



The Inclusion of Gender Minority Communities in
Survey Research: A Report for the Consortium on
Electoral Democracy

Quinn M. Albaugh, Queen's University

Report 2023-10-23

The Inclusion of Gender Minority Communities in Survey Research: A Report for the Consortium on Electoral Democracy

Quinn M. Albaugh

Original Version: August 12, 2021
Published Online: October 23, 2023

Executive Summary

Survey researchers are working on updating the measurement of gender to be more inclusive of transgender, non-binary, Two-Spirit, and gender minority respondents. In December 2020, Consortium of Electoral Democracy (C-Dem) Co-Directors Allison Harell and Laura Stephenson commissioned this report on improving the gender questions on C-Dem surveys. I conducted a review of the measurement of gender in the Canadian Election Study (CES), along with current proposals for improving the measurement of gender in survey questionnaires across disciplines, including political science, psychology, computer science, and health research. I examined survey questions on the 2019 CES and alternative proposals across fields based on three normative standards: (1) dignity for transgender, non-binary, Two-Spirit, and gender minority respondents; (2) ability to identify transgender, non-binary, Two-Spirit, and gender minority respondents; and (3) ability to adapt to future changes in identity labels.

The current C-Dem approach does not meet these three standards. It does not meet the first standard because (1) respondents who are not men or women cannot refer to themselves using the terms that they use for themselves and (2) the “Other” response option literally otherizes respondents who are not men or women (Bauer et al., 2017). It does not meet the second standard because it cannot disaggregate among respondents who are not either men or women and because the 2019 CES gender identity question is likely to undercount transgender men and transgender women. It does not meet the third standard because there is no open-ended option to allow respondents to mention identity labels that may not fit into researchers’ imposed categories. If new identity labels emerge

or gain increased currency in the future, as they have in the past, the CES cannot record and track them over time.

I recommend that C-Dem adopt a new standard for gender questions that includes

- a close-ended gender self-categorization question with an additional “Non-Binary” response option that replaces the “Other” category with an open-ended “Another gender, please specify:” response option,
- a close-ended transgender identity question (“Are you transgender?”), as in the 2020 U.S. Collaborative Multi-Racial Post-Election Survey,
- a separate Two-Spirit identity question (“Are you Two-Spirit?”) asked only of Indigenous respondents (Pruden and Salway, 2020), and
- two separate seven-point masculinity and femininity scales (Gidengil and Stolle, 2020).

In addition, I recommend that C-Dem undertake the following steps to evaluate competing proposals to improve gender questions, including

- running question wording experiments to assess whether close-ended or open-ended gender identity questions perform better (in terms of response times, non-response, and missing data from hostile or bigoted responses) in the context of an online election study fielded by commercial vendors,
- running a question wording experiment on whether it is necessary to include an “explainer” of what transgender means in a question on transgender identity,
- asking gender minority respondents identified in the next CES’s campaign period survey to evaluate proposed gender items in the post-election survey, and
- holding a workshop that includes both C-Dem partners and trans, non-binary, Two-Spirit, and gender minority scholars across disciplines on how to improve gender questions.

These steps will help inform future revisions to gender questions on C-Dem surveys, provide opportunities for C-Dem researchers to contribute to the growing cross-disciplinary literature on the measurement of gender in survey research, and capture more faithfully the lived experiences of gender in society.

Acknowledgements

When I originally wrote this report, I held a postdoctoral fellowship from the University of Toronto and Max Planck Institute for Religious and Ethnic Diversity. I am especially grateful to Elizabeth Baisley and Michael Donnelly for useful conversations about the contents of the report and Elisabeth Gidengil and Elizabeth Goodyear-Grant for their thoughtful reviews of the report before publication. Of course, no one who gave comments or suggestions is responsible for the contents of this report. Finally, I am deeply grateful to all the trans and non-binary people who've shared with me their experiences and thoughts on sex/gender in survey research. I hope that this report will be useful for them. Ultimately, I hope it will contribute to reducing the harms embedded in survey research for trans, non-binary, Two-Spirit, and gender minority respondents.

Contents

- 1 Introduction** 5
- 2 A Note on Terminology** 6
- 3 Standards for Evaluating Gender Questions** 8
- 4 Existing Proposals for Measuring Gender in Surveys** 10
 - 4.1 Interviewer Coding of Respondents' Perceived Gender 11
 - 4.2 Three-Option Close-Ended Gender Identity Questions 12
 - 4.3 An Open-Ended Gender Identity Question 13
 - 4.4 The Two-Step Model 15
 - 4.5 The 3x3 Approach 18
 - 4.6 A Two-Spirit Identity Question 19
 - 4.7 Masculinity and Femininity Scales 20
- 5 Recommendations for Future C-Dem Surveys** 21
- 6 Next Steps in Improving Gender Questions** 24
 - 6.1 Question Wording Experiments 24
 - 6.2 Asking Gender Minority Respondents to Evaluate the Questions 26
 - 6.3 A Workshop on Measuring Gender in Survey Research 27

1 Introduction

Survey researchers increasingly recognize the need for new questions to measure different aspects of respondents' genders.¹ However, there is no agreed-upon standard for measuring gender in survey research. What are the best ways to improve the inclusion of gender minority respondents, including transgender (or trans), non-binary, Two-Spirit, and gender minority people, in survey research? And what additional information do we need to make decisions about improving gender questions?

In this report, I review alternative approaches to measuring gender in survey research across disciplines, including close-ended gender identity questions that include non-binary response options (Medeiros, Forest, and Öhberg, 2020), open-ended gender identity questions (Fraser et al., 2020; Lindqvist, Sendén, and Renström, 2020), the “two-step approach” (Badgett et al., 2014), the “3x3 approach” (Bleischel et al., 2021), separate Two-Spirit identity questions (Pruden and Salway, 2020), and masculinity and femininity scales (Alexander, Bolzendahl, and Öhberg, 2021; Bittner and Goodyear-Grant, 2017a,b; Cassino, 2020; Gidengil and Stolle, 2020, 2021). I assess each of these alternative approaches in relation to three main standards – (1) dignity for gender minority respondents (or **dignity**, for short), (2) the ability to identify gender minority respondents (or **identifiability**), and (3) the prospects for using these questions over the long-run to gather longitudinal data (or **reusability**).

Based on this review of the literature, along with my personal experiences as an active member of trans communities for over a decade, I make a series of recommendations. First, I propose a base set of gender questions for future C-Dem surveys, including

- a close-ended gender identity question that includes an explicit non-binary option and an “another gender, please specify:” option,
- a transgender identity question (“Are you transgender?”),
- a separate Two-Spirit identity question (“Are you Two-Spirit?”) asked only of Indigenous respondents (Pruden and Salway, 2020), and

¹Scholars of gender and political behaviour often draw a sharp distinction between “sex” and “gender” (Bittner and Goodyear-Grant, 2017b; Cassino, 2020; Gidengil and Stolle, 2020, 2021; McDermott and Hatemi, 2011). I avoid making a sharp distinction between sex and gender in this report because these two broad concepts are not fine-grained enough to capture the specific aspects of sex and gender that matter for trans, non-binary, Two-Spirit, and/or gender minority respondents. Instead, I focus on a series of more specific concepts, such as assigned sex at birth, perceived gender, gender identity, gender modality, and transgender identity, that help clarify which aspects of sex and gender are important to measure and how they relate to one another, which I discuss in more detail below. I generally use gender as an umbrella term rather than “sex/gender” to make the report more accessible to a variety of audiences.

- two separate seven-point masculinity and femininity scales (Gidengil and Stolle, 2020).

The full set of questions are available in Table 2. Second, I propose a series of studies for C-Dem that can help resolve issues with the literature on improving gender questions and provide opportunities for C-Dem partners and researchers. Third, I suggest holding a workshop on improving gender questions in survey research. This workshop should include not only C-Dem partners but also trans, non-binary, Two-Spirit, and gender minority researchers with interests in survey methodology across disciplines.

2 A Note on Terminology

Since this report draws on a number of specialized terms that may not be familiar to all readers, this section provides a brief overview of key aspects of sex and gender that are relevant for understanding different types of survey questions and their impacts on gender minority respondents.

At birth, medical professionals, families, governments, and other social institutions and actors coercively categorize all babies as “male” or “female,” largely based on the appearance of babies’ genitalia. Indeed, if babies have “ambiguous” genitalia, medical professionals may push for their genitalia to conform more to normative views of how “male” and “female” genitalia ought to look (Davis, 2015; Karkazis, 2008). **Assigned sex** refers to the category assigned to individuals by these institutions – usually male or female.

In everyday life, individuals learn to interpret others’ appearances as boys or girls or as men or women (Kessler and McKenna, 1985). Individuals go through social spaces categorizing other people into binary gender categories that shape how they interact with those people. **Perceived gender** refers to categories assigned to people by using social cues to ascertain an individual’s gender. Of course, in a society that takes cisgender people as the norm, the usual assumption is that perceived gender and assigned sex are the same. When survey interviewers classify respondents’ genders based on vocal cues in telephone interviews or vocal and visual cues in face-to-face interviews, they are measuring perceived gender – specifically, how the interviewer perceives the respondent’s gender. Not all people in society will necessarily agree about how to classify the same person in the same way.

Gender identity refers to a person’s self-conception of their gender. For example, individuals may consider themselves to be men, women, non-binary, Two-Spirit, or some other gender identity. Agender people may consider themselves as not having a gender identity. Both cis (non-trans) and trans people have gender identities. Gender identities

can be binary (man or woman) or non-binary (neither man nor woman). This view of gender identity draws on social identity theory (Tajfel et al., 1971; Turner, 1975; Turner et al., 1987). If gender identity is a social identity, then it is nearly impossible to understand and talk about it without knowing whether and how an individual self-categorizes their gender identity. Recent work on developing non-categorical gender questions in survey research sometimes refers to individuals' self-placements in relation to the concepts of masculinity or femininity as gender identity. In this report, I follow the general practice in trans communities of using gender identity to refer to individuals' self-categorizations of their gender. These self-categorizations usually are *not* as transgender or trans.² In trans communities, these terms are commonly used as adjectives about someone's gender identity but rarely used as answers to the question, "what is your gender (identity)?" In order to distinguish these self-categorizations from masculinity and femininity self-placements, I refer to the latter as **gender role self-conceptions**.³

Gender modality "refers to how a person's gender identity stands in relation to their gender assigned at birth" (Ashley, 2022, 22). That is, an individual's gender modality can be cis (the same as their gender assigned at birth), trans (different from their gender assigned at birth), or possibly something else entirely (Ashley, 2022). Despite the term's recent coinage, it is gaining currency as a way of discussing the aspect of gender on which cis and trans people differ from one another without treating cisness as normal and unmarked and transness as abnormal and marked. Importantly, people whose identities are different from their assigned sex at birth do not necessarily identify with the term transgender. As a result, it is possible to distinguish between an ascriptive view of gender modality and **transgender identity**. Transgender identity is a self-categorization as transgender, using the same social identity approach as gender identities.

Sexual identity refers to how individuals self-categorize themselves in relation to their sexual and romantic attractions.⁴ Just as cis people can have any sexual identity, trans people can have any sexual identity. For example, trans men can identify as straight, gay, bisexual, or any other sexual identity.

Not all respondents will necessarily accept the distinction between gender and sexual identities. In particular, some **Two-Spirit** people may reject the idea that it is possible to draw a sharp distinction between gender and sexual identities as settler thinking. Accord-

²Since the original writing of this report, Statistics Canada released early results from the 2021 Census of Canada, which asked separate sex at birth and gender identity questions. It provides strong evidence that few individuals use transgender as a primary gender identity term. Instead, about two-thirds of individuals who aren't men or women use the term non-binary to describe themselves (Canada, 2022).

³I acknowledge that some individuals may not place themselves on masculinity and femininity scales purely based on how they relate to gender roles. However, masculinity and femininity generally refer to the social roles associated with men and women—or gender roles.

⁴I do not use the term "sexual orientation" because it is important to distinguish between identity, attraction, and behaviour in studying sexuality (Sexual Minority Assessment Research Team, 2009; Geary et al., 2018; Wolff et al., 2017).

ing to Pruden and Salway (2020), “Two-Spirit is a term coined by Indigenous lesbian, gay, bisexual, transgender, queer, and/or non-heterosexual (LGBTQ+) leaders at the Third Annual Intertribal Native American/First Nations Gay and Lesbian Conference in Winnipeg in 1990.” Two-Spirit can refer to both non-heterosexual and/or non-cisgender Indigenous identities. It may also be a spiritual identity. While not all Two-Spirit people may accept that they are “gender minority” respondents in the Western social scientific sense, some do. As a result, it is important to pay particular attention to Two-Spirit respondents in devising more inclusive gender questions.

3 Standards for Evaluating Gender Questions

The highest standard for evaluating gender questions is that they must respect the **dignity of gender minority respondents** (the Dignity Standard). This standard leads to at least two concrete recommendations. First, researchers should not ask respondents invasive questions about their personal or medical history unless it is relevant to the study. Gender identity, transgender identity, and sexual identity are all relevant for explaining political outcomes, such as vote choice or political attitudes. Questions about assigned sex, hormone use, genitalia, and other medical or “sex” information are not generally politically relevant. Second, researchers should allow respondents to self-identify with terms they actually use. Third, researchers should avoid otherizing language, such as forcing respondents who are neither men nor women into an “other” gender category rather than treating non-binary gender identities as a normal part of life (Bauer et al., 2017).

The next standard for evaluating gender and sexuality questions is that they must be able to **identify gender minority communities** (the Identifiability Standard). If researchers do not include questions that identify these communities, then they are erasing them from research (Namaste, 2000). For example, an important step in addressing the Dignity Standard is allowing trans men and trans women to identify as men and women (respectively). This approach allows them to use the language they use for their gender identities. However, it would not allow survey researchers to study inequalities facing them or analyze their attitudes and behaviours separately from cis men and women, which effectively treats transness as irrelevant to research.⁵ Researchers should pay particular attention to several groups traditionally under-represented in research, including (1) trans men and trans women; (2) people who are neither men nor women (such as non-binary, genderqueer, and agender people); and (3) Two-Spirit people. Researchers need to adopt different approaches to including these communities in survey research and identifying respondents from these communities because they diverge in different ways

⁵Indeed, it is possible that not meeting the Identifiability Standard might also mean necessarily not meeting the Dignity Standard. I thank Elizabeth Goodyear-Grant for making this point.

from the prevailing societal assumptions about gender. Trans men and women people challenge the expectation that assigned sex and gender identity are the same. For that reason, it is impossible to identify trans men and women people with a single gender question. Both gender identity and transgender identity are necessary to identify these people. Non-binary people challenge the presumption that everyone is either a man or a woman. Binary gender categories are inadequate for including non-binary people in survey research. Two-Spirit people challenge the distinction between gender and sexual identity and the presumption that sexual identities are universal across cultures. Two-Spirit identities highlight the importance to be aware of cultural contexts in designing gender and sexual identity questions.

Finally, an ideal set of gender questions for C-Dem would be **reusable over time** (the Reusability Standard). Reusability is important for generating longitudinal data on gender minority respondents, along with allowing for comparisons across provincial and territorial studies. The CES has traditionally provided a long time series of data, and C-Dem has begun conducting provincial and territorial election studies. Unless there are good reasons to change questions – as in this case – it is useful to ask the same question in all C-Dem surveys to allow for comparisons over time and across provinces. For the gender questions to meet this standard, they need to be able to evolve with changing language and understandings of gender. We know that the language individuals use to describe their gender and that understandings of gender can change over time (Stryker, 2017). For example, although people who did not see themselves as either men or women have existed for a long time (Bornstein, 2016), the term non-binary rose in prominence during the 2010s. For this reason, survey questions about gender must be able to adapt over time as new identity labels emerge. If they are not able to capture new identities as they emerge, then survey researchers will have to revise their questions again in the future, which makes it difficult to conduct longitudinal analyses. It is almost certainly impossible to design entirely close-ended response categories that will last through the coming decades (Magliozzi, Saperstein, and Westbrook, 2016).

When there are multiple sets of questions that meet all three standards, then it is reasonable to shift the discussion to matters of survey methodology. For this reason, I discuss whether each proposal meets these standards before discussing matters of response times, non-response, or missing data biases. Given that my review suggests that there may be multiple approaches that meet these three standards, I conclude by proposing a series of C-Dem initiatives that may help in assessing the trade-offs between different approaches for survey methodology.

4 Existing Proposals for Measuring Gender in Surveys

In this section, I discuss different approaches to measuring gender in survey research, including the traditional approach in face-to-face and telephone surveys (interviewer coding) and six recent proposals to improve upon the traditional approach. Each of these proposals has some advantages. A summary of whether each approach meets this standard is in Table 1. However, as Table 1 shows, no single existing approach identifies trans, non-binary, Two-Spirit, and other gender minority respondents at the same time. Ultimately a combination of approaches is likely to be necessary to meet the three standards of interest in this report – dignity for gender minority respondents, ability to identify gender minority communities, and reusability over time.

Table 1: Summary of Evaluation of Approaches to Measuring Gender Based on Dignity, Identifiability, and Reusability Standards

Approach	Dignity	Standard Identifiability	Reusability
Interviewer Coding	Not Met	Not Met	Not Met
Three-Option Close-Ended Gender Identity Question	Not Met	Not Met	Not Met
Open-Ended Gender Identity Question	Met	Not Met	Met
Two-Step Model (Version 1: Assigned Sex + Gender Identity)	Not Met	Not Met	Not Met
Two-Step Model (Version 2: Gender Identity + Transgender Identity)	Met	Not Met	Met
3x3 Approach	Met	Not Met	Met
Two-Spirit Identity Question (for Two-Spirit Respondents only)	Met	Met	Met
Masculinity and Femininity Scales	Met	Not Met	Met

4.1 Interviewer Coding of Respondents' Perceived Gender

Since the early years of survey research, survey interviewers have generally coded respondents' perceived genders – as men or women – based on visual or vocal cues. In 2019, the CES telephone survey still used this approach. However, this method falls short on all three standards outlined above:

Dignity Standard (Not Met) It does not respect the dignity of gender minority respondents because it forces respondents who are not men or women into those categories and does not allow respondents to self-identify at all. This approach misgenders many trans and non-binary respondents, which is psychologically harmful (McLemore, 2015, 2018).

Identifiability Standard (Not Met) It does not allow researchers to identify trans, non-binary, Two-Spirit, and gender minority respondents. As a result, it effectively erases those identities as meaningful topics of study (Namaste, 2000).

Reusability Standard (Not Met) Interviewer coding already does not reflect changes in language and understandings around gender because it treats gender minority respondents as if they do not exist. As a result, interviewer coding is only going to become more out of step with society.

Many of these problems with interviewer coding also take place when researchers use self-reported binary “sex” questions. These approaches share the same problem is slotting everyone into two binary categories. More importantly, however, it is often unclear what a “sex” question actually measures. Sex refers to a variety of differences in bodies – including chromosomes, hormones, primary and secondary sex characteristics, prenatal testosterone exposure, and so forth – that do not all go together.⁶ For trans and intersex respondents, it is often unclear how to respond to a “sex” question, especially one that does not specifically mention assigned sex at birth. Many trans people respond to sex questions with their gender identity rather than their assigned sex (as researchers may assume). In addition, intersex people with similar bodies may provide different answers from one another.

⁶For this reason, it's useful to think of sex as a social construct that categorizes people based on their bodies (Butler, 1990).

4.2 Three-Option Close-Ended Gender Identity Questions

One solution to the problems with binary interviewer coding is to allow three-option close-ended gender identity questions (Medeiros, Forest, and Öhberg, 2020). These questions have become increasingly prevalent in the shift from telephone to online surveys. Online surveys do not have interviewers and, as a result, cannot rely on interviewer coding of respondents' perceived gender. However, these close-ended gender identity questions also present pitfalls in meeting the standards outlined above. In particular, on their own, they cannot meet the dignity and identifiability standards.

The 2019 CES online panel provides a useful example of these pitfalls. The 2019 CES online panel was the first CES survey to offer respondents the opportunity to identify as neither men nor women. It used a three-category gender identity question:

Are you...?

1. A man
2. A woman
3. Other (e.g. Trans, non-binary, two-spirit, etc.)

This question treats these survey responses as mutually-exclusive. For instance, a trans man must choose between selecting “A man” or “Other (e.g. Trans, non-binary, two-spirit, etc.)” It also does not provide an open-ended response for the “Other” response category to allow respondents to specify their identities. It was mandatory for all respondents to the 2019 CES online panel because of the desire to weight the sample on gender identity. Unfortunately, the 2019 CES question does not meet the standards outlined in the previous section.⁷

Dignity Standard (Not Met) The 2019 CES online panel's three-option close-ended gender identity questions typically does not meet the standard of dignity for gender minority respondents for two reasons. First, it does not allow gender minority respondents to use explicit language that reflects their identities. Second, it uses otherizing language for gender minority respondents by lumping them in an “Other” category (Bauer et al., 2017).

⁷While the 2019 CES online panel does not meet these standards, it at least avoids certain pitfalls in designing gender identity questions with more than two categories. For example, some surveys include three-option Male/Female/Transgender questions. Trans respondents do not react well to these response options. For example, in 2016 and 2017, the U.S. Census Bureau conducted focus groups with trans and non-binary people to improve data collection on the Current Population Survey. One focus group participant said, “If I have to take another study where it says ‘male/female/transgender’ I might just rip out all my hair” (Holzberg et al., 2017, 37). I immediately stop responding to any survey when I see these response categories, as do many other trans academics and activists.

Identifiability Standard (Not Met) The 2019 CES online panel’s three-option gender identity question does not meet the standard of identifying gender minority communities. Since the gender categories are not mutually exclusive, this gender identity question forces trans men and women to choose between identifying as trans or identifying as men or women. Trans men and trans women may quite reasonably select “A man” or “A woman” rather than “Other.” Since the 2019 CES online panel did not include a transgender question, trans men and women respondents are largely invisible in the survey responses (Namaste, 2000). While it is very likely that there were trans men and trans women, we do not know who they are or how they answered. The exceptions are the rare few respondents who provided that information in open-ended responses to the sexual identity question. The same problem can arise for Two-Spirit respondents, who may also identify as men or women. This invisibility of trans and Two-Spirit men and women respondents is particularly problematic given that the 2019 CES online panel aims to allow researchers to study small populations, such as gender minority communities (Stephenson et al., 2021).

Reusability Standard (Not Met) The 2019 CES online panel question does not meet the standard of adapting to future changes in language and culture. It is possible that some of the identities explicitly mentioned in the other category will fall out of use in the future.

Ultimately, while additional categories besides “man” and “woman” are certainly necessary and feasible (Medeiros, Forest, and Öhberg, 2020), they are not enough to ensure the inclusion of gender minority communities in survey research (Westbrook and Saperstein, 2015).

4.3 An Open-Ended Gender Identity Question

Other scholars advocate including an open-ended gender identity question, usually “What is your gender?” (Fraser et al., 2020; Fraser, 2018; Lindqvist, Sendén, and Renström, 2020). Open-Ended gender identity questions generally meet the Dignity Standard and the Reusability Standard but fall short on the Identifiability Standard.

Dignity Standard (Met) Of all the possible gender identity questions, an open-ended gender identity question performs the best in allowing respondents the freedom to use language they prefer. When the U.S. Census Bureau conducted focus groups with trans and non-binary people on updating the gender questions on the Current Population Survey, trans and non-binary participants generally preferred this option because it allowed respondents to use language that resonated for them (Holzberg et al., 2017).

Identifiability Standard (Not Met) First, it does not meet the standard of identifying gender minority respondents because it does not measure transgender identity. That is, like the three-category CES question, it may under-count trans men and women respondents who respond as M/male/man or F/female/woman without mentioning their trans identities. If transgender identity is politically relevant, then it is important to identify trans men and women respondents.⁸

Reusability Standard (Met) Of all the possible gender identity questions, an open-ended question is the most well-suited to adapting to changes in language and culture over time, since it does not have preset categories. If the socially understood categories change, respondents can still use the labels they prefer, and researchers can change their coding of the open-ended questions.

There are two main objections to open-ended gender identity questions that rely on claims about survey methodology. The first is that they require researchers to code open-ended responses, which can take some time and expert knowledge. However, recent studies show that it is not very costly to code responses to an open-ended gender identity question (Fraser et al., 2020; Lindqvist, Sendén, and Renström, 2020). With relatively minimal costs in terms of time, C-Dem can provide a ready-to-use coding of the open-ended gender identity questions and guidance on how to use the various items. The second is that online survey respondents may not answer the question cooperatively or even provide hostile or bigoted responses (Jaroszewski et al., 2018; Saperstein and Westbrook, 2021), which create missing data problems. However, it is unclear whether these findings of hostile or bigoted findings apply to the C-Dem surveys. Jaroszewski et al. (2018) found these problems were much more common in surveys fielded through Fantasy Football than through Tumblr in 2016, and Saperstein and Westbrook (2021) found similar results in GfK samples of American adults from 2014 and 2016. Given the differences in the time and geography of the samples, it is unclear how frequent these hostile or bigoted responses would be on C-Dem surveys. As a result, it is difficult to assess the trade-offs between close-ended and open-ended gender identity questions. For this reason, I recommend that C-Dem conduct a study designed to examine the prevalence of hostile and bigoted responses and identify which respondents are likely to provide hostile or bigoted responses.

Even if these methodological critiques make purely open-ended gender identity questions undesirable in practice, they still suggest a need to include an open-ended response category within a close-ended gender identity question. Close-ended gender identity questions that do not allow respondents to use the language they prefer are unlikely to

⁸As with other gender identity questions, it is possible to mitigate this downside by adding an additional transgender identity question after an open-ended gender identity question. I discuss this possibility below.

ever meet the dignity and resuability standards.

4.4 The Two-Step Model

One of the major problems with surveys that rely on a single gender identity question—including both close-ended and open-ended questions—is that many trans men and women people do not answer these questions by self-identifying as trans men or trans women but simply as men and women. When researchers do not include additional questions to identify trans men and women respondents, they may under-count trans people by only counting trans men and women who will go out of their way to identify trans in a gender identity question. Trans activists in Philadelphia developed the “two-step model” to avoid this problem of under-counting. The “two-step model” involves asking at least two gender questions to identify trans respondents. The first step is to ask about assigned sex at birth. The second step is to ask about current gender identity. Researchers then identify transgender respondents by coding whether respondents’ assigned sex matches their current gender identity.

Today, the two-step model can take two forms. The first is the original version developed in Philadelphia, which asks about **assigned sex** and **gender identity**. This version does not meet any of the three standards.

Dignity Standard (Not Met) The original two-step approach poses unnecessary harms to the dignity of gender minority respondents. The two-step approach originally focused on identifying trans people in health-care settings. Health researchers and medical providers typically argue that assigned sex provides relevant information for understanding health. However, there is simply no reason to ask about assigned sex in most political surveys.⁹ Gender identity and transgender identity are relevant for political behaviour as social identities, not as externally enforced categories. If political behaviour scholars accept that it is more appropriate to identify sexual minority respondents based on their identities rather than based on their attractions or behaviours, then we should also accept that we should identify gender minority respondents, based on their identities. In addition, some gender minority respondents view assigned sex questions as invasive (Badgett et al., 2014; Holzberg et al., 2017).

Identifiability Standard (Not Met) The original two-step approach identifies trans respondents by looking for respondents whose gender identities do not “correspond” to their assigned sex at birth. However, some individuals whose gender identities do not

⁹For scholars of biopolitics, it is often more useful to have more specific biological indicators, such as hormone levels.

correspond to their assigned sex at birth reject the idea that they are transgender. For example, many gender-diverse youth do not necessarily see themselves as transgender (Kidd et al., 2021). For this reason, some researchers argue that the original two-step approach (assigned sex + gender identity) is superior on measurement grounds to alternative approaches that include a transgender identity question because it identifies a larger number of transgender respondents, including both people who identify as transgender and people whose gender modalities are transgender (Saperstein and Westbrook, 2021). However, this argument relies on the idea that being transgender is something for researchers to ascribe onto respondents (as in health research) rather than a social and political identity. For the same reasons that political behaviour researchers measure the identities of racial or sexual minority respondents rather than ascriptively categorizing them, we need to measure the identities of gender minority respondents. The alternative risks categorizing people whose gender identities do not correspond to their assigned sex at birth as transgender even though they would not identify as such. That is, it risks inflating the number of respondents categorized as transgender to include people who explicitly reject that identity label. This decision is likely consequential for political attitudes and behaviour. If someone takes on the transgender identity label, that identity is likely to shape the development of their attitudes and behaviours. By contrast, if someone has medically transitioned but does not identify as transgender, they may respond to politics in ways that resemble cisgender men or women.

In addition, if C-Dem includes an assigned sex question, it is possible that many potential gender minority respondents will not participate in C-Dem surveys. If gender minority respondents view assigned sex questions as invasive, they may conclude that the researchers who use those questions do not have their best interests at heart and stop responding. This concern is particularly acute because gender minority communities have had negative experiences interacting with many social institutions, including researchers. Anecdotally, my conversations with other trans and non-binary people indicate that many of them refuse to participate in surveys that do not use inclusive language.

Reusability Standard (Not Met) At first glance, this version of the two-step approach seems well-suited for use over time. In 2018, Statistics Canada adopted this measurement approach as its new standard for measuring sex and gender and used this approach on the 2021 Census, which would make it possible to weight the CES to questions that mirror the 2021 Census questions. However, it is debatable whether the 2021 Census will provide high-quality data on gender minority communities. It uses a single questionnaire for an entire household, rather than private individual survey responses, which makes it likely that it will dramatically under-count gender minority respondents who may not have

disclosed their identities to their families or roommates.¹⁰ In addition, earlier this year, Statistics Canada engaged the public in another round of consultations on improving its standard for measuring sex and gender. If C-Dem adopted a similar approach to Statistics Canada, it may have to change the questions again, which would make it difficult to gather longitudinal data. As a result, this approach might not meet the reusability standard.

The second version of the two-step approach is to ask about **gender identity** and **transgender identity**. This approach comes closer to meeting the dignity, identifiability, and reusability standards than past work.

Dignity Standard (Met) This approach meets the dignity standard because it asks about gender identity and transgender identity in ways that allow respondents to use identity labels they use and does not include invasive questions, such as assigned sex at birth.

Identifiability Standard (Not Met) When researchers combine a gender identity question with a transgender identity question, they are able to identify trans men and trans women. In addition, unlike the original two-step approach, this approach does not categorize respondents as transgender if they do not identify with that language. However, this approach may undercount Two-Spirit people who do not see Two-Spirit identity as strictly a gender identity to write into the open-ended response category. For this reason, it may be necessary to add an additional question on Two-Spirit identity.

Reusability Standard (Met) This combination of questions is likely to be reusable for the foreseeable future. While it is possible that the term transgender will go out of fashion or come under critiques in the future, it is socially well-understood, and another term is likely to take its place.

As a result, I recommend using a version of the two-step approach that includes a transgender identity question rather than an assigned sex question.

¹⁰Although the 2021 Census technically offered an option for individuals to request a separate, private questionnaire, it required calling a telephone help line to request a separate questionnaire code in the mail. Many trans people do not like to make phone calls where they may interact with someone who may misgender them based on their voice, and receiving a separate code by mail may raise unwanted questions about why someone wants to fill out their Census questionnaire in private.

4.5 The 3x3 Approach

In February 2021, a team of psychologists proposed an alternative approach – that measures gender identity, gender modality (as identities), and gender “binarity” in three separate questions (Bleischel et al., 2021). These three questions allow researchers to categorize respondents into a 3x3 table. The first question asks about gender identity using a single open-ended item. The second questions asks about gender modality (allowing responses of transgender, cisgender, or something else). The third questions asks about gender binarity (binary, non-binary, or something else). Of all these options, the 3x3 model allows respondents the most options to self-identify. For example, this approach allows respondents to refuse to classify themselves as either transgender or cisgender.

Dignity Standard (Met) The 3x3 meets the dignity standard because it allows respondents to use their preferred language in describing their gender identities. The additional questions on gender modality and gender binarity may actually be better at ensuring dignity for gender minority respondents than more established approaches because they make it easier for gender minority respondents to report more complex self-understandings of themselves in relation to transgender or non-binary identity labels.

Identifiability Standard (Not Met) The 3x3 approach meets the identifiability standard. In fact, the 3x3 standard arguably performs better than the two-step approach in making it easier to identify respondents who may identify as non-binary, transgender, and women, for example. It remains unclear whether the 3x3 approach would identify Two-Spirit respondents. While they can report their Two-Spirit identities in an open-ended gender identity question, this approach may still under-count Two-Spirit respondents who use other terms when reporting their gender identities to survey researchers because they see Two-Spirit identity as something other than a gender identity.

Reusability Standard (Met) Given that the 3x3 approach uses an open-ended gender identity question, the 3x3 approach is likely to be reusable in the face of changing language around gender. As with the two-step approach, it is likely that the terms transgender and non-binary will be socially understood – especially within gender minority communities – for the foreseeable future. While these terms may fall out of favour, the 3x3 approach is likely to be reusable over time.

Given the relatively little data on how this set of questions performs in practice, it likely requires additional testing to see how different groups respond to these questions in practice. For example, these questions may seem strange to survey respondents, which may

increase response times or produce non-response biases. As a result, I recommend testing the 3x3 approach against other possible models. In addition, I recommend including an additional question to help identify Two-Spirit respondents.

4.6 A Two-Spirit Identity Question

As discussed above, traditional survey questions, along with many of the proposed alternatives, may not allow Indigenous respondents to identify as Two-Spirit. It is important for C-Dem to be inclusive of Indigenous respondents, given the understandable lack of trust many Indigenous people have in academic research. These concerns are particularly important for Two-Spirit respondents who face intersecting systems of marginalization.

Some surveys, such as the 2019 CES online panel, make reference to Two-Spirit identity within a gender identity question. This approach poses two main problems. First, it requires researchers to impose a conception of Two-Spirit identity as either a gender identity or sexual identity that will not work for all Two-Spirit people. Second, it risks the possibility that non-Indigenous people will either appropriate or mistakenly select an Indigenous identity label.

[Pruden and Salway \(2020\)](#) recommend asking a Two-Spirit identity question (“Are you Two-Spirit?”) that is separate from both the gender identity and sexual identity questions. [Pruden and Salway \(2020\)](#) also propose only presenting this question to Indigenous respondents. The risk in offering a Two-Spirit identity option within a standard gender self-categorization question is that non-Indigenous respondents will either mistakenly select the Two-Spirit option or appropriate the term Two-Spirit. The Two-Spirit identity question should therefore come after a full race/ethnicity question *or* after an Indigenous screener early in the questionnaire.

Dignity Standard (Met) This approach allows Two-Spirit respondents to identify as such without forcing them to do so in a gender identity or a sexual identity question, which imposes a settler understanding of gender and sexual identities as analytically separate on Indigenous respondents. More importantly, it helps avoid potential harms to Indigenous communities that may exist in alternative approaches that allow non-Indigenous respondents to appropriate the term.

Identifiability Standard (Met) A Two-Spirit identity question makes it possible to identify Two-Spirit respondents who may also identify with other gender identity terms. This question also avoids over-counting Indigenous respondents by preventing non-Indigenous respondents from appropriating the term.

Reusability Standard (Met) A Two-Spirit identity question is likely to be reusable over time, given that the term has become increasingly prominent over the past three decades. It is unlikely to require an overhaul in the foreseeable future.

4.7 Masculinity and Femininity Scales

The final approach for improving the measurement of gender in survey research focuses on measuring respondents' self-conceptions in relation to masculinity and femininity. Usually, these studies rely on one or two question masculinity and femininity scales as a way to tap into gender role self-conceptions without allocating survey space to a lengthy battery (Bittner and Goodyear-Grant, 2017a,b; Cassino, 2020; Gidengil and Stolle, 2020, 2021). These gender role self-conceptions provide useful information on how respondents of the same gender identity see themselves differently in relation to established gender roles. For example, these questions can allow researchers to identify “hyper-masculine men,” who exhibit different attitudes and behaviour from other men (Gidengil and Stolle, 2021).

While these studies have demonstrated that self-conceptions of masculinity and femininity have considerable predictive power for many outcomes of interest, the usual approach of combining these masculinity and femininity scales with interviewer coding of perceived gender (Cassino, 2020), self-reported binary “sex” (Bittner and Goodyear-Grant, 2017a,b), or three-option gender identity questions (Gidengil and Stolle, 2020, 2021) does not meet the three standards outlined above.

Dignity Standard (Met) When researchers include separate masculinity and femininity scales, they allow respondents to say that they do not feel masculine or feminine. Gender minority respondents who reject these concepts altogether can respond “Not at all” to both questions. It may help to add an explicit “Don’t know” for gender minority respondents – and those questioning their gender – to provide them an opportunity not to answer, though it is unclear from past work how many respondents would choose this option.

Identifiability Standard (Not Met) Although masculinity and femininity scales move beyond a purely binary view of gender, they cannot identify gender minority respondents on their own. For example, non-binary respondents may perceive themselves as (1) not masculine and not feminine, (2) both masculine and feminine, (3) more masculine than feminine, or (4) more feminine than masculine. As a result, additional questions are necessary to identify gender minority respondents.

Resuability Standard (Met) Masculine and feminine gender roles are so deeply entrenched as social constructs. As a result, masculinity and femininity scales will likely

remain useful for the foreseeable future.

Given the evidence that these questions are useful for understanding political behaviour, C-Dem should include them alongside other questions that are important for meeting the three standards outlined above.

The body of work on masculinity and femininity scales suggests some possible best practices for C-Dem. Studies using masculinity and femininity scales differ in the number of response options offered to respondents, including six (Cassino, 2020), seven (Magliozzi, Saperstein, and Westbrook, 2016; Gidengil and Stolle, 2020, 2021), and one hundred and one (Bittner and Goodyear-Grant, 2017a,b). However, a small number of response categories – such as seven categories – is likely more appropriate to C-Dem’s purposes. Cassino (2020) demonstrates that a six-response question can be feasible in a telephone context. Since C-Dem has run both telephone and online surveys, a seven-category question is likely to be adaptable to both telephone and online survey modes.

Past studies also differ in putting masculinity or femininity on opposing ends of one response scale (Bittner and Goodyear-Grant, 2017a,b; Cassino, 2020) or in two separate response scales (Magliozzi, Saperstein, and Westbrook, 2016; Gidengil and Stolle, 2021). However, Magliozzi, Saperstein, and Westbrook (2016) and Gidengil and Stolle (2021) correctly point out that masculinity and femininity are conceptually separable. That is, it is possible for someone to see themselves as both very masculine and very feminine and not at all masculine and not at all feminine. On a single response scale, these are not distinguishable. For this reason, if researchers have an interest in measuring gender role self-conceptions, two questions are more appropriate than one.

5 Recommendations for Future C-Dem Surveys

In this section, I outline my recommendations for future C-Dem surveys, including the next CES. I pull together insights from different existing proposals into a set of questions that meet the the three criteria outlined above – dignity for gender minority respondents, ability to identify gender minority respondents, and adapting to changes in understandings and language around gender over time.

Table 2 displays the five proposed gender questions, as formatted for online surveys. (While all five items should appear on both online and telephone surveys, telephone surveys require some adaptations to make response options clear to respondents.) The first asks about gender identity. This question uses the two most common responses – a man or a woman. It also includes the option to identify as non-binary, which is by far the most common identity label used by individuals who identify as neither men nor women today. This recommendation follows the longstanding practice in survey research of providing more commonly-used terms as close-ended response categories. In order to allow

Table 2: Proposed Items for Gender Identity, Transgender Identity, Sexual Identity, and Gender Role Self-Conceptions for the Next CES

Variable	Question	Response Options
Gender Identity	Are you...?	A man A woman Non-binary Another gender, please specify: [open-ended text box]
Transgender Identity	[Immediately after the Gender Identity question on the same screen:] Are you transgender?	Yes No Don't know
Two-Spirit Identity	[If respondent is Indigenous:] Are you Two-Spirit?	Yes No Don't know
Masculinity Scale (Self-Appraisal)	How masculine do you feel?	1-7 scale from 1 ("Not at all masculine") to 7 ("Very masculine") with an explicit Don't know
Femininity Scale (Self-Appraisal)	How feminine do you feel?	1-7 scale from 1 ("Not at all feminine") to 7 ("Very feminine") with an explicit Don't know

respondents to use the language they use to describe their gender, which is important for preserving the dignity of gender minority respondents, this question also includes an open-ended gender identity option. This open-ended option uses the language “Another gender” rather than “Other” to avoid literally othering gender minority respondents (Bauer et al., 2017). This option allows researchers to identify gender minority respondents by giving them the option to identify as non-binary or with another gender identity. The “another gender, please specify” option also helps the question meet the standard of being able to adapt to future changes in understandings or language around gender by allowing gender minority respondents to write in their own self-descriptions.

The second question asks about transgender identity. This question uses a simple yes-or-no question (“Are you transgender?”). Other large-scale collaborative political science surveys, such as the 2020 Collaborative Multiracial Post-Election Survey, have adopted this question. This question does not include an explainer of the term “transgender,” as suggested in some past work (Badgett et al., 2014), because the term transgender has become increasingly common in public discussions in recent years. It is unlikely that there are many respondents who have no sense of what the word means, and it can take survey time to write longer questions. (As I discuss later in this report, I recommend running a question-wording experiment in a future C-Dem survey to test this possibility.) This question includes a “Don’t know” option out of a concern that some gender minority respondents may not feel like either response option applies to them.

The third question asks about Two-Spirit identity (“Are you Two-Spirit”). Based on recommendations from Two-Spirit researchers, only Indigenous respondents would receive the Two-Spirit identity question (Pruden and Salway, 2020). This approach recognizes Two-Spirit identity as an Indigenous identity and prevents non-Indigenous respondents from appropriating the label. If the racial identity question is toward the end of the questionnaire, the “Another gender, please specify:” option in the gender identity question would allow Two-Spirit respondents to self-identify early in the survey if that is how they would prefer to describe their gender identity. The “Another gender, please specify:” option would also allow Indigenous respondents who use other terms from their own language or nation if they so choose.

The fourth and fifth questions present respondents with separate masculinity and femininity scales. Two separate scales are better than one combined masculinity-femininity scale on both theoretical and practical grounds (Gidengil and Stolle, 2020). Scales with a handful of response options are more suitable for telephone surveys than 0-100 scales (Cassino, 2020), which would allow the scales to appear on both telephone and online surveys. I recommend testing an explicit “Don’t know” option in online surveys to examine the extent to which it produces missing data for these scales.

All five questions should appear in all C-Dem surveys, including the telephone and online CES surveys, provincial and territorial election studies, and Democracy Checkups. While it may seem that there will not be enough transgender or Two-Spirit respondents to

make it worthwhile to ask these questions on all C-Dem surveys, there are three reasons to include these questions. First, these questions are necessary to ensure the dignity of gender minority respondents. Second, it is likely that these populations will increase over time, given that young people are increasingly likely to take on gender minority identities.¹¹ Third, these questions will allow C-Dem researchers to pool data across surveys to gather enough survey responses to study these populations on their own. These will contribute to C-Dem’s existing goal of gathering data on small populations (Stephenson et al., 2021).

When there are multiple waves within the same survey, as in the CES, the first three questions – gender identity, transgender identity, and Two-Spirit identity – should all appear on the campaign period survey (CPS). Given the small percentage of gender minority respondents in the general population, it is particularly important to avoid losing respondents to attrition between survey waves. I have no specific recommendation about the masculinity and femininity scales, though it may be useful to present the same respondents with these scales in different survey waves to examine their response stability.¹²

6 Next Steps in Improving Gender Questions

While I have proposed a set of questions that meets the three standards, my review of the existing proposals for improving the measurement of gender in survey research suggests that there are some unresolved debates around different sets of questions that may also meet these standards. As a result, I propose three main C-Dem initiatives to improve gender questions in survey research, including a series of question wording experiments, asking gender minority respondents to evaluate gender questions, and bringing together researchers from other major political science survey teams and from other disciplines to discuss alternative proposals.

6.1 Question Wording Experiments

C-Dem should consider running question wording experiments to evaluate new questions on gender and sexual identity. These question wording experiments would help resolve outstanding questions from the literature on gender questions:

1. **How well does an open-ended gender identity question perform in comparison with a close-ended gender identity question?** Open-ended questions allow

¹¹For example, in a survey of high school students in a school district in the northeastern United States, over nine percent of respondents reported a gender identity that did not correspond to their assigned sex at birth. (Kidd et al., 2021).

¹²I thank Michael Donnelly for this suggestion.

respondents more freedom to self-identify their gender identities, and they are particularly well-suited to adapt to changes in gender identity terms over time. Past research on open-ended gender identity questions suggest that the costs of asking these questions are minimal. However, survey researchers in practice report that some respondents provide hostile responses.

2. **Is it necessary to include an “I don’t understand the question” option in transgender identity or gender modality questions?** Researchers started including these question options several years ago, before the rise in “trans visibility.” As a result, it is unclear how many cis respondents actually do not understand the term.
3. **How well does the 3x3 approach to asking about gender perform in comparison with more traditional approaches?** Psychologists have recently developed a new approach to measuring gender that asks about gender identity, gender modality, and gender binarity as three dimensions for understanding gender identity. While it is promising, the format of the questions may strike respondents as unusual, and there is little evidence of its effectiveness in population surveys like the CES. An experiment would allow for a comparison of this approach with more traditional measures of gender identity and transgender identity.

C-Dem should examine these questions with a series of question-wording experiments on the next CES online panel. I propose to address the first question – close-ended or open-ended – by randomly assigning respondents to receive the close-ended and the open-ended responses in different waves of the survey. This would allow both a between-subjects comparison of responses within the CPS and a within-subjects comparison of responses across different gender identity questions. I propose to address the second question – the “transgender explainer” question – by randomly assigning respondents to receive or not to receive a version of the transgender identity question that includes an explainer. Finally, since the 3x3 approach includes a pre-existing series of questions on gender identity, gender modality, and gender binarity, I propose to compare the entire 3x3 approach with more traditional two-step approaches. If we fully cross the gender identity and gender modality experiments, that produces a 2x2 of conditions, along with one additional condition for the 3x3 approach. Table 3 displays the five experimental groups.

These question wording experiments can examine response times for each question, item non-response, and total non-response after the question as outcomes to evaluate the various conditions. For the 3x3 approach, I propose to evaluate response times across the entire set of questions in each condition, rather than each question separately.

For the open-ended questions, I recommend starting by coding the most common responses to open-ended gender questions – M/Male/Man and F/Female/Woman – using an automated dictionary-based approach. These six most common responses account

Table 3: Proposed Question Wording Experiments for Gender Identity and Transgender Identity/Gender Modality Items

Variable	Group 1	Group 2	Group 3	Group 4	Group 5
Gender Identity	CPS: Close-Ended PES: Open-Ended	CPS: Open-Ended PES: Close-Ended	CPS: Close-Ended PES: Open-Ended	CPS: Open-Ended PES: Close-Ended	3x3 Approach
Transgender Identity/Gender Modality	No Transgender Explainer	Transgender Explainer	No Transgender Explainer	Transgender Explainer	3x3 Approach
Gender Binariness	(Included in Gender Identity)	(Included in Gender Identity)	(Included in Gender Identity)	(Included in Gender Identity)	3x3 Approach

for the overwhelming majority of respondents (Fraser et al., 2020; Lindqvist, Sendén, and Renström, 2020). Two researchers can then code each of the remaining open-ended responses manually. I propose to code these responses into the following categories: (1) Men, (2) Women, (3) Non-Binary People, (4) Two-Spirit People (if they give Two-Spirit as a gender identity label), (5) Hostile or Non-Cooperative Responses, and (6) Other Responses. If several respondents choose the same gender identity terms, then it is possible to add additional categories as needed. The between-subjects and within-subjects design will allow for two different approaches to assessing the prevalence of hostile or non-cooperative approaches, as well as providing data on which respondents engage in hostile or non-cooperative respondents.

6.2 Asking Gender Minority Respondents to Evaluate the Questions

The views of gender minority respondents are particularly important for understanding the extent to which gender questions preserve their dignity. One way to elicit their views would be to ask questions that allow researchers to identify gender minority respondents in the campaign-period wave of the CES (as discussed above) and then present a series of open-ended questions that ask gender minority respondents to evaluate the close-ended and open-ended gender identity questions, along with the 3x3 approach. After these open-ended questions, the survey could ask gender minority respondents to pick which

of these approaches they would prefer in a close-ended question.

6.3 A Workshop on Measuring Gender in Survey Research

After C-Dem researchers finish collecting and analyzing the data from these question wording experiments and questions for gender minority respondents, C-Dem should hold a workshop on how to improve gender questions for future C-Dem surveys. This workshop should include presentations from C-Dem researchers who evaluate the gender questions from the CES. However, the invitees should not only be C-Dem researchers. In particular, C-Dem should ensure that the workshop attendees reflect the particular interests of gender minority respondents by inviting trans, non-binary, and Two-Spirit researchers who work on measuring gender across fields. This workshop could lead to a special issue of a journal or an edited volume on measuring sex and gender in survey research.

Based on this research, C-Dem can adopt an updated set of gender questions that will ensure the dignity of gender minority respondents, identify gender minority respondents, and allow for the collection of comparable data across C-Dem surveys.

References

- Alexander, Amy C., Catherine Bolzendahl, and Patrik Öhberg. 2021. "Gender, Socio-Political Cleavages and the Co-Constitution of Gender Identities: A Multidimensional Analysis of Self-Assessed Masculine and Feminine Characteristics." *European Journal of Politics and Gender* 4(1): 151–171.
- Ashley, Florence. 2022. "'Trans' Is My Gender Modality: A Modest Terminological Proposal." In *Trans Bodies, Trans Selves*, ed. L. Erickson Schroth. 2nd ed. Oxford University Press, 22.
- Badgett, M. V. Lee, Kellan E. Baker, Kerith J. Conron, Gary J. Gates, Alison Gill, Emily Greytak, and Jody L. Herman. 2014. *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys*. Los Angeles, CA: The Williams Institute.
- Bauer, Greta R, Jessica Braimoh, Ayden I Scheim, and Christoffer Dharma. 2017. "Transgender-Inclusive Measures of Sex/Gender for Population Surveys: Mixed-Methods Evaluation and Recommendations." *PloS one* 12(5): e0178043.
- Bittner, Amanda, and Elizabeth Goodyear-Grant. 2017a. "Digging Deeper into the Gender Gap: Gender Salience as a Moderating Factor in Political Attitudes." *Canadian Journal of Political Science* 50(2): 559–578.
- Bittner, Amanda, and Elizabeth Goodyear-Grant. 2017b. "Sex isn't Gender: Reforming Concepts and Measurements in the Study of Public Opinion." *Political Behavior* 39(4): 1019–1041.
- Bleischel, Will, Zach C. Schudson, Rhea Ashley Hoskin, and Sari M. van Anders. 2021. "Recommendations for Inclusive Gender/Sex Demographic Questions using the Gender/Sex 3x3." Paper presented at the Society for Personality and Social Psychology, San Francisco, CA.
- Bornstein, Kate. 2016. *Gender Outlaw: On Men, Women, and the Rest of Us*. Revised and updated edition ed. New York, NY: Vintage Books.
- Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. New York, NY: Routledge.
- Canada, Statistics. 2022. *Canada is the First Country to Provide Census Data on Transgender and Non-Binary People*. Ottawa: Statistics Canada.

- Cassino, Dan. 2020. "Moving Beyond Sex: Measuring Gender Identity in Telephone Surveys." *Survey Practice* 13(1).
- Davis, Georgiann. 2015. *Contesting Intersex: The Dubious Diagnosis*. New York, NY: NYU Press.
- Fraser, Gloria. 2018. "Evaluating Inclusive Gender Identity Measures for Use in Quantitative Psychological Research." *Psychology & Sexuality* 9(4): 343–357.
- Fraser, Gloria, Joseph Bulbulia, Lara M Greaves, Marc S Wilson, and Chris G Sibley. 2020. "Coding Responses to an Open-Ended Gender Measure in a New Zealand National Sample." *The Journal of Sex Research* 57(8): 979–986.
- Geary, Rebecca S, Clare Tanton, Bob Erens, Soazig Clifton, Philip Prah, Kaye Wellings, Kirstin R Mitchell, Jessica Datta, Kirsten Gravningen, and Elizabeth Fuller. 2018. "Sexual Identity, Attraction and Behaviour in Britain: The Implications of Using Different Dimensions of Sexual Orientation to Estimate the Size of Sexual Minority Populations and Inform Public Health Interventions." *PLOS ONE* 13(1): e0189607.
- Gidengil, Elisabeth, and Dietlind Stolle. 2020. "Comparing self-categorisation approaches to measuring gender identity." *European Journal of Politics and Gender* 4: 31–50.
- Gidengil, Elisabeth, and Dietlind Stolle. 2021. "Beyond the 'Gender Gap': The Role of Gender Identity." *The Journal of Politics*: Forthcoming.
- Holzberg, Jessica, Renee Ellis, Matt Virgile, Dawn Nelson, Jennifer Edgar, Polly Phipps, and Robin Kaplan. 2017. *Assessing the Feasibility of Asking about Gender Identity in the Current Population Survey: Results from Focus Groups with Members of the Transgender Population*. Washington, DC: Bureau of Labor Statistics.
- Jaroszewski, Samantha, Danielle M. Lottridge, Oliver L. Haimson, and Katie Quehl. 2018. "'Genderfluid' or 'Attack Helicopter': Responsible HCI Research Practice with Non-Binary Gender Variation in Online Communities." Paper presented at the The 2018 CHI Conference, Montreal, QC.
- Karkazis, Katrina. 2008. *Fixing Sex: Intersex, Medical Authority, and Lived Experience*. Durham, NC: Duke University Press.
- Kessler, Suzanne ., and Wendy McKenna. 1985. *Gender: An Ethnomethodological Approach*. Chicago, IL: University of Chicago Press.
- Kidd, Kacie M., Gina M. Sequeira, Claudia Douglas, Taylor Paglisotti, David J. Inwards-Breland, Elizabeth Miller, and Robert W. S. Coulter. 2021. "Prevalence of Gender-Diverse Youth in an Urban School District." *Pediatrics* 147(6): e2020049823.

- Lindqvist, Anna, Marie Gustafsson Sendén, and Emma A. Renström. 2020. "What is Gender, Anyway: A Review of the Options for Operationalising Gender." *Psychology & Sexuality*: Forthcoming.
- Magliozzi, Devon, Aliya Saperstein, and Laurel Westbrook. 2016. "Scaling Up: Representing Gender Diversity in Survey Research." *Socius* 2: 2378023116664352.
- McDermott, Rose, and Peter K. Hatemi. 2011. "Distinguishing Sex and Gender." *PS: Political Science & Politics* 44(1): 89–92.
- McLemore, Kevin A. 2015. "Experiences with Misgendering: Identity Misclassification of Transgender Spectrum Individuals." *Self and Identity* 14(1): 51–74.
- McLemore, Kevin A. 2018. "A Minority Stress Perspective on Transgender Individuals' Experiences with Misgendering." *Stigma and Health* 3(1): 53.
- Medeiros, Mike, Benjamin Forest, and Patrik Öhberg. 2020. "The Case for Non-Binary Gender Questions in Surveys." *PS: Political Science & Politics* 53(1): 128–135.
- Namaste, V. 2000. *Invisible Lives: The Erasure of Transsexual and Transgendered People*. Chicago, IL: University of Chicago Press.
- Pruden, Harlan, and Travis Salway. 2020. "What and Who is Two-Spirit?" in Health Research." CIHR Meet the Methods Series, Volume 2 (October), <https://cihr-irsc.gc.ca/e/52214.html> (accessed 8 April 2021).
- Saperstein, Aliya, and Laurel Westbrook. 2021. "Categorical and Gradational: Alternative Survey Measures of Sex and Gender." *European Journal of Politics and Gender* 4(1): 11–30.
- Sexual Minority Assessment Research Team. 2009. *Best Practices for Asking Questions About Sexual Orientation on Surveys*. Los Angeles, CA: The Williams Institute.
- Stephenson, Laura B, Allison Harell, Daniel Rubenson, and Peter John Loewen. 2021. "Measuring Preferences and Behaviours in the 2019 Canadian Election Study." *Canadian Journal of Political Science* 54(1): 118–124.
- Stryker, Susan. 2017. *Transgender History: The Roots of Today's Revolution*. Seal Studies second ed. New York, NY: Basic Books.
- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1(2): 149–178.

- Turner, John C. 1975. "Social Comparison and Social Identity: Some Prospects for Inter-group Behaviour." *European Journal of Social Psychology* 5(1): 1–34.
- Turner, John C., Michael A. Hogg, Penelope J. Oakes, Stephen D. Reicher, and Margaret S. Wetherell. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. New York, NY: Basil Blackwell.
- Westbrook, Laurel, and Aliya Saperstein. 2015. "New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys." *Gender & Society* 29(4): 534–560.
- Wolff, Margaret, Brooke Wells, Christina Ventura-DiPersia, Audrey Renson, and Christian Grov. 2017. "Measuring Sexual Orientation: A Review and Critique of U.S. Data Collection Efforts and Implications for Health Policy." *The Journal of Sex Research* 54(4-5): 507–531.