school counselors, therapists, pediatricians, school principals, gynecologists, and others) because you are a woman?
	How many times have you been treated unfairly by neighbors because you are a woman?
	How many times have you been treated unfairly by teachers or professors because you are a woman?
Sexist Discrimination in the Workplace	How many times were you denied a raise, a promotion, tenure, a good assignment, a job, or other such things at work that you deserved because you are a woman?
	How many times were you forced to take drastic steps (such as filing a grievance, filing a lawsuit, quitting your job, moving away, and other actions) to deal with some sexist thing that was done to you?
	How many times have you been treated unfairly by your employer, boss or supervisors because you are a woman?

	How many times have you been treated unfairly by your co-workers, fellow students or colleagues because you are a woman?
Sexism in Close Relationships	How many times have you been treated unfairly by your boyfriend, husband, or other important man in your life because you are a woman?
	How many times have you been treated unfairly by your family because you are a woman?
	How different would your life be now if you HAD NOT BEEN treated in a sexist and unfair way?

Campus Climate Survey Validation Study (CCSVS)

As part of an initiative by the White House Task Force to Protect Students from Sexual Assault, funding was allocated to develop a measure to that colleges and universities could use to assess the climate and prevalence of sexual harassment and sexual assault within their institutions (Krebs et al., 2019). The measure was tested on and designed for an undergraduate population. The questions in Table A7 below represent the items that form the constructs for sexual harassment and coerced sexual contact. For the questions asking about experiences of sexual harassment, the question stem read: "Since the beginning of the current academic year in [FILL: August/September], [YEAR], has anyone done the following to you either in person or by phone, text message, e-mail, or social media? Please include things regardless of where they happened." For the questions asking about experiences of coerced sexual contact, the question stem read: "For this next question, please think about the entire time since you entered college. If you have attended more than one school, please think about the time since you first entered any college or university. At any point since you entered college, has anyone had any of the following types of unwanted sexual contact with you (i.e., sexual contact without your consent and that you did not want to happen)?" All responses options were a binary choice of yes or no.

Table A7. Survey Items from Campus Climate Survey Validation Study (CCSVS)

Construct	Question
Sexual Harassment	Made sexual advances, gestures, comments, or jokes that were unwelcome to you
	Flashed or exposed themselves to you without your consent
	Showed or sent you sexual pictures, photos, or videos that you didn't want to see
	Showed or sent sexual photos/videos of you or spread sexual rumors about you that you didn't want shared
	Watched or took photos/videos of you when you were nude or having sex, without your consent
Coerced Sexual Contact	Forced touching of a sexual nature (forced kissing, touching of private parts, grabbing, fondling, rubbing up against you in a sexual way, even if it is over your clothes)
	Oral sex (someone's mouth or tongue making contact with your genitals or your mouth or tongue making contact with someone else's genitals)
	Anal sex (someone putting their penis in your anus)
	[RESPONSE WILL NOT DISPLAY IF D3=MALE] Sexual intercourse (someone putting their penis in your vagina)
	Sexual penetration with a finger or object (someone putting their finger or an object like a bottle or a candle in your [IF D3= FEMALE OR TRANSGENDER OR SOMETHING ELSE OR MISSING, FILL: "vagina or"] anus)

ARC3 Campus Climate Survey (ARC3, 2015)

Like the CCVS, the ARC3 measure of sexual harassment was informed by the American Association of University Women (AAUW) sexual harassment survey (Hill & Kearl, 2011; Krebs et al., 2016), as well as the SEQ (Fitzgerald et al., 1995). The measure assesses two types of sexual harassment: physical harassment (e.g., made unwanted attempts to establish a romantic sexual relationship) and nonphysical harassment (e.g., repeatedly told sexual stories or jokes that were offensive; Tilley et al., 2020). The ARC3 survey leverages a modular approach to allow universities to tailor it to their specific needs. As a result, the survey's length and content may vary across contexts in which it is fielded. The sexual harassment measure is administered in two separate modules, one capturing behaviors perpetrated by faculty/staff and another capturing the same behaviors carried out by students (Swartout et al., 2019). The questions in Table A8 and Table A9 below represent the items that form the constructs for sexual harassment among faculty/staff and students, respectively. The response options for each question is : 0 (never), 1 (once or twice), 2 (sometimes), 3 (often), 4 (many times).

Table A8. Survey Items from ARC3 - Faculty

Construct	Question
Sexist Hostility/Sexist Gender Harassment	Treated you "differently" because of your sex (for example, mistreated, slighted, or ignored you)?
	Displayed, used, or distributed sexist or suggestive materials (for example, pictures, stories, or pornography which you found offensive)?
	Made offensive sexist remarks (for example, suggesting that people of your sex are not suited for the kind of work you do)?
	Put you down or was condescending to you because of your sex?
Sexual Hostility/Crude Gender Harassment	Repeatedly told sexual stories or jokes that were offensive to you?
	Made unwelcome attempts to draw you into a discussion of sexual matters (for example, attempted to discuss or comment on your sex life)?
	Made offensive remarks about your appearance, body, or sexual activities?
	Made gestures or used body language of a sexual nature which embarrassed or offended you?
Unwanted Sexual Attention	Made unwanted attempts to establish a romantic sexual relationship with you despite your efforts to discourage it?
	Continued to ask you for dates, drinks, dinner, etc., even though you said "No"?
	Touched you in a way that made you feel uncomfortable?
	Made unwanted attempts to stroke, fondle, or kiss you?
Sexual Coercion	Made you feel like you were being bribed with some sort of reward or special treatment to engage in sexual behavior?
	Made you feel threatened with some sort of retaliation for not being sexually cooperative (for example, by mentioning an upcoming review)?
	Treated you badly for refusing to have sex?

	Implied better treatment if you were sexually cooperative?

Table A9. Survey Items from ARC3 - Students

Construct	Question
Sexist Hostility/Sexist Gender Harassment	Treated you "differently" because of your sex (for example, mistreated, slighted, or ignored you)?
	Displayed, used, or distributed sexist or suggestive materials (for example, pictures, stories, or pornography which you found offensive)?
	Made offensive sexist remarks (for example, suggesting that people of your sex are not suited for the kind of work you do)?
	Put you down or was condescending to you because of your sex?
Sexual Hostility/Crude Gender Harassment	Repeatedly told sexual stories or jokes that were offensive to you?
	Made unwelcome attempts to draw you into a discussion of sexual matters (for example, attempted to discuss or comment on your sex life)?
	Made offensive remarks about your appearance, body, or sexual activities?
	Made gestures or used body language of a sexual nature which embarrassed or offended you?
Unwanted Sexual Attention	Made unwanted attempts to establish a romantic sexual relationship with you despite your efforts to discourage it?
	Continued to ask you for dates, drinks, dinner, etc., even though you said "No"?
Sexual Harassment Aia Electronic Communication	Sent or posted unwelcome sexual comments, jokes or pictures by text, email, Facebook or other electronic means?
	Spread unwelcome sexual rumors about you by text, email, Facebook or other electronic means?
	Called you gay or lesbian in a negative way by text, email, Facebook or other electronic means?

Campus Climate Survey (CCS; Follingstad et al., 2020)

Like the CCVS Campus Climate Survey, CCS was created for institutes of higher education and includes multiple scales that address different constructs (e.g., rape myth acceptance, intimate partner violence, stalking). The CCS's six-item scale for sexual harassment is based on past measures of sexual harassment, with items being adapted from Fitzgerald et al. (1995), Decker and Hudson (1994), and Campbell (2014). The stem for all six-items is "(During the past 12 months), how often has someone (NOT someone you are dating or a spouse/partner) done any of the following to you?" The response options for each question is : 0 (never), 1 (once), 2 (sometimes), 3 (often).

Table A10. Survey Items from Campus Climate Survey (Follingstad et al., 2020)

Question
During the past 12 months), how often has someone (NOT someone you are dating or a spouse/partner) done any of the following to you?
• Said sexual things to you that you did not want to hear?
• Sent sexual messages or pictures that you did not want (including porn)?
• Asked or pressured you for a date, hook up, or sexual favors even though you had already said no?
• Made unwanted sexual gestures or imitated sexual motions when you did not want them to?
• Touched you sexually (breasts, buttocks, or genitals) when you did not want them to?
• Exposed themselves to you (breasts, buttocks, or genitals) when you did not want them to?

Association of American Universities (AAU) Campus Climate Survey (Cantor et al., 2020)

Similar to the CCVS described above, the AAU Campus Climate Survey is intended to be a tool to assist colleges and universities to understand the climate at their institution. The intended audience for the survey is undergraduate and graduate students, as well as faculty and staff at institutes of higher education. The survey includes measures of sexual assault, intimate partner violence, stalking, and sexual harassment. Students are first asked about their knowledge of university resources related to sexual assault and sexual misconduct prior to assessments of experiences.

Knowledge of specific university services and resources is assessed through the following question:

Are you aware of the services and resources provided by the following? (Mark all that apply)

- [UNIVERSITY SPECIFIC LIST]
- None of the above

Next, knowledge of resources is gauged further through the following four questions. Response options for all questions included "Not at all," "A little," "Somewhat," "Very," and "Extremely."

Table A11. Survey items from AAU Campus Climate Survey Assessing Knowledge of Sexual Assault and Sexual Misconduct University Resources (Cantor et al., 2020)

Question
How knowledgeable are you about how sexual assault and other sexual misconduct are defined at [University]?
How knowledgeable are you about where to get help at [University] if you or a friend experienced sexual assault or other sexual misconduct?
How knowledgeable are you about where to make a report of sexual assault or other sexual misconduct at [University]?
How knowledgeable are you about what happens when a student reports an incident of sexual assault or other sexual misconduct at [University]?

In designing the section focused on sexual harassment, researchers used Leskinen and Cortina (2014)'s sexual harassment measure, with some modifications. In addition, to ensure the AAU Campus Climate Survey's measure of sexual harassment aligned with the legal definition,[15] the following stem was included on all questions in the section:

"These next questions ask about situation in which a student at [University] or someone employed by or otherwise associated with [University] said or did something that

- Interfered with your academic or professional performance,
- Limited your ability to participate in an academic program, or
- Created an intimidating, hostile, or offensive social, academic or work environment."

Response options for all questions were a binary "yes," or "never experienced."

Table A12. Survey Items from AAU Campus Climate Survey (Cantor et al., 2020)

Question
Since you have been a student at [University], has a student, or someone employed by or otherwise associated with [University] made sexual remarks or told jokes or stories that were insulting or offensive to you?
Since you have been a student at [University], has a student, or someone employed by or otherwise associated with [University] made inappropriate or offensive comments about your or someone else's body, appearance or sexual activities?
Since you have been a student at [University], has a student, or someone employed by or otherwise associated with [University]said crude or gross sexual things to you or tried to get you to talk about sexual matters when you didn't want to?
Since you have been a student at [University], has a student, or someone employed by or otherwise associated with [University]emailed, texted, tweeted, phoned, or instant messaged offensive sexual remarks, jokes, stories, pictures or videos to you that you didn't want?
Since you have been a student at [University], has a student, or someone employed by or otherwise associated with [University]continued to ask you to go out, get dinner, have drinks or have sex even though you said, "No"?

If a respondent answered "yes" to any of the question in Table A12, the following questions were assessed to understand reporting of sexual harassment to the university.

[15] The U.S. Equal Employment Opportunity Commission (EEOC) and U.S. Department of Education (ED) states the behavior must create a "hostile or offensive work or academic environment," and the behavior needs to be "frequent or severe."

Table A13. Survey Items from AAU Campus Climate Survey Assessing Reporting of Sexual Harassment (Cantor et al., 2020)

Skip logic	Question	Response options
If yes to any questions in Table A12	Q1. Since you have been a student at [UNIVERSITY], have you contacted any of the following about any of these experiences? (Mark all the apply)	[UNIVERISTY SPECIFIC LIST] None of the above
If a program is selected in Q1	Q2. When did you most recently contact [Program] about these experiences?	Fall of 2018 – present Fall of 217 – Summer of 2018 Fall of 2016 – Summer of 2017 Prior to Fall of 216
If program NOT selected in Q1	Q3. Why did you decide not to contact any of these programs or resources? (Mark all that apply)	I did not know where to go or who to tell I felt embarrassed, ashamed or that it would be too emotionally difficult I did not think anyone would believe me I did not think it was serious enough to contact any of these programs or resources I did not want the person to get into trouble I feared negative academic, social or professional consequences I feared it would be kept confidential I could handle it myself I feared retaliation I didn't think these resources would give me the help I needed Incident occurred while school was not in session Other
Selected "not serious enough" or "other" in Q3	Q3a. You said you did not contact any of these programs or resources (because it was not serious enough/for an 'other' reason/because it was not serious enough and for an 'other' reason). Please review the list below and mark any of the reasons that may better describe why you didn't contact any of these programs or resources (Mark all that apply).	I was not injured or hurt The reaction by others suggested that it wasn't serious enough to contact any of these programs or services I contacted other programs or services that I felt were appropriate I had trouble reaching the program or service I was too busy The event happened in a context that began consensually Because of the person's gender, I thought it would be minimized or misunderstood I might be counter-accused Alcohol and/or other drugs were present Events like this seem common My body showed involuntary arousal Other:

If yes to any questions in Table A12	Q4. Which of the following person, if any, did you (also) tell about this? (Mark all that apply)	Friend Family member Faculty member or instructor Resident advisor (RA), or other live-in residential staff Other administrative staff Spiritual or religious advisor, leader, or clergy Therapist or counselor Sexual or romantic partner Program or resource outside the University (e.g., a hotline) Physician Someone else I didn't tell anyone (else)

Workplace Incivility Scale (Cortina, Magley, Williams, & Langhout, 2001)

The Workplace Incivility Scale (Cortina et al., 2001) measures interpersonal mistreatment in the workplace. It focuses on rude, condescending, or disrespectful behaviors employees may encounter in their place of work from their supervisor or coworkers. It was initially used to describe the experiences of employees in federal circuit court. The stem for all the items listed below reads, "During the past 5 years while employed by [WORKPLACE], have you been in a situation where any of your superiors or coworkers?" Response options for each item ranged from 0 ("never") to 4 ("most of the time"). Items were summed to create an overall incivility score.

Table A14. Survey Items from Workplace Incivility Scale (Cortina, Magley, Williams, & Langhout, 2001)

Question
Put you down or was condescending to you?
Paid little attention to your statement or showed little interest in your opinion?
Made demeaning or derogatory remarks about you?
Addressed you in unprofessional terms, either publicly or privately?
Ignored or excluded you from professional camaraderie?
Doubted your judgement on matter over which you have responsibility?
Made unwanted attempts to draw you into a discussion of personal matters?

Workplace Heterosexist Experiences Questionnaire (WHEQ; Waldo, 1999)

The WHEQ measures experiences of direct (e.g., called a derogatory slur) and indirect (e.g., assumptions of sexuality) heterosexism in the workplace using behaviorally based questions. Respondents indicate the frequency (0 = Never, 1 = One or Twice, 2 = Sometimes, 3 = Often, 4 = Most of the time) with which they experienced each of the behaviors below during the past year (i.e., "during the past 12 months,"). The questions from the WHEQ are presented in Table A15 below.

Table A15 Survey Items from the WHEQ (Waldo, 1999)

Question
Told offensive jokes about lesbians, gay men, or bisexual people (e.g., "fag" or "dyke" jokes, AIDS jokes)?
Made homophobic remarks in general (e.g., saying that gay people are sick or unfit to be parents)?
Ignored you in the office or in a meeting because you are gay/lesbian/bisexual?
Made crude or offensive sexual remarks about you either publicly (e.g., in the office) or to you privately?
Made homophobic remarks about you personally (e.g., saying you were abnormal or perverted)?
Call you a "dyke," "faggot," "fence-sitter" or some similar slur?
Avoided touching you (e.g., shaking your hand) because of your sexual orientation?
Denied you a promotion, raise, or other career advancement because of your sexual orientation?
Made negative remarks based on your sexual orientation about you to other coworkers?
Tampered with your materials (e.g., computer files, telephone) because of your sexual orientation?
Physically hurt (e.g., punched, hit, kicked, or beat) you because of your sexual orientation?
Set you up on a date with a member of the other sex when you did not want it?
Left you out of social events because of your sexual orientation?
Asked you questions about your personal or love life that made you uncomfortable (e.g., why don't you ever date anyone or come to office social events)?
Displayed or distributed homophobic literature or materials in your office (e.g., electronic mail, flyers, brochures)?
Made you afraid that you would be treated poorly if you discussed your sexual orientation?
Implied faster promotions or better treatment if you kept quiet about your sexual orientation?
Made you feel it was necessary for you to pretend to be heterosexual in social situations (e.g., bringing an other-sex date to a company social event, going to a heterosexual "strip" bar for business purposes)?
Made you feel it was necessary for you to lie about your personal or love life (e.g., saying that you went out on a date with a person of the other sex over the weekend or that you were engaged to be married)?
Discouraged your supervisors from promoting you because of your sexual orientation?
Made it necessary for you to "act straight" (e.g., monitor your speech, dress, or mannerisms)

Made you feel as though you had to alter discussions about your personal or love life (e.g., referring to your partner as a “roommate”)
Made unwanted attempts to draw you into a discussion of personal or sexual matters (e.g., attempted to discuss or comment on your sex life)?
Gave you unwanted sexual attention
Made unwanted attempt to stroke or fondle you?

Racialized Sexual Harassment Scale (RSHS; Buchanan, 2005b)

The RSHS captures experiences of racial harassment, sexual harassment, and racialized sexual harassment. The scale consists of 21 items measured on a 5-point scale, with response options ranging from 0 ("never") to 4 ("many times"). The scale can be used in whole or part (e.g., only the subsection for racial harassment). The RSHS was initially used with college and university students and asks about their experiences in the past 12 months. However, since the RSHS's initial development, it has also been used in other contexts, including workplaces. The stem provided in this example reflects one used for higher education; however, it can be modified: "During the time that you have attended your school, were you ever in a situation in which any of your teachers, classmates, advisors, students or staff:"

Table A16. Survey Items from the RSHS (Buchanan, 2005b)

Construct	Question
Sexual Harassment	Said things to insult people of your *gender* (for example, saying women/men aren't good at a particular job)
	Told jokes or stories that described people of your *gender* negatively
	Displayed pictures or cartoons that showed people of your *gender* negatively
	Made comments about your body that emphasized your *gender* (for example, comments about the size of your breasts or penis)
	Made comments about your clothing/accessories emphasizing your *gender* (for example, women wear skirts that are too revealing)
	Said they expected you to behave certain ways because of your *gender* (for example, as a man, expected you to always control your emotions, or for a woman, expected you to wear make-up or smile a lot)
Racial Harassment Subscale	Said things to insult people of your *ethnicity* (for example, saying people of your race/ethnicity can't handle certain jobs)
	Told jokes or stories that described people of your *ethnicity* negatively
	Displayed pictures or cartoons that showed people of your *ethnicity* negatively
	Called you insulting names that referred to your *ethnicity* (for example, "nigger," "spic," "cracker," "white trash," "chink," etc.)
	Made comments about your body that emphasized your *ethnicity* (for example, comments on "kinky" hair, "slant eyes," skin color)
	Made comments about your clothing/accessories emphasizing your *ethnicity* (e.g., covering one's head with a scarf/turban, associating certain brands or styles of clothing with your ethnic group)
	Said they expected you to behave certain ways because of your *ethnicity* (for example, as a Black person that you would be loud and rude, as a Latino speak poor English, as an Asian person that you would like math, or as a White person that you would be prejudiced, etc.)
Racialized Sexual Harassment	Said things to insult people of your *gender and ethnicity* (for example, Black women are rude, Asian men are wimpy, Latino men are violent, White women are dumb, etc.)
	Told jokes or stories that described people of your *gender and ethnicity* negatively
	Displayed pictures or cartoons that showed people of your *gender and ethnicity* negatively

	Called you insulting names that referred to your gender and ethnicity (“black bitch,” “white whore,” “geisha,” “mamacita”)
	Made comments about your body that emphasized your *gender and ethnicity* (for example, for Black women, comments about one’s “Black ass,” for Black men “large penis,” “skinny white bitch,” etc.)
	Made comments about your clothing/accessories emphasizing your *gender and ethnicity* (for example, for Black women, comments about clothing with animal prints or lots of color, traditional dress for women of your ethnicity, for Black/Latino men, comments on gold necklaces/chains)
	Said they expected you to behave certain ways because of your *gender and ethnicity* (for example, expected you as a Black or Latina woman to wear inappropriate clothes, expected you as an Asian man to be self-controlled and disciplined, as an Asian woman to try to please others, as a Latino man that you would be unfaithful in relationships, etc.)

Technology-Facilitated Sexual Violence Victimization Scale (Powell & Henry, 2019)

The Technology-Facilitated Sexual Violence Victimization Scale was developed to capture online gendered violence across four constructs: digital sexual harassment, image-based sexual abuse, sexual aggression and/or coercion, and gender/sexuality-based harassment. The scale consists of 21 dichotomous items, was reviewed by external subject matter experts for face validity, and has previously exhibited high internal consistency. Respondents were asked if any of the following experiences happened to them during their lifetime, with the response options "yes" or "no."

Table A17. Survey Items from the Technology-Facilitated Sexual Violence Victimization (Powell & Henry, 2019)

Construct	Question
Digital Sexual Harassment	Sexually harassed you
	Unwanted sexually explicit images, comments, emails, or text messages
	Partner has checked up on location/activities multiple times a day
	Repeated and/or unwanted sexual requests online or via email or text message
	Partner gained access to your emails or other online accounts without permission
	Publicly posted online an offensive sexual comment about you
	Posted personal details online saying you are available to have sex
Image-Based Sexual Abuse	Nude or semi-nude image taken without permission
	Nude or semi-nude image posted online/sent onto others without permission
	Nude or semi-nude image threat to post online/send onto others
Sexual Aggression and/or Coercion	Image/video of an unwanted sexual experience taken
	Image/video of an unwanted sexual experience posted online/sent onto others
	Image/video of an unwanted sexual experience threat to post online/send onto others
	Unwanted sexual experience with someone met online
	Unwanted sexual experience with someone met dating site/app
Gender/Sexuality-Based Harassment	Gender-based offensive and/or degrading messages, comments, or other content
	Sexuality or sexual identity-based offensive and/or degrading messages, comments, or other content
	Gender-based offensive and/or degrading messages, comments, or other content in virtual world
	Sexually violent threats
	Described or visually represented unwanted sexual act against your avatar or game character
	Described or visually represented unwanted sexual act against you using an online/email/messages

B Appendix B: Message Board Methodology Overview and Key Themes

Appendix B provides a high-level overview of the message board methodology and the key themes we identified in coding the message board responses. Although providing a message board methodology and findings report is outside of the scope of this effort, this appendix is intended to provide an overview of the approach. In addition to this brief description, Appendix C provides the full message board protocol. Fors Marsh also provided the full message board transcripts to the National Center for Science and Engineering Statistics (NCSES) for their records for future thematic coding efforts if desired.

Message Board Methodology Overview

Participants were recruited using a third-party vendor and paid $75 to answer questions on workplace culture and gender relations in STEM as well as to gauge concerns or considerations regarding fielding questions about sexual harassment among STEM students and professionals (see Appendix C for complete message board protocol). Participants were screened by the third-party vendor for project fit and assigned to a board based on their status (i.e., student vs. professional), and gender identity (woman vs. man vs. sexual or gender minority [SGM]). Emphasis was placed on recruiting racial/ethnically diverse participants and SGMs (see Tables B2, B3, and B4 for complete participant demographics). All participants answered a series of questions spaced out over a week (see Appendix C for complete message board protocol). Every message board was assigned a primary and alternate moderator who observed their assigned boards during data collection to ensure participants stayed focused on the questions and respectful in their discourse. Table B1 presents the fielding windows for each of the message boards.

Table B1. Message board fielding windows

Dates	Message Board
17–21 OCT 2022	Woman Graduate Students
	Early-Career Women
	Mid-Career Women
	Late-Career Women
24–28 OCT 2022	Men Graduate Students
	Men Early Career
	Men Mid-Career
	Men Late Career
31 OCT–4 NOV 2022	Women Undergraduate Students
	Men Undergraduate Students
	SGM Students
	SGM Professionals

Message Board Participant Demographic Overview

In this section, we provide an overview of message board participant demographics. Table B2 provides the racial/ethnic demographics and fields of study for the student message board participations. Table B3 provides demographic information and fields of employment for professional message board participants. Table B4 provides demographic and field of study/employment information for SGM message board participants.

Table B2. STEM Student Message Board Participant Demographics and Fields of Study

	Undergraduate		Graduate		Total (n = 58)
	Women (n = 14)	**Men (n = 13)**	**Women (n =16)**	**Men (n = 15)**	
Race					
White	10 (71.43%)	10 (76.92%)	8 (50.00%)	9 (60.00%)	37 (63.79%)
Black/African American	3 (21.43%)	1 (7.69%)	3 (18.75%)	1 (6.67%)	8 (13.79%)
Asian	1 (7.14%)	2 (15.38%)	3 (18.75%)	4 (26.67%)	10 (17.24%)
Not Listed: Write In[16]	0 (0.00%)	0 (0.00%)	1 (6.25%)	1 (6.67%)	2 (3.45%)
Two or More Races	0 (0.00%)	0 (0.00%)	1 (6.25%)	0 (0.00%)	1 (1.72%)
Hispanic, Latino, or of Spanish Origin					
Yes	4 (28.75%)	2 (15.38%)	1 (6.25%)	2 (13.33%)	9 (15.52%)
No	10 (71.43%)	11 (84.62%)	15 (93.75%)	13 (86.67%)	49 (94.48%)
Field of Study					
Engineering	1 (7.14%)	2 (15.38%)	0 (0.00%)	1 (6.67%)	4 (6.90%)
Biological, Agricultural, and Environmental Life Sciences	2 (14.29%)	1 (7.69%)	2 (12.50%)	2 (13.33%)	7 (12.07%)
Social Science	1 (7.14%)	0 (0.00%)	2 (12.50%)	1 (6.67%)	4 (6.90%)
Physical Sciences, Geosciences, Atmospheric, and Ocean Science	1 (7.14%)	0 (0.00%)	1 (6.25%)	0 (0.00%)	2 (3.45%)
Health	3 (21.43%)	1 (7.69%)	7 (43.75%)	6 (40.00%)	17 (29.31%)
Psychology	3 (21.43%)	0 (0.00%)	4 (25.00%)	0 (0.00%)	7 (12.07%)
Computer and Information Sciences	1 (7.14%)	4 (30.77%)	0 (0.00%)	1 (6.67%)	6 (10.34%)
Mathematics and Statistics	2 (14.29%)	2 (15.38%)	0 (0.00%)	2 (13.33%)	6 (10.34%)
Other: Write In	0 (0.00%)	3 (23.08%)	0 (0.00%)	2 (13.33%)	5 (8.62%)

Table B3. STEM Professionals Message Board Participant Demographics and Fields of Employment

	Early Career		Mid-Career		Late- Career		Total (n = 92)
	Women (n = 17)	**Men (n = 16)**	**Women (n = 15)**	**Men (n = 16)**	**Women (n = 14)**	**Men (n =14)**	
Race							
White	7 (41.18%)	9 (56.25%)	6 (40.00%)	7 (43.75%)	11 (78.57%)	9 (64.29%)	49 (53.26%)
Black/African American	6 (35.29%)	4 (25.00%)	4 (26.67%)	2 (12.50%)	3 (21.43%)	1 (7.14%)	20 (21.74%)
Asian	3 (17.65%)	3 (18.75%)	4 (26.67%)	4 (25.00%)	0 (0.00%)	0 (0.00%)	14 (15.22%)

[16] The graduate woman self-identified as Hispanic, and the graduate man self-identified as Central Asian.

Native Hawaiian or other Pacific Island	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (6.25%)	0 (0.00%)	0 (0.00%)	1 (1.09%)
Not Listed: Write In[17]	1 (5.88%)	0 (0.00%)	0 (0.00%)	1 (6.25%)	0 (0.00%)	2 (14.29%)	4 (4.35%)
Two or More Races	0 (0.00%)	0 (0.00%)	1 (6.67%)	1 (6.25%)	0 (0.00%)	2 (14.29%)	4 (4.35%)
Hispanic, Latino, or of Spanish Origin							
Yes	2 (11.76%)	4 (25.00%)	0 (0.00%)	2 (12.50%)	0 (0.00%)	1 (7.14%)	9 (9.78%)
No	15 (88.24%)	12 (75.00%)	15 (100%)	16 (100%)	14 (100%)	14 (100%)	92 (100%)
Field of Employment							
Engineering	2 (11.76%)	4 (25.00%)	3 (20.00%)	4 (25.00%)	3 (21.43%)	2 (14.29%)	18 (19.57%)
Biological, Agricultural, and Environmental Life Sciences	0 (0.00%)	0 (0.00%)	1 (6.67%)	4 (25.00%)	1 (7.14%)	0 (0.00%)	6 (6.52%)
Social Science	2 (11.76%)	2 (12.5%)	1 (6.67%)	0 (0.00%)	0 (0.00%)	1 (7.14%)	6 (6.52%)
Physical Sciences, Geosciences, Atmospheric, and Ocean Science	1 (5.88%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (7.14%)	2 (2.17%)
Health	7 (41.18%)	5 (31.25%)	3 (20.00%)	4 (25.00%)	6 (42.86%)	3 (21.43%)	28 (30.43%)
Psychology	2 (11.76%	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (7.14%)	0 (0.00%)	3 (3.26%)
Computer and Information Sciences	0 (0.00%)	3 (18.75%)	4 (26.67%)	3 (18.75%)	2 (14.29%)	6 (42.86%)	18 (19.57%)
Mathematics and Statistics	0 (0.00%)	1 (6.25%)	2 (13.33%)	0 (0.00%)	1 (7.14%)	0 (0.00%)	4 (4.35%)
Other: Write In	3 (17.65%)	1 (6.25%)	1 (6.67%)	1 (6.25%)	0 (0.00%)	1 (7.14%)	7 (7.61%)

Table B4. STEM SGM Student and Professionals Message Board Participant Demographics and Fields of Study/Employment

	Students (*n* = 11)	**Professionals (*n* = 16)**	**Total (*n* = 27)**
Gender			
Women	4 (36.36%)	6 (37.50%)	10 (37.04%)
Man	4 (36.36%)	9 (56.25%)	13 (48.15%)
Non-binary	1 (9.09%)	0 (0.00%)	1 (3.70%)
Prefer to Self-Describe	0 (0.00%)	1 (6.25%)	1 (3.70%)

[17] Early career woman self-identified as Hispanic. Mid-career man self-identified as Arab. One late career man self-identified as Hispanic and the other self- identified as Multi Racial.

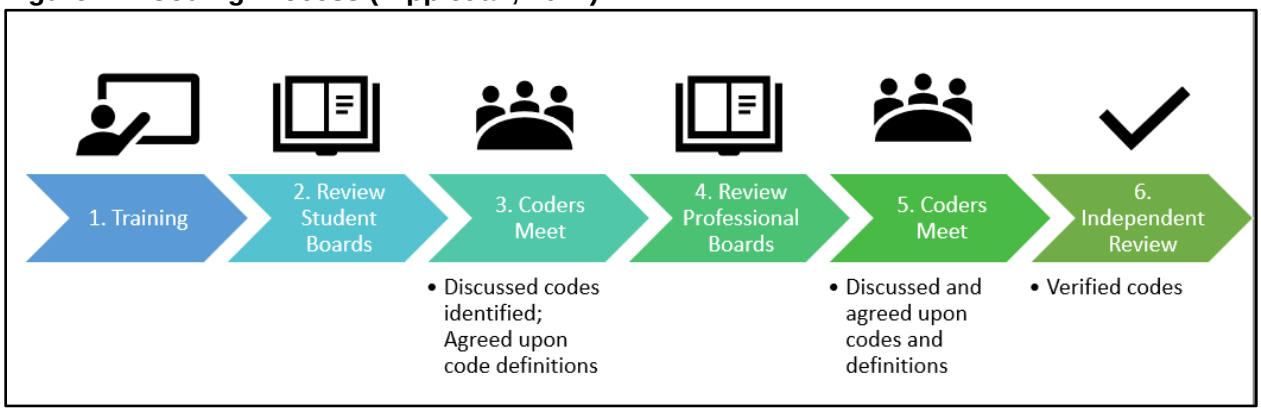
Prefer not to say	2 (18.18%)	0 (0.00%)	2 (7.41%)
Transgender			
Yes	0 (0.00%)	0 (0.00%)	0 (0.00%)
No	9 (81.82%)	16 (100.00%)	25 (92.59%)
Self-Describe	2 (18.18%)	0 (0.00%)	2 (7.41%)
Sexual Orientation			
Bisexual	6 (54.55%)	7 (43.75%)	13 (48.15%)
Gay	2 (18.18%)	6 (37.50%)	8 (29.63%)
Lesbian	0 (0.00%)	1 (6.25%)	1 (3.70%)
Queer	1 (9.09%)	0 (0.00%)	1 (3.70%)
Pansexual	2 (18.18%)	0 (0.00%)	2 (7.41%)
Other	0 (0.00%)	2 (12.50%)	2 (7.41%)
Race			
White	6 (54.55%)	8 (50.00%)	14 (51.85%)
Black/African American	0 (0.00%)	4 (25.00%)	4 (14.81%)
American Indian or Alaska Native	0 (0.00%)	1 (6.25%)	1 (3.70%)
Asian	2 (18.18%)	1 (6.25%)	3 (11.11%)
Two or More Races	3 (27.27%)	2 (12.50%)	5 (18.52%)
Hispanic, Latino, or of Spanish Origin			
Yes	2 (18.18%)	1 (6.25%)	3 (11.11%)
No	9 (81.82%)	15 (93.75%)	24 (88.89%)
Fields of Study/Employment			
Engineering	1 (9.09%)	3 (18.75%)	4 (14.81%)
Biological, Agricultural, and Environmental Life Sciences	2 (18.18%)	3 (18.75%)	5 (18.52%)
Social Science	2 (18.18%)	0 (0.00%)	2 (7.41%)
Physical Sciences, Geosciences, Atmospheric, and Ocean Science	1 (9.09%)	1 (6.25%)	2 (7.41%)
Health	1 (9.09%)	3 (18.75%)	4 (14.81%)
Psychology	2 (18.18%)	1 (6.25%)	3 (11.11%)
Computer and Information Sciences	1 (9.09%)	2 (12.50%)	3 (11.11%)
Mathematics and Statistics	1 (9.09%)	1 (6.25%)	2 (7.41%)
Other: Write In	0 (0.00%)	2 (12.50%)	2 (7.41%)

Thematic Analysis

Once all of the message boards closed, trained coders reviewed the transcripts using the thematic analysis process described below. Thematic analysis is a method used to identify themes or patterns from qualitative data, such as focus groups, interviews, or qualitative message boards (Lochmiller, 2021). For this effort, we used thematic analysis to identify themes in message board transcripts focused on the perceptions of sexual harassment in STEM among STEM students and professionals. Themes identified as part of this effort informed this implementation plan. We followed

standard thematic coding protocol where coders first familiarized themselves with the data set and generated their initial codes individually (Braun and Clarke, 2006) and then iteratively reviewed and processed the message board data together (see Hipp et al., 2017).

Figure B1. Coding Process (Hipp et al., 2017)



We first coded the STEM student boards. Coders were each assigned student message boards to review and generate initial codes, then coders reconvened to discuss observations and created a coding scheme. Once generated, coders applied the coding scheme to their assigned transcripts. Coders met periodically during the coding process to reach consensus on codes and definitions within the transcripts. Once student boards were coded, coders reviewed and coded the STEM professional boards. After all transcripts were coded, coders discussed and consolidated findings. Once completed, independent reviewers who did not work on this project conducted an independent review on half of the message board transcripts to provide an assessment (Hipp et al., 2017).[18] The independent review confirmed the coding scheme. Categories and themes were identified from the coding scheme, mirroring the process followed by Hipp et al. (2017) and the recommendations put forward by Braun and Clarke (2006). Themes were identified first by population (i.e., student or professional) and then compared across population. Major themes were defined as topics that were identified by the majority of participants in both the student and professional boards. Table B5 provides the major themes identified with brief descriptions of those themes and how they informed the implementation plan.

[18] After a review of the literature and discussions with qualitative experts, it was decided that an independent review of half the coded message boards would produce the same results as a full review.

Table B5. High-Level Description of Message Board Themes

Theme	Description	Implications
Sexual harassment is unwanted.	Acts of sexual harassment are not desired by or are forced upon the victim without their consent.	Survey questions about sexual harassment may cause discomfort, and therefore it is important to consider how questions of this sensitive nature would fit within a larger survey on an unrelated topic.
Sexual harassment is uncomfortable.	When someone is being sexually harassed, it causes discomfort, which can interfere with their feelings of safety and satisfaction in the workplace.	
There is a sexual nature to sexual harassment.	Sexual harassment is a form of harassment that targets the individual's sexuality through comments or acts that are sexual in nature or are about their sexual/gender identity.	
Sexual harassment includes actions that reinforce heterosexist work view/demean others based on sexual orientation/identity.	One form of sexual harassment identified by respondents was biased behaviors and discrimination against others based on their sexual orientation or gender identity.	Constructs related to discrimination based on sexual orientation/gender identify should be considered for inclusion in the survey. Additionally, individuals who identify as SGMs would be an important population of interest for understanding sexual harassment in STEM.
Online/social media is a high-risk environment.	Message board respondents consistently identified social media platforms as a location where sexual harassment is more likely to occur.	It would be useful to include survey constructs specifically about forms of harassment that may have been experienced over social media or in social settings outside of work.
Social/non-work sites, especially settings where there is alcohol, are high-risk environments.	Another location where respondents indicated sexual harassment was more likely to occur was social settings outside of work, such as bars, happy hours, or parties.	
People would take the survey to make an impact.	One consistent motivator for responding to the survey is the desire to contribute to research that will lead to positive change, inform prevention practices, and reduce negative outcomes.	These themes provide helpful understanding of the motivations that could impact survey participation and can help inform the communication strategy, especially during recruitment. For instance, this highlights the importance of communicating how the results of the survey will be used.
People would take the survey to promote awareness of the issue.	Another motivator for completing the survey is to help provide data that will illuminate the prevalence and effects of sexual harassment in STEM for a wide audience.	
People would hesitate to the take survey because of lack of impact.	One reason people would not want to take the survey is because of concerns about how the data/results will be used and whether they will lead to any meaningful change.	
People would hesitate to take the survey because of anonymity/confidentiality concerns.	Another reason people would not want to participate in the survey is if they thought their responses to sensitive questions might be linked back to them and shared publicly.	It is highly important that survey respondents' privacy and anonymity is protected and that this is communicated during survey recruitment as well as contained within the survey itself.

Fear over retaliation would prevent people from participating in the survey.	A concern people had about taking the survey was whether information they reveal in their responses might lead to negative consequences for them, especially in their career.	
Most respondents would not hesitate to participate in the survey.	In general, the message board respondents did not have reasons they would not want to participate in the survey.	It is a positive indicator for survey recruitment that the target population is generally not hesitant to participate.

In assessing the message board themes, the team coded with the goal of identifying themes to inform the implementation plan. The message board responses are a rich data source and have further potential utility toward providing knowledge related to climate, culture, and harassment among STEM students and professionals. Fors Marsh provided the full message board transcripts to NCSES that could be further analyzed in future efforts.

Appendix C: Message Board Protocol

Appendix C provides the protocol used for the message boards. The protocol was provided by the Fors Marsh team to the vendor supporting the message boards and then that protocol was uploaded by the vendor onto their message board platform.

NCSES Discussion Guide

TYPE: QualBoard 4.0
PROJECT TOPIC: Workplace Culture & Gender Relations in STEM

PROJECT LEVEL PROGRAMMING:

QUALBOARD 4.0® SAMPLE GROUP DISCUSSION GUIDE
GROUP DISCUSSION EVENT:
PROJECT TOPIC: Workplace Culture & Gender Relations in STEM

GROUP DISCUSSION TOPIC: Workplace Culture & Gender Relations in STEM

SETUP NOTES

PROJECT TITLE: Workplace Culture & Gender Relations in STEM
TIMEZONE: Eastern Standard Time
PROJECT OPEN TIME: Approximately October 10, 2022 – October 28, 2022 (4 sessions during 10/10-10/14; 4 sessions during 10/17-10/21; and 4 sessions during 10/24-10/28)
PROJECT END TIME: Approximately October 28, 2022
PARTICIPANTS: 15 per group (12 groups -- 180 total participants)

QUALBOARD TYPE: QualBoard Group Discussion Event

EVENT START: 10/10; 10/17; 10/24

EVENT END: 10/14; 10/21; 10/28

Programming Instructions -- Schlesinger Group
We indicated where language will need to change per group in green highlight. Much of the information we are asking is the same across group;, however, the professional groups language will refer to the group as "professionals," whereas the student groups are referred to as "students." Additionally, we have two versions of section 2—one for the professional group and one for the student group. We indicate which group by added the population in green highlight next to the question type. Last, there is a question that is asked to only the sexual and gender minority (SGM) groups (Section 3, Question 7) and is indicated by the green highlight.

In total, we will have 12 separate boards:

1. Undergraduate Men in STEM
2. Undergraduate Women in STEM
3. Graduate Men in STEM
4. Graduate Women in STEM
5. Early-Career Men in STEM
6. Early-Career Women in STEM
7. Mid-Career Men in STEM
8. Mid-Career Women in STEM
9. Late-Career Men in STEM
10. Late-Career Women in STEM
11. Sexual and Gender Minority Professionals in STEM
12. Sexual and Gender Minority Students in STEM

SECTION 1: Welcome & Introductions. Day 1, 7:00AM EST
QUESTION 1 - NOTICE ONLY

Question Title: Introduction, Expectations, & Ground Rules

Hello, thank you for taking the time to participate in the [Undergraduate Men in STEM/Undergraduate Women in STEM/Graduate Men in STEM/Graduate Women in STEM/Early-Career Men in STEM/ Early-Career Women in STEM/Mid-Career Men in STEM/Mid-Career Women in STEM/ Late-Career Men in STEM/ Late-Career Women in STEM/Sexual and Gender Minority Professionals in STEM/ Sexual and Gender Minority Students in STEM] Message Board. If you think you may be in the wrong group, please reach out to CONTACT INFO and we will correct the issue.

This discussion board is being moderated on behalf of the National Science Foundation's National Center for Science and Engineering Statistics. Our discussion will focus on your perspectives as [professionals/students] in science, technology, engineering, and mathematics (STEM), specifically on issues related to sexual harassment and gender discrimination. Your perspectives and opinions gathered from this forum will help inform future efforts to better understand these behaviors in the STEM field and how to best measure these experiences, so we encourage you to candidly share your thoughts and opinions on today's topics.

Here are some important things to know as we get started:

1. Questions for each session will be posted at 7:00AM EST and should take around 30 minutes to complete. Please respond to each discussion question openly and honestly as we are interested in understanding your experiences in STEM. You do not have to answer any question you do not feel comfortable answering, however we encourage your full participation. We encourage you to share your perspectives and reasoning-- please note that one-word responses, unless requested, are not typically helpful in these discussions. Board moderators will also be responding to your comments, so look for replies from them as well.
2. This is intended to be a conversation among peers. You will be asked to respond to and interact with others' responses. There are no wrong answers, and we would like hear everyone's perspectives on the topics we discuss, even if they differ from what someone else has said. In the spirit of this, please remain respectful in your responses to each others' opinions during our discussion.
3. What we talk about on the message board is confidential. That means that you will not be personally identified in any of the summary reports or other materials we might prepare based on our discussions here. Please respect the privacy of everyone on this board and would therefore ask that you please not share any of our discussions with others outside of this group.
4. Over the next few days, we will discuss your experiences as [professionals/students] in STEM, to include a discussion of climate and culture, common behaviors that characterize your [workplace/primary academic department], and your thoughts on a potential survey instrument for STEM students and professionals.

QUESTION 2 – NOTICE ONLY

Question Title: Tech Support & Helpful Tips
If you get stuck or can't get something to work, click the "Help" button in the upper right of the page. Our support staff will take care of you. Here are some helpful tips for you:

- New to all this? Visit this site for some great "Getting Started" information https://help.qualboard.com/docs/participant-help
- Check in at least twice a day--once to answer questions and once to respond to other participants or moderator questions.
- Please be as detailed as possible in your response.
- If you run into any trouble, visit https://help.qualboard.com or click on the "Help" button for assistance and our support staff will take care of you.

You can return to the Dashboard now and answer a discussion question.

QUESTION 3 – TEXT RESPONSE

Question Title: Profile and Introduction
Please update your profile to include:

- Your first name or an alias. For the purposes of maintaining privacy, please do not share any personally identifying information, such your last name, the names of others, or your [place of employment/school].
- If you feel comfortable, please add a profile picture to your profile. To do so:
  - Click on your name in the upper right, and select "Profile"
  - Click on the little pencil above your name (next to the grey head)
  - Find a picture you wish to use and click on it and hit "Open"
  - It should now appear above your name on your profile page
  - Click the "Save and Update" button on the profile page

If you run into any trouble, visit https://help.qualboard.com or click on the "Help" button for assistance.
To get started, we'd like to know a bit more about you. Please tell us:

1. Your first name or an alias and preferred pronouns
2. Something you like to do in your free time
3. What factors influenced you to want to pursue STEM?

SECTION 2 (STUDENT GROUPS) - DAY ONE 7:01AM EST

DAY 1 QUESTION 1 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Climate & Culture

First, we would like to ask about the climate and culture in your primary academic department. When we say primary academic department, we mean the academic department in which you are receiving your main degree. As you answer today's questions, please do not share any personally identifying information, such your last name, the names of others, or your [place of employment/school] to help maintain privacy.

What comes to mind when you hear the words climate and culture?
[Possible probes: How does climate differ from culture? What does [participant response] mean to you?]

Note to moderator: If no participant(s) answer or if they are unsure of what we mean by climate and culture, please provide the following high level definitions.

Culture: The overall values and beliefs of the organization or university
Climate: Perceptions and experiences of individuals within the organization or university

DAY 1 QUESTION 2 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Positive Climate & Culture

What are some positives that come to mind when you think of *your primary academic department*'s climate and culture?
[Possible probe: When you wrote [participant term], what do you mean?]

DAY 1 QUESTION 3 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Negative Climate & Culture
What are some negatives that come to mind when you think of *your primary academic department*'s climate and culture?
[Possible probe: When you wrote [participant term], what do you mean?]

DAY 1 QUESTION 4 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Influences on Climate & Culture
What factors influence the climate and culture in your primary academic department?

[Possible probe: When you wrote [participant term], what do you mean?]

DAY 1 QUESTION 5 -TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Differences in Climate & Culture Part 1
How important are social identities (e.g., gender, race/ethnicity, and sexual orientation) in how people experience the culture in your primary academic department?
[Possible probe: When you wrote [participant term], what do you mean?]

DAY 1 QUESTION 6 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Day 1, Social Identities in Climate & Culture Part 2
How do you think your experience with the culture in your primary academic department compares to those of other social identities (e.g., gender, race/ethnicity, and sexual orientation)?
[Possible probe: When you wrote [participant term], what do you mean?]

DAY 1 QUESTION 7 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Environment Differences in Climate & Culture

What role does the physical environment (e.g., research labs, field work, conferences, or the classroom) play in the culture in your primary academic department?
[Possible probes: What environments might be described more in a negative way than others? What environments might be described in a more positive way than others?]

DAY 1 QUESTION 8 – TEXT RESPONSE, STUDENT GROUPS

Question Title: Thank you!

Thank you for your contributions today! We have learned a lot from you and hope that you have learned from others in the group as well. Please log in later today to see any posts that you may have missed while away and to add to what others have to say in the discussion.

We may ask you to clarify or build on your answers as well. Please look for our replies to your posts and follow up when we ask additional questions. This is optional, but if you have any additional thoughts from today's discussion, please let us know below!

SECTION 2 (PROFESSIONAL GROUPS) - DAY ONE. 7:01AM EST

DAY 1 QUESTION 1 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Climate & Culture

First, we would like to ask about climate and culture in the workplace. When we say workplace, we mean the higher-level organization you work for, which could include companies, universities, organizations, or agencies. We are also going to ask you questions about your more immediate work environment, but let's start with focusing on the higher-level organization you work for. As you answer today's questions, please do not share any personally identifying information, such your last name, the names of others, or your [place of employment/school] to help maintain privacy.

What comes to mind when you hear the words climate and culture?
[Possible probes: How does climate differ from culture? What does [participant response] mean to you?]

Note to moderator: If no participant(s) answer or if they are unsure of what we mean by climate and culture, please provide the following high level definitions.

Culture: The overall values and beliefs of the organization or university
Climate: Perceptions and experiences of individuals within the organization or university

DAY 1 QUESTION 2 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Positive Workplace Climate & Culture

Now that we've discussed the differences between climate and culture, what are some positives that come to mind when you think of your *workplace's* climate and culture?

As a reminder, by workplace, we mean the higher-level organization you work for which could include companies, universities, organizations, or agencies.

DAY 1 QUESTION 3 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Negative Workplace Climate & Culture

What are some negatives that come to mind when you think of your *workplace's* climate and culture?

As a reminder, by workplace, we mean the higher-level organization you work for which could include companies, universities, organizations, or agencies.

DAY 1 QUESTION 4 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Work Division Positive and Negative Climate & Culture

We would now like you to think about your *work division's* climate and culture. By work division, we mean the work environment in which you interact with on a regular basis, which could include divisions, departments, or units.

What are some positives that come to mind when you think of your *work division's* climate and culture? What are some negatives that come to mind when you think of your *work division's* climate and culture?

DAY 1 QUESTION 5 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Work Division Comparison

How would you compare the climate and culture of your *work division* to the climate and culture of your overall workplace?

DAY 1 QUESTION 6 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Social Identities in Climate & Culture Part 1

How important are social identities (e.g., gender, race/ethnicity, and sexual orientation) in how people experience the culture in your work division or workplace?

[Possible probe: You wrote [participant term], can you say more about that?]

DAY 1 QUESTION 7 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Social Identities in Climate & Culture Part 2

How do you think your experience with the culture in your work division or workplace compares to those of other social identities (e.g., gender, race/ethnicity, and sexual orientation)?

[Possible probe: You wrote [participant term], can you say more about that?]

DAY 1 QUESTION 8 -TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Day 1, Environmental Differences in Climate & Culture

How might climate and culture, either at the workplace or work division level, differ across certain environments (e.g., other work divisions, other corporate locations, off-site work, conferences, etc.)?

[Possible probes: What environments might be described more in a negative way than others? What environments might be described in a more positive way than others?]

DAY 1 QUESTION 9 – TEXT RESPONSE,PROFESSIONAL GROUPS

Question Title: Thank you!

Thank you for your contributions today! We have learned a lot from you and hope that you have learned from others in the group as well. Please log in later today to see any posts that you may have missed while away and to add to what others have to say in the discussion.

We may ask you to clarify or build on your answers as well. Please look for our replies to your posts and follow up when we ask additional questions. This is optional, but if you have any additional thoughts from today's discussion, please let us know below!

SECTION 3: DAY TWO. 7:01AM EST

DAY 2 QUESTION 1 – NOTICE ONLY

Question Title: Welcome Back!

In the previous session you were asked about your [workplace's and work division's / primary academic department's] climate and culture. Today, we're going to talk more about the types of behaviors that occur within your [workplace or work division / primary academic department]. Again, please feel free to contact me with any questions.

As a reminder, to help maintain privacy, please do not share any personally identifying information, such your last name, the names of others, or your [place of employment/school] when answering questions.

Let's begin!

DAY 2 QUESTION 2 – TEXT RESPONSE

Question Title: Day 2, Respectful Behaviors

We'd like to discuss the types of behaviors that you think characterize respect in your [workplace or work division / primary academic department]. What does respect look like at [workplace or work division / primary academic department]? Who gets respect at [workplace or work division / primary academic department] and how is it shown? How, if at all, does this differ by social identity (e.g., gender, race/ethnicity, and sexual orientation)?

[Possible probes: What do behaviors lacking respect in your [workplace or work division / primary academic department] look like? What do you mean by [participant response]?]

DAY 2 QUESTION 3 – TEXT RESPONSE

Question Title: Day 2, Inclusion

What does it mean to be included at [workplace or work division / primary academic department]? Who gets included and how is this shown? How, if at all, does this differ by social identity (e.g., gender, race/ethnicity, and sexual orientation)?

[Possible probes: What does this exclusion look like and how is it shown?]

DAY 2 QUESTION 4 – TEXT RESPONSE

Question Title: Day 2, Problematic Behaviors

We would now like to discuss harassment in the [workplace or work division / primary academic department]. Specifically, we want to focus our discussion on sexual harassment. While sexual assault may also occur in the workplace, the purpose of today's discussion is focused specifically on sexual harassment. How would you define sexual harassment?

[Possible probes: What about gender discrimination? How might these two terms differ?]

Note to moderator: If participant asks for examples of sexual harassment and/or gender discrimination or you are unsure what might qualify, a list of possible behaviors are presented at the end of this document in a table called "List of Possible SH/GD Behaviors" that can be used to provide examples.

DAY 2 QUESTION 5 –TEXT RESPONSE

Question Title: Day 2, Behaviors in the Workplace/Primary Academic Department

Below is a list of possible behaviors that may occur in your [workplace or work division/primary academic department]. What are some behaviors that are *not* included on this list that you consider sexual harassment in the [workplace or division / primary academic department]?

- Telling sexual jokes
- Touching someone
- Making sexual gestures
- Making comments based on gender
- Sharing sexual pictures or videos of themselves
- Sharing sexual pictures or videos of other people
- Making sexual advances
- Talking about their sexual activity
- Asking about sexual activity or preferences
- Making repeated attempts to establish a romantic or sexual relationship with the same person
- Telling offensive jokes about sexual or gender minorities
- Using homophobic names or slurs

Sharing homophobic materials in your office

[Possible probes: What forms of harassment would you consider specific to your field? How does this [new behavior] differ from some of the others listed? What do you mean by [new behavior]?]

DAY 2 QUESTION 6 – TEXT RESPONSE

Question Title: Day 2, Behaviors in the [Workplace/Primary Academic Department]
Without naming specific names, who or which [employee/student] level tends to engage in the behaviors discussed above (e.g., telling sexual jokes, making sexual gestures, asking about sexual activity or preferences, etc.) in your [workplace or division / primary academic department]? By [employee/student] level, we mean, peers, colleagues, students, teacher assistants, professors, interns, research managers, supervisors, directors, administrators, executives, etc.)
[Possible probes: How do various levels of leadership in your [workplace or division / primary academic department] respond to these behaviors or issues?]

DAY 2 QUESTION 7 -TEXT RESPONSE, SEXUAL AND GENDER MINORITY GROUPS ONLY

Question Title: Day 2, Negative Behaviors in the [Workplace/Primary Academic Department]
What negative behaviors, if any, do you encounter in your workplace because of your sexual orientation or gender identity?
[Possible probe: You wrote [participant term], can you say more about that?]

DAY 2 QUESTION 7/8 – MULTIPLE ANSWER

Question Title: Day 2, Locations of Behaviors, Part 1
We'd like to better understand more about the contexts in which you witnessed or experienced the behaviors identified above (e.g., telling sexual jokes, making sexual gestures or advances, asking about sexual activity or preferences, used homophobic names or slurs, etc.).
Listed here are potential locations and settings where these behaviors might occur. Please indicate all the locations where ***you have witnessed or experienced*** these behaviors. There is an "some other place" option in case the location you have in mind is not listed. If you select 'some other place,' please specify the additional locations you have witnessed or experienced these behaviors next to that selection.

Social media/Online	Dorm, on-campus apartment, or other university housing
Labs	Events outside the classroom

Fieldwork Location	Academic conferences
Office	Some other place, please specify:
Classroom or similar space for teaching and/or learning	I have not witnessed or experienced these behaviors.

DAY 2 QUESTION 8/9 – TEXT RESPONSE

Question Title: Day 2, Locations of Behaviors, Part 2

What location or setting do you consider to be the highest risk for these behaviors to occur? What are some reasons that location or setting is high risk?

[Possible probes: How does this location differ from other locations? Without naming specific names, who or which employee/student level might tend to be in this location? When do these behaviors tend to occur in this location? What other locations on this list do you consider to be high risk?]

DAY 2 QUESTION 9/10 – TEXT RESPONSE, PROFESSIONAL GROUPS

Question Title: Thank you!

Again, thank you for your contributions! Please log in later today to see any posts that you may have missed while away and to add to what others have to say in the discussion.

We may ask you to clarify or build on your answers as well. Please look for our replies to your posts and follow up when we ask additional questions. This is optional, but if you have any additional thoughts from today's discussion, please let us know below!

SECTION 4: DAY THREE. INSTRUCTIONS. Dec 2nd, 7:00AM EST

DAY 3 QUESTION 1 – NOTICE ONLY,

Question Title: Final Day Instructions

Over the last two days, we've talked about several experiences that can occur in [the workplace/primary academic departments]. Now, we would like your thoughts about a potential survey questionnaire for students and professionals in science, technology, engineering, and mathematics (STEM).

Survey Concept Description: There is interest in developing a survey to understand STEM workplace and educational experiences. If developed, this survey would measure the incidence and impact of sexual harassment, discrimination, and other harmful workplace behaviors in STEM environments. It would also be used to understand the extent and implications of sexual harassment and discrimination in the workplace and academia, and would be administered to both students and professionals in STEM.

When answering today's questions, keep this description of a survey instrument in mind. As a reminder, please do not share any personally identifying information, such your last name, the names of others, or your [place of employment/school] to help maintain privacy.

DAY 3 QUESTION 2 – TEXT RESPONSE

Question Title: Day 3, Pros of this Survey Instrument

Based upon the survey concept described above, what are some reasons you might be interested in participating?

[Possible probe: You wrote [participant term], can you say more about that?]

DAY 3 QUESTION 3 – TEXT RESPONSE

Question Title: Day 3, Cons of the Survey Instrument
Based upon the survey concept described above, what are some reasons you might be hesitant to participate?

DAY 3 QUESTION 4 – TEXT RESPONSE,

Question Title: Day 3, Survey Contact Method
What would be the best way to contact someone interested in the survey (e.g., personal email, school or work email, by mail or phone call)? When considering this question, reflect on your likelihood of taking the survey based on how you could be contacted.

[Possible probes: What are some ways you would not want to be contacted?]

DAY 3 QUESTION 5 – TEXT RESPONSE

Question Title: Day 3, Time of Year

When would be the best time of year (e.g., winter, spring, summer, fall) to administer this survey? When answering this question, consider how the time of year may impact your likelihood of taking the survey and when you would not like the survey to be administered.
[Possible probes: How does time of year impact your likelihood of taking the survey? When would you not like the survey to be administered?]

DAY 3 QUESTION 6 – TEXT RESPONSE

Question Title: Day 3, Organization Sponsoring the Survey
How might the organization sponsoring the survey impact your likelihood of participating in the survey? For example, if it was sponsored by a government agency vs. a contracting company vs. university.
[Possible probes: What are some reasons why [participant response] would make you more/less likely to participate?]

DAY 3 QUESTION 7 – TEXT RESPONSE

Question Title: Day 3, Capturing Experiences of Sexual Harassment in the Survey
How comfortable would you feel answering questions about sexual harassment in [the workplace/academia] on this survey? When answering this question, consider what types of experiences or behaviors you would like to see captured in the survey.
[Possible probes: What, if any, questions related to sexual harassment would you not want to answer?]

DAY 3 QUESTION 8 – TEXT RESPONSE

Question Title: Day 3, Stand Alone Survey or Part of Existing Survey
Would you prefer a survey measuring sexual harassment and similar behaviors to be a stand-alone survey or included as part of an existing survey (that is, not specifically focused on workplace climate and culture)?
[Possible probes: If questions related to sexual harassment, gender discrimination, and the like appeared on a survey you were participating in that is not directly related to these topics, how would you react?]

DAY 3 QUESTION 9 – MATRIX MULTIPLE CHOICE

Question Title: Day 3, Willingness to Take Survey
If a survey were to field with questions related to sexual harassment and other harmful behaviors…

	Yes	No	Not Sure
1. Would you answer those questions?			
2. Would you take a whole survey about sexual harassment?			

DAY 3 QUESTION 10 – NOTICE ONLY

Question Title: Day 3, Thank You for Your Participation!

You have now completed our pre-set questions for the discussion board!

Again, please check in throughout the day to add your opinion to any new discussions and respond to the moderator's probes. The discussions will remain open until Midnight EDT tonight, DATE. Please remember to not share anything discuss here with anyone outside of this group, to protect everyone's privacy. Your incentive will be sent to your email address within 4-6 weeks. If you do not receive the incentive by DATE, please email EMAIL address.

List of Possible SH/GD Behaviors (Day 2 Question 4)
Note: Below are a list of examples for the moderator to provide to participants if they are unsure of what harassing behaviors look like. This is not an exhaustive list.

Telling sexual jokes
Touching someone in a sexual manner
Making sexual gestures
Making comments based on gender
Sharing sexual pictures or videos of themselves
Making sexual advances
Sharing sexual pictures or videos of other people
Talking about their sexual activity
Asking about sexual activity or preferences
Making repeated attempts to establish a romantic or sexual relationship with the same person
Told offensive jokes about sexual or gender minorities

D Appendix D: Considerations and Recommendations in Surveying Minors

Appendix D describes considerations for including minors in a survey assessing sexual harassment and related constructs. This is primarily a concern because the sampling frame for the National Training, Education, and Workforce Survey (NTEWS) includes individuals ages 16 through 75 in the skilled technical workforce. This appendix discusses—at a high-level—the complexities in including minors in the sampling frame, including the complex legal and research definitions of minors, the varying disclosure requirements for reports of sexual harassment, and the additional privacy considerations that come with sampling minors. Finally, we recommend consulting with NCSES's legal counsel to determine if the sample should include population members under the age of 18.

Considerations

In identifying the target population for a survey of sexual harassment within STEM as STEM students (i.e., undergraduate and graduate) and STEM professionals, the National Center for Science and Engineering Statistics (NCSES) will need to contend with the considerations surrounding the surveying of minors (i.e., individuals under 18 years of age) since individuals falling within this age range have the potential to overlap with the target population. The inclusion of minors in the survey sample—regardless of whether NCSES is leveraging a new survey, fielding a supplemental survey, or adding questions to an existing survey—requires navigating implications for minors in the survey sample. Even if fielding a new survey that includes undergraduate students, it is likely that some members of the sample frame will be under 18 years of age. If leveraging a sample of an existing survey (e.g., the NTEWS) it is possible that minors (both students and non-students) may be in the survey's sample. In fact, NCSES already conducts surveys in which the target population includes minors (e.g., the NTEWS surveys individuals ages 16 through 75). Therefore, we have outlined some of the high-level considerations that may be required to navigate, regardless of which survey option NCSES selects.

The first complication involved with the inclusion of minors in the sampling approach is the definition of *who* is a minor from both a legal and research perspective. Who is legally considered a minor may vary from state to state and may also vary based on the particulars of the research—in some cases, individuals who are 16 or older may be permitted to legally consent to and participate in certain types of research.[19] This may necessitate multiple sets of research protocols and procedures that depend on the state of residence of the sample members.

A second complication involves the varying disclosure requirements depending on the minor's state of residence. In situations where a minor discloses being a victim of sexual misconduct, there may be legal obligations to promptly report these disclosures to law enforcement, and, again, these reporting requirements may vary by state, particularly since a government entity is collecting this data. These legal requirements will need to be fully understood, documented, and monitored as statutes are updated over time. This will necessitate development of clear protocols to govern data collection covering the specific survey items related to experiences of sexual harassment and any

[19] See e.g., *Research with Children FAQs*. (n.d.). HHS.gov. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/children-research/index.html; *Research with Minors | Research Compliance and Integrity*. (n.d.). https://rci.ucmerced.edu/irb/researchers/research-vulnerable-populations/research-minors.

marginal comments or whitemail that may ensue due to the prompting of these survey items.[20] With these protocols in place, it will be necessary to inform minors who are sampled and potentially their parents (if parental consent is necessary) about the types of disclosures that will be reported to law enforcement, which may result in a follow-up investigation or wellness checks.

In situations where sample members are deemed minors, there may be a need to obtain assent from the minor participants, and potentially their parents, in which they are informed that the survey contains items measuring experiences of sexual harassment. Furthermore, obtaining new Institutional Review Board (IRB) approval is likely to be required when modifying existing survey protocols to address these sensitive topics.[21]

If parental consent is necessary in some instances, privacy policies may need to be updated to ensure that the minors' survey response are not disclosed to parents. The privacy policy, possibly including the consent process, will need to clearly communicate the protections in place to prevent parental access of sensitive survey data. If strong protections are not possible, there may be limitations to the types of questions minors can respond to.

Recommendations

Due to the complexities of this issue, we recommend consulting with NCSES's legal counsel to define the legal parameters within which this research project must operate. Additionally, we suggest conducting additional research to thoroughly understand the ethical considerations of measuring this sensitive topic among minors, particularly the associated benefits to the population (dependent upon the adequacy of the sample frame to provide enough cases for reliable measurement) and risks to the individual sample members. Based on this understanding of legal requirements, benefits, and risks, NCSES can determine if the sample should include population members

[20] In fact, this is potentially a consideration for adult respondents as well. We might expect that there would be an increase in whitemail or marginal comments that require some type of follow-up action from the organization fielding the survey. As the survey is developed, NCSES should consider developing a protocol guiding such responses.

[21] Obtaining new approvals from both IRB and the Office of Management and Budget (OMB) is expected—even if adding questions directly to existing surveys—due to the sensitive nature of the additional measures.