

Inclusive Measures, Reductive Analysis: Gender Multiplicity and the Impact of Forced Categorization on Transgender and Gender-Diverse Participants

Theodore Forest Quinn, B.A.^{1,5} & Erin Leigh Courtice, Ph.D.¹, The Garden¹, Mojdeh Jankouk, B.A.¹, Alex Abramovich, Ph.D.^{2,3}, Jeffrey D. Wardell, Ph.D., C.Psych.^{2,3,4}, Trevor A. Hart, Ph.D., C.Psych.^{1,5}, Sarah S. Dermody, Ph.D., C.Psych.¹

¹Department of Psychology, Toronto Metropolitan University, Toronto, Ontario, Canada

²Institute for Mental Health Policy Research, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

³Department of Psychiatry, University of Toronto

⁴Department of Psychology, York University, Toronto, Ontario Canada

⁵Dalla Lana School of Public Health, University of Toronto, Toronto Ontario, Canada

Author Note

Theodore F. Quinn and Erin L. Courtice contributed equally to this work and share first authorship. The Garden is the Community Advisory Board for the Predictors of Alcohol Use and Consequences among Transgender (PACT) Adults Study. Members of The Garden who contributed to the current article and wish to be named include Andriy Hrabchuk, Asher Koshkin, Angelo Luu, and Finley Rogers; one member elected to remain anonymous. For transparency, the authors disclose that as the Director of the HOPE Centre, Dr. Hart has received funding from Gilead and ViiV; and a gift donation from DECIEM.

This research was supported by a grant from the Canadian Institutes of Health Research (CIHR; Funding reference number AI3 180919; PI: Sarah S. Dermody). The views expressed herein do not necessarily represent the official policy of CIHR.

Correspondence concerning this article should be addressed to Sarah S. Dermody,
Department of Psychology, Toronto Metropolitan University, 350 Victoria St., Toronto,
Ontario, Canada M5B 2K3; Email: ssdermody@torontomu.ca

PROOF COPY

Abstract

Researchers have increasingly adopted inclusive practices in measuring gender identity, such as select-all-that-apply (SATA) question formats. However, little is known about what is lost when gender multiplicity (identifying with multiple gender labels simultaneously) is collapsed for quantitative analysis. Even less is known about how transgender and gender-diverse (TGD) participants experience this process of forced reduction. The current mixed-methods study explores how Canadian TGD adults responded to forced gender identity categorization in a study examining alcohol use. Participants ($N = 99$) completed a SATA gender identity item. Those who selected multiple labels ($n = 71$) were then asked to choose a single label for analytic purposes, and describe how this process made them feel via an open-ended prompt. Quantitative responses were visualized with an alluvial diagram to illustrate flow from multi-label to single-label identities. Open-ended responses were analyzed with inductive, reflexive thematic analyses. In total, 32 unique gender identity combinations were reported, with two participants opting not to select a single identity. Nearly all multi-label participants (98.6%) responded to the open-ended question. Two thematic categories emerged: (1) overall emotional responses to being asked to select one label (e.g., positive, indifferent, indecisive, negative), and (2) underlying reasons for those responses (e.g., one label felt insufficient, context mattered, identity-related uncertainty, or strategic simplification). Findings highlight that gender identity multiplicity may be high among TGD individuals, raising critical questions about how reductive analytic procedures may undermine inclusive measurements of gender among TGD individuals. Implications for ethical, participant-centred approaches to gender data are discussed.

Keywords. Transgender, gender diversity, survey design, inclusion, mixed-methods research.

Public Significance Statement

Transgender and gender-diverse (TGD) people can and do use multiple terms to describe their gender identity. Yet, psychological researchers often force these complex identities into single categories for data analyses, which may erase important aspects of people's gender. This study found that over 70% of TGD adults used multiple gender labels. Although most were willing to select a single category when asked, many experienced this forced categorization as distressing or inadequate representations of their gender. These findings highlight the need for researchers to obtain informed consent before recategorizing participants' gender identities and to be transparent about the diversity that is lost in this process.

Inclusive Measures, Reductive Methods: Gender Multiplicity and the Impact of Forced Categorization on Transgender and Gender-Diverse Participants

Psychology researchers have increasingly adopted inclusive measurement approaches that better reflect the complexity of transgender and gender-diverse (TGD) identities. These approaches most commonly include two-step gender/sex measures, open-ended or write-in responses, and select-all-that-apply (SATA) checklists, each designed to capture aspects of multidimensional gender constructs in slightly different ways (Fraser, 2018; Herman et al., 2024; Puckett et al., 2020).¹ Nonetheless, researchers often recategorize responses into a single higher-order category to facilitate analysis (Bauer et al., 2017; Lombardi & Banik, 2016; Magliozzi et al., 2016; Urlacher, 2022). To clarify how gender recategorization and forced choice impact TGD participants and the validity of gender identity data, we examined the prevalence of multi-label gender identification among TGD adults. We also examined how participants who endorsed multiple gender identity labels selected a single label and qualitatively explored their reactions to this process.

Common Approaches to Gender Measurement

One of the most widely adopted gender measurement approaches is the two-step method, which asks participants to report their sex assigned at birth (SAAB) and current gender identity in two separate questions (e.g., Bauer et al., 2017; Cahill et al., 2014; Deutsch, 2016; Patterson et al., 2017; Reisner et al., 2014; Tate et al., 2013; The GenIUSS Group, 2013; Tordoff et al., 2019; 2025). Although question phrasing varies across studies, the most common implementation uses an SAAB item with male, female, and sometimes intersex as response options, followed by a categorical list of gender identities (e.g., man,

¹ Appropriate gender identity measurement varies across research contexts and disciplines. In medical settings where biological characteristics are clinically relevant (e.g., hormone therapy), questions about sex assigned at birth or physiology may be necessary. Our discussion focuses on psychological research contexts where gender identity is examined as a social identity or demographic characteristic rather than as a proxy for sex. We recognize that researchers in other fields may prioritize different measurement strategies based on their specific research aims.

woman, nonbinary) (Bauer et al., 2017; Fraser, 2018). For instance, as Herman et al. (2024) point out, a 2022 report from the National Academies of Science, Engineering, and Medicine (NASEM; Committee on Measuring Sex, Gender Identity, and Sexual Orientation et al., 2022) recommended the following approach: (p. 22): (a) “What sex were you assigned at birth, on your original birth certificate?” and (b) “What is your current gender?”. Researchers are advised to compare participant responses to both questions in order to determine whether individuals are cisgender or TGD.

The core strength of the two-step method lies in its ability to identify TGD participants through a comparison of SAAB and current gender identity, even when individuals do not explicitly identify as “transgender” (Bauer et al., 2017; Conron et al., 2008; Kidd et al., 2022; Suen et al., 2020). Empirical evidence supports the two-step method’s improved sensitivity; for instance, Tate et al. (2013) found that it identified transgender participants at nearly twice the rate of a single-question format. Additionally, research suggests that both cisgender and TGD participants generally understand and are comfortable responding to two-step questions (Lombardi & Banik, 2016; Reisner et al., 2014). Consequently, multiple major institutions have recommended the two-step method, including Statistics Canada, the Gender Identity in US Surveillance Group through the Williams Institute, the Center for Excellence in Transgender Health at the University of California, San Francisco, and the World Professional Association of Transgender Health Electronic Medical Record Working Group (Deutsch et al., 2013; The GenIUSS Group, 2013; Harrison-Quintana et al., 2012; Kidd et al., 2022; Lagos & Compton, 2021; Park, 2016; Statistics Canada, 2022).

Despite these strengths, the two-step method also has notable limitations. First, some TGD participants report that being asked about their SAAB can be distressing or invalidating in medical and non-medical contexts, especially when the information seems irrelevant

(Abramovich & Michel, 2025; Conron et al., 2008; Fraser, 2018; Puckett et al., 2020; Suen et al., 2020). Second, the two-step method does not fully capture the nuance and multiplicity within TGD identities (Abramovich & Michel, 2025; Pao et al., 2025), such as by overlooking individuals who identify with more than one gender and/or those who use culturally specific or emergent labels that are less easily reduced (Bauer et al., 2017; Fraser, 2018; Hillier et al., 2010; Mills et al., 2019). Similarly, a two-step method might miss the nuances of code-switching, a phenomenon in which people adapt their descriptions of gender identity based on contextual factors (e.g., in public versus with trusted individuals; Glen & Hurrell, 2012; Suen et al., 2020; Vivienne et al., 2023). Abramovich and Michel (2025) also note that the two-step method may further misclassify non-cisgender people who don't identify as "trans" (i.e., those who identify as a woman or a man, rather than as a specifically "transgender" woman or "transgender" man). Third, the two-step approach may alienate intersex individuals, as intersex identity at birth is often unknown, ignored, and/or omitted, and can vary between countries as a sex designation at birth (Bauer et al., 2017).

Given the limitations of the two-step method, researchers have recently explored alternative approaches to assess gender identity. For instance, some researchers have used write-in questions without any pre-established categorized list; this approach allows participants to describe their gender in their own words, offering maximal freedom and flexibility (Glick et al., 2018; Kuper et al., 2012; Pinto et al., 2019; Rankin et al., 2010). Similarly, SATA formats present a comprehensive list of gender identity labels, allowing participants to select as many as they feel apply (Broussard et al., 2018; Callander et al., 2021; Rankin et al., 2010). Both write-in and SATA question formats enable participants to express gender multiplicity (identifying with multiple terms simultaneously) and may better accommodate emerging or community-specific language relative to fixed-category options (Kuper et al., 2012; Rankin et al., 2010). Other researchers have combined write-in and SATA

approaches, generating variations of new multi-step methods. For instance, Ker et al. (2025) expanded on the two-step approach by asking participants to (a) write in a word or phrase to describe their gender, (b) select a single TGD identity from a researcher-created list of options, and (c) report their SAAB. Herman et al. (2024) used what they term as Gender EXPAND (EXPANsive responsive genDer), asking participants to (a) report whether or not they identify as TGD, (b) select their gender from a SATA list, which included a write-in response, and (c) select one option from a shorter list of researcher-generated gender categories.

Preliminary evidence suggests that, overall, TGD participants view these alternative approaches positively. For example, exploratory qualitative research has found that TGD participants explicitly recommended both write-in and SATA formats as ways to improve inclusivity (Puckett et al., 2020; Marlar & Willcoxon, 2023). Across studies, participants have emphasized the importance of having language that reflects their identities and of being able to choose among multiple terms. However, using SATA and write-in formats also introduces new analytic challenges, which researchers are only beginning to address. In particular, it remains unclear how best to categorize and analyze responses when participants endorse multiple identity terms or provide open-ended responses that resist standard classification. For instance, Herman et al. (2024) found that researchers miscategorized 9.5% of SATA responses when a single gender identity was reported and 13.6% when two or more were reported.

Using Inclusive Gender Identity Data in Statistical Analyses

Despite growing consensus on inclusive gender identity questions, substantial challenges remain in using these data for statistical analyses. Most quantitative approaches to gender-based analyses require reducing diverse data to simplified, mutually exclusive categories (Guyan, 2022; Suárez, 2024). However, this simplification can directly undermine

the inclusivity of the original measurement approach. For example, the two-step method tends to yield relatively straightforward categories for analysis (most commonly "man," "woman," and a catch-all "TGD" or "other" group that includes transgender and nonbinary participants). This strategy is analytically convenient, but it is also reductive in that it treats gender as static and singular, and obscures meaningful variation within the TGD umbrella (Rankin & Garvey, 2015). As such, (re)classifying transgender and nonbinary people together as one group may be problematic when researchers seek to examine mechanisms like minority stress and social determinants of health, which appear to operate differently across diverse gender identities (Dermody et al., 2025; Goldberg et al., 2019; Jacobsen et al., 2024; Johnson et al., 2020; Poquiz et al., 2021; Pradell et al., 2024).

Write-in and SATA questions may yield richer, more flexible data. However, they also require researchers to interpret and recode responses, which often involves collapsing responses into a small number of categories for statistical power or ease of comparison. Through recategorization, researchers may end up with similarly reductive groupings despite richer data. Importantly, little research has examined what is lost when participants who report multiple gender identities are forced into a singular analytic category. Consider two participants who both identify as nonbinary: one also identifies as transfeminine and the other as transmasculine. Though both may be categorized under a shared "nonbinary" (Goldberg et al., 2019) or "TGD" label for analyses, these participants likely navigate different social expectations or forms of stigma through processes such as femmephobia (the devaluation, denigration, and regulation of femininity in any and all genders and sexes, Hoskin, 2020). Similarly, these two participants may have access to different sources of gender-minority resilience (Poquiz et al., 2021). Although many best-practice guidelines recommend that participants be consulted or allowed to self-select a single analytic identity (e.g., Cameron &

Stinson, 2019; Veldhuis et al., 2024), the frequency with which researchers follow this recommendation, or how participants experience this process, remains unclear.

More fundamentally, emergent literature suggests that a large portion of TGD individuals may commonly identify with multiple gender labels at a given time. Galupo et al. (2017) found that many participants' gender identities could not be reduced to a single category, with descriptions that blended, challenged, or moved fluidly across established terms. Their findings echo earlier work in sexual identity research (Galupo et al., 2015), where individuals with plurisexual identities (e.g., bisexual, pansexual, etc.) and identities that may not have been as commonly recognized at the time of study (e.g., asexual, polyamorous, etc.) often used multiple labels or more elaborate self-descriptors. Other researchers have similarly noted that, when investigating gender diversity among TGD participants, over 70% use multiple labels to describe their gender identity (Callander et al., 2021; Kuper et al., 2012). Likewise, Puckett et al. (2020) found that TGD participants recommended multiple response formats (including SATA and write-in) as more reflective of their lived experience, even while acknowledging the complexity they introduce for researchers. Early survey work by Gallup (Marlar & Willcoxon, 2023) suggests that multi-label identification occurs in a small but non-negligible proportion of the general population. In their experimental format, 0.8% of respondents selected two gender labels and 0.1% selected three. Notably, these were self-administered surveys in the general population, and not targeted toward TGD individuals (for whom multiplicity rates may be substantially higher). These findings suggest that although inclusive measurement tools are necessary to capture the complexity of gender, current analytic practices largely erase these nuances.

The Current Study

Despite advances in how gender identity is measured (e.g., Herman et al., 2024; Ker et al., 2025; Puckett et al., 2020), less attention has been given to how these data are used

once collected, particularly in cases where participants identify with more than one gender label. Recent studies have documented that TGD people use multiple labels to describe their gender (Galupo et al., 2017; Ker et al., 2025) and that researcher-driven categorization can lead to miscategorization of TGD identities (Herman et al., 2024). However, these studies focused primarily on technical categorization processes and did not capture participants' perspectives on forced-choice reduction. As researchers increasingly adopt inclusive formats such as SATA checklists, important questions remain about how such data may be recoded for analysis, how TGD participants experience this process, and what information is lost when multi-label identities are collapsed. The current study addresses this gap by integrating quantitative examination of multi-label endorsement patterns with qualitative data examining the impact of forced categorization on participants. Our research questions were:

1. What is the prevalence of multi-label gender identification among TGD adults responding to a SATA gender identity question?
2. Among TGD participants who endorse multiple gender identity labels, which single gender labels do they subsequently select when asked to choose one for analytic purposes?
3. How do TGD participants describe their experiences of being asked to reduce their multi-label gender identities to a single category for research purposes?

Method

Participants and Procedure

Participants were drawn from a larger 21-day Ecological Momentary Assessment (EMA) study that examined predictors and consequences of alcohol use among 99 transgender and gender-diverse (TGD) Canadian adults. Data collection for this study took place from November 2023 to October 2024. Eligibility criteria for the study included: (1) self-identifying as TGD; (2) residing in Canada with proof of residency (e.g.,

government-issued ID, piece of mail, or bank statement); (3) being of legal drinking age (18 or 19+, depending on province/territory); and, (4) endorsing frequent alcohol use (i.e., once per week or more).

All participants completed an individual two-hour baseline appointment, conducted with a trained research assistant and held on a secure online video conferencing platform. After receiving the participant's informed consent, the research assistant conducted a timeline follow-back procedure (Sobell & Sobell, 1992) to review the participant's substance use patterns over the past 30 days. Next, the research assistant reviewed the EMA study procedure (which was not relevant to the current study) with the participant. Participants then completed the baseline survey (the data source for the present study) via an online Qualtrics form. The research assistant remained present (via the video conferencing platform) while the participant completed the baseline survey to address any queries the participant had. The baseline survey included a demographic questionnaire (including questions on gender identity and SAAB), health-related information (e.g., use of gender-affirming hormones), and various questionnaires assessing gender minority stress, resilience, substance use, and mental health. After completing the baseline survey, participants completed a clinical interview assessing for possible alcohol use disorder with a trained clinical psychology graduate student (not relevant to the current study). Participants were compensated with a \$60 (CAD) gift card after the baseline appointment. Debriefing information was sent to participants upon completing the 21-day EMA period. This study protocol was approved by the Toronto Metropolitan University Research Ethics Board.

Measures

Gender Identity

Select All That Apply (SATA) Question. We developed gender identity items in consultation with the study's community advisory board (CAB), "The Garden," which

comprises five TGD adults from across Canada, aged 18–38 at the time the study began. In the baseline survey, participants were first asked: “*How do you describe your gender? Please select all that apply.*” They then selected all relevant options from 15 response options (listed in Table 1), including “gender identity not listed (please specify),” which allowed for a text-entry response.

Single-Response Question. Participants who endorsed multiple gender identity labels on the SATA question were presented with a second question, asking them to select just one gender identity label. Participants who received this follow-up question were first shown the following explanation:

In research literature, trans folks have historically been categorized based on their sex assigned at birth, or a single gender identity that they list in surveys like these, despite the fact that experiences related to gender can often be quite nuanced and unique. Researchers are then faced with an issue: how do we accurately represent these nuanced experiences, while also making sure that the data collected can be analyzed? For data analysis purposes, please select one label that you feel most accurately describes your gender identity/you would prefer to be categorized with for this study.

After reading the explanation above, participants were asked to select only one label that best aligned with their gender identity. The list of gender identity labels that participants could select from was similar to the SATA question (see Table 1), with the exception of two additional response options based on The Garden’s feedback: “I cannot pick one identity” and “I will not pick one identity.”

Open-Ended Item

To evaluate the potential impact of the gender identity questions, participants who endorsed multiple gender identity labels on the SATA question were invited to respond to an

open-ended prompt: *“How did being asked to choose only one label to represent your gender identity make you feel? Please be as specific as you would like.”*

Analytic Approach

We used descriptive analyses to characterize the sample's demographic makeup and response patterns to the nominal gender identity questions. To visually represent and compare the response patterns between the two nominal gender identity variables, we created an alluvial diagram using the `'ggalluvial'` package in R (Brunson, 2020; Brunson & Read, 2023).

We used an inductive, reflexive thematic analytic approach (Braun & Clarke, 2006; 2021; 2023; Levitt, 2018) to analyze the free-text responses expressing feelings about being asked to select a single gender identity label. First, the co-first authors Theodore Forest Quinn (TFQ) and Erin Leigh Courtice (ELC) engaged in independent data familiarization by reading through all of the participant responses. Focusing on a subset of the first 20 entries, the researchers then collaboratively engaged in systematic data coding, generating an initial set of codes to reflect prominent themes in the data. Here, two major themes with subsequent codes emerged: 1) feelings overall (positive, negative, indifferent, and indecision) and 2) reasons for feelings (one identity is insufficient, one identity is sufficient, and context-dependent). To ensure coding reliability, TFQ and ELC both independently coded the first set of 20 responses. Next, they met to discuss coding decisions, resolve discrepancies, and refine the coding framework by generating additional codes to capture additional subthemes. This iterative, collaborative coding process continued with subsequent independent coding of batches of responses (another set of 20, then a set of 10), then meeting after each batch to resolve any discrepancies, integrate newly emerging subthemes, and merge those with similar connotations (e.g., “fine” “alright” “okay”). Thematic saturation was achieved after 50 responses were coded. The remaining 20 responses were similarly independently coded, with a final meeting to resolve any remaining discrepancies.

Throughout the coding process, TFQ and ELC used memos to note unique interpretations of the data and how their own positionality may have influenced these interpretations (Levitt, 2018). This process resulted in a finalized coding scheme reflecting all identified themes. Once coding was complete, participants were randomly assigned colours as pseudonyms to aid in differentiating responses when presenting results (MacEntee et al., 2024). Our use of pronouns throughout this manuscript reflects those used by participants during data collection. Data, materials, and R script will be provided to individual researchers upon request.

Research Team and Reflexivity

In alignment with qualitative reporting standards (Levitt et al., 2018), we offer readers the following to help contextualize how this research and its analyses were approached. Specifically, we reflect on the identities represented across our authorship and Community Advisory Board (CAB) and how that diversity shaped methodological decisions throughout the study. This study (and the larger study from which it derives) adopts a health promotion lens, following community-based participatory research (CBPR) practices and principles (Collins et al., 2018). Namely, TGD community members with varying degrees of expertise and lived/living experience have had ongoing, intentional involvement throughout the study, from idea conception through to data dissemination.

Diversity of the research team is paramount to the success of CBPR; it allows for an exchange of various expertise, and the knowledge gained can then be actively used to effect positive, lasting change that directly benefits the community of focus (Collins et al., 2018; Springer & Skolarus, 2019). Community knowledge is valuable to academic knowledge, and the partnerships developed throughout CBPR should foster empowerment among its members (Collins et al., 2018; Springer & Skolarus, 2019). Decisions made without community input risk misclassification and the perpetuation of epistemic harms. Therefore, in

the current study, the lived experience represented across the authorship and CAB directly informed decisions about measurement, recruitment language, and interpretation of preliminary findings.

One of the first authors, TFQ, is a white, transmasculine, genderqueer, queer-identified, and neurodivergent person; the other first author, ELC, is a queer-identified, white cisgender woman. The third listed author (The Garden) represents the CAB for the current study, which collectively provided feedback on research methodology, gender-identity question design, interpretations of the qualitative data, and manuscript editing. Members of The Garden all hold various TGD and ethno-racial identities, with varying degrees of lived/living experience as community members, researchers/academics/post-secondary students, and/or affiliations with other 2SLGBTQ+² organizations and community work. Alex Abramovich (AA) and Sarah S. Dermody (SSD) are the co-Primary Investigators for the current research and provided feedback on analyses and interpretation of the qualitative data. AA is a multi-ethnic, queer-identified, transgender man, and SSD is a multi-racial, cisgender, heterosexual woman. The remainder of the research team (including co-authors, student trainees, and volunteer/research assistants) hold various positionalities, all of which include TGD allyship and experience with community-engaged research involving TGD people.

Furthermore, in alignment with the principles of CBPR, open communication and shared decision-making (Collins et al., 2018) were central in developing the final gender identity question used in the larger study. The CAB was presented with three options and invited to share any alternative suggestions. Input gathered from CAB members was prioritized and ultimately informed the design of the final question; members voted anonymously on the options presented, and written feedback was incorporated where possible (as described below). Questions were then revised and finalized with the CAB prior to data

² An acronym commonly used in Canada which includes two-spirit identities. Placing “2S” at the beginning of the acronym acknowledges that Indigenous knowledge and ways of being predate colonial ideas of gender identity and sexual orientation.

collection. As CBPR is a cyclical, ongoing collaboration (Collins et al., 2018), CAB members were provided with updates about the performance of the gender identity question throughout the data collection process. In preparation for this manuscript, preliminary findings were shared with the CAB for their commentary and input. Finally, those interested in authorship contributed to the final edit and review of the manuscript's rough draft.

Results

Descriptive Statistics

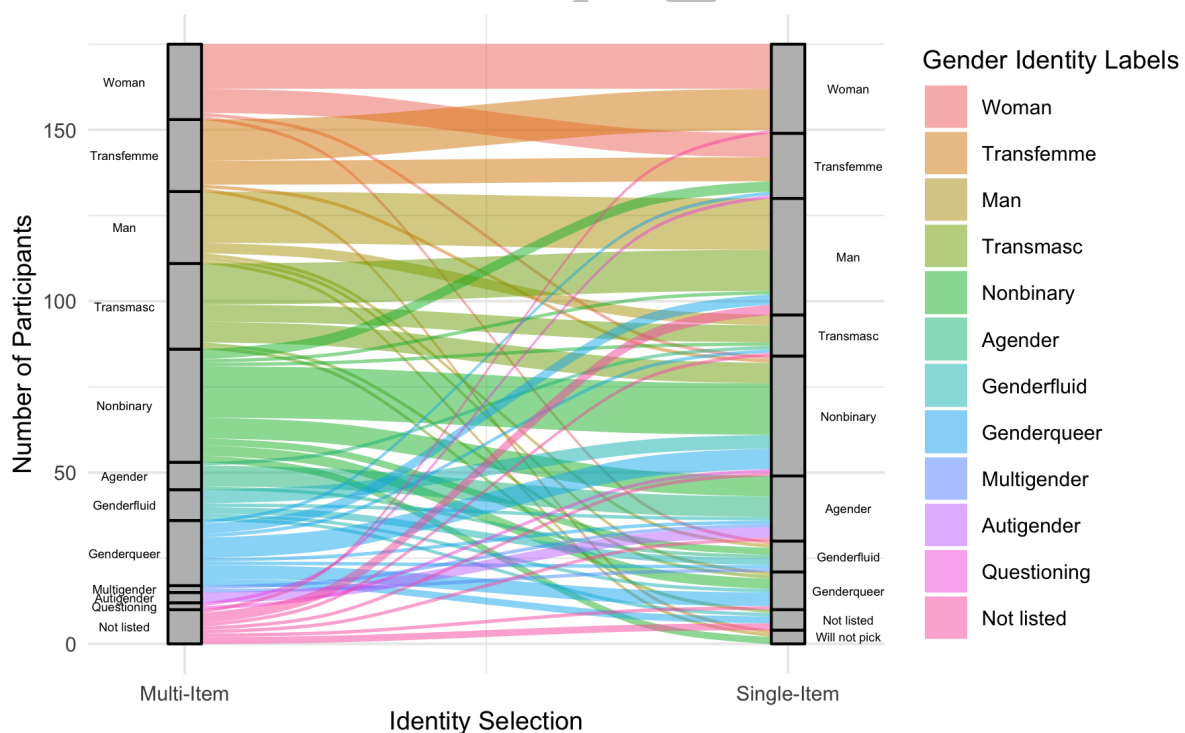
Of the 99 participants included, the mean number of gender identities selected was 2.05 ($SD = 0.99$), and 71 participants used multiple labels to describe their gender identity. The mean age of this subsample was 27.91 ($SD = 7.63$); we provide detailed demographic information in Table 1. The single gender identity labels selected by these participants are presented in Table 2. Among participants who used multiple gender identity labels, 32 unique combinations of gender identity labels were reported. Of these, the most common combinations were transfemme and woman ($n = 15$, 21.1%), transmasculine and man ($n = 11$, 15.5%), and nonbinary and transmasculine ($n = 4$, 0.06%). Among the remaining 41 participants, the remaining 29 unique gender identity label combinations were reported by three or fewer participants, with the majority (20 unique combinations) each reported by only one participant.

Figure 1 presents an alluvial diagram (created with `'ggalluvial'` and `'ggplot2'` in R; Brunson & Read, 2023; Wickham, 2016) that illustrates the flow from multi-select endorsements to forced single identities, with the width of each flow proportional to the number of participants. We selected an alluvial diagram because it effectively illustrates the flow and redistribution of participants from multi-select gender identity endorsements to forced single-choice selections. Unlike traditional bar charts or tables, an alluvial diagram preserves the relational connections between participants' multiple identity labels and their

eventual forced-choice selection. Therefore, these flows underscore the vast diversity with which multidimensional gender identities consolidate into singular categories under forced-choice conditions. A few notable patterns emerge in this diagram. First, most participants who endorsed both transfemme and woman on the SATA question selected woman ($n = 12$) on the single-select question. Similarly, the majority of participants who endorsed both transmasc and man on the SATA question selected man ($n = 10$) on the single-select question. Second, all participants who endorsed both nonbinary and transmasc on the SATA question selected nonbinary on the single-select question ($n = 4$); however, of those who endorsed both nonbinary and transfemme ($n = 2$), one selected nonbinary and the other selected transfemme.

Figure 1.

Alluvial diagram of multi-select to single-select gender identity flows.



Note. This figure depicts the mapping from participants' selections using a multi-select SATA (Select All That Apply) gender identity question format to their selections in a single-select forced-choice format. Each colour represents a specific gender identity label chosen in the

SATA format, shown on the left-hand side of the diagram. The width of each ribbon reflects the number of participants who selected that identity combination in the SATA format ($N = 71$). Ribbons then flow to the single-select forced-choice label each participant chose, shown on the right-hand side. The legend applies to both the multi-item SATA (left) and single-item forced-choice (right) components of the diagram.

Qualitative Results

Of the 71 participants who endorsed multiple gender identity labels on the SATA question, 70 responded to the open-ended item. From these responses, we identified two overarching thematic categories capturing participants' reactions to selecting a single gender identity label: (1) Overall Feelings about Selecting a Single Identity Label and (2) Reasons Underlying Feelings about Selecting a Single Identity Label. Within both themes, distinct categories emerged, reflecting a wide range of experiences and sentiments regarding the single-label approach. More details about the categories of responses within these two themes are presented below; Table 3 provides exact counts of participants who were coded into each category. To maintain authenticity, we present our extracted participant quotations as written (i.e., without editing for punctuation or grammar).

Overall Feelings about Selecting a Single Identity Label

Participants' feelings about selecting a single label to reflect their gender identity ranged from positive acceptance to negative distress, with some expressing indifference or ambivalence, and others expressing indecisiveness about selecting one identity.

Positive. A few participants described feeling positively or affirmed by the opportunity to choose a single identity. For these individuals, selecting a label appeared straightforward and aligned with their sense of self. For instance, Magenta wrote: “[...] *I feel most strongly as just male and so it felt affirming and positive to choose just one label*” (33

years old, Man, Genderqueer, Gender not listed, please specify: [FTM³, Trans Man]). Another participant said, *"i think it was helpful in making me feel out how i really feel i identify"* (Cyan, 43 years old, Man, Genderfluid, Non-binary, Genderqueer).

Indifferent. Many participants reported feeling neutral or unbothered by the request to select a single gender identity label. For instance, Purple shared, *"It didn't particularly bother me"* (25 years old, Non-binary, Genderfluid). Similarly, Teal (27 years old, Non-binary, Transmasc) said: *"It's fine, just picked the one most relevant to where I am now."* Within these responses, many participants used language indicating that selecting a single gender identity label was "fine", "alright", "okay," or that they "didn't mind." We initially differentiated these specific terms as potential indicators of subtle discomfort, but later determined that we could not impose this broad assumption. Therefore, we maintained these initial sub-codes only as indicative of broadly neutral or indifferent responses.

Indecisive. Some participants expressed uncertainty or indecisiveness about choosing one label. For example, as Maroon (23 years old, Non-binary, Agender, Transmasc, Autigender, Genderqueer) expressed: *"It was a struggle to decide because I personally use so many different words to describe myself [...]"* Within this theme, sub-themes emerged related to confusion about how to appropriately represent their identity and to feelings of hesitation about the validity or accuracy of their selected identity. Speaking to the sub-theme of confusion, Orange (28 years old, Man, Transmasc, Gender Non-Conforming) wrote:

Kind of confused--I wasn't really sure whether to put male or transmasc as I identify with both. [...] I find a lot of this language is used imprecisely. Like when putting transmasc as a separate category than man it makes me wonder--what are they basing the distinction on? Genitals? Assigned sex at birth? Presumed experiences of discrimination?

³ This is language that the participant used to (partially) describe their gender identity on the SATA question. Most commonly, FTM refers to someone who has transitioned from "female-to-male".

Hesitation was expressed in about half of all responses within the indecisiveness category. For example, one participant described: *“A little conflicted being unable to select both transfemme and woman. Primarily I associate this to my discomfort in occupying women's spaces and general inexperience living as my desired gender (less than 1 year being fully "out")”* (Burgundy, 33 years old, Woman, Transfemme). Importantly, this sub-theme was reflected across gender identities and ages; as another participant shared:

I feel a bit confused and indecisive. I would say I am a man in ordinary situations because that is how I would like to be perceived. I find that if I say otherwise people will see me as my assigned gender at birth and misgender me, either consciously or subconsciously [...]. (Ivory, 21 years old, Man, Transmasc)

Negative. Other participants explicitly described negative emotional reactions or distress when asked to pick a single identity label. Although many expressed generally negative sentiments, some described more detailed emotional responses such as discomfort, frustration, or anxiety. For instance, as Bronze (28 years old, Man, Non-binary, Transmasc) shared: *“I immediately felt somewhat panicked and overwhelmed -- I felt my heart start beating a bit faster, and my chest felt tight [...].”*. Another participant, Green (24 years old, Woman, Transfemme) said:

not super great honestly... i picked woman because that's what i am and thats what always wanted to be and be percieved as. But ONLY choosing woman feels like erasing my unique path to my gender. i am proud to be trans and with the right people i want to be seen as a trans woman [...].

Reasons Underlying Feelings about Selecting a Single Identity Label

In their responses, many participants offered contextual reasoning for their feelings. The underlying reasons included satisfaction or dissatisfaction with available labels, the influence of contextual factors, experiences of gender identity (e.g., pride and self-assurance,

‘imposter syndrome’), a desire to simplify disclosure for others, and personal or perceived community sentiments toward gender identity labels.

One Identity Will Suffice. Several participants felt that selecting a single label was not problematic, as they already primarily identified with one label in other contexts. For instance, one participant wrote: “[...] *I've identified as non-binary for 10 years and feel it is most fitting*” (Yellow, 31 years old, Non-binary, Transmasc). Subthemes included expressions of pride or self-assurance in one’s identity and appreciation that the available options felt representative and inclusive of their experience. For example, Indigo (28 years old, Woman, Transfemme, Gender Questioning) explained: *“I did not mind picking one specific label to represent my gender identity. I felt that the options provided gave a solid selection that I felt matched my identity.”* For these individuals, a single label was sufficient to capture their gender identity.

One Identity is Insufficient. Other participants found the task challenging or invalidating because no single label could fully capture their gender identity. These responses often reflected a deep sense of complexity or uniqueness in their gender experience. Subthemes included concerns about identity erasure, beliefs that their gender was too complex or personal to fit one label, and difficulty, frustration, or refusal to choose a label that did not accurately represent their identity. As Mauve (25 years old, Non-binary, Transfemme) stated:

Using one label to represent my identity would be nothing short of perjury. My identity, gender, sex, or not - is not quantifiable in a single label. My experiences are myriad and my representation is too minimal to pin down a singular label that aptly identifies me.

Context Dependent. Some participants' comfort with selecting a single label depended on the question's context. Participants expressed that they were more willing to

choose one label in the context of research or data analysis, even if they would not do so in other personal or social settings: *"I don't really like it that much, but i do truly understand why you chose to [make] it that way so i'm fine with it in that regards if it makes sense"* (Viridian, 24 years old, Non-binary, Agender). Notably, some participants referenced the disclosure prompt before being asked to choose one identity directly. In these responses, participants acknowledged the limitations of current research and data analysis practices and/or expressed comfort with choosing one identity due to researcher transparency. As noted by Periwinkle (19 years old, Man, Transmac):

Like you said, most surveys typically categorize trans participants based on their assigned sex at birth - not reflecting the unique and nuanced lived experiences of trans participants. This made me feel like my identity was more respected and that researchers recognize the distinction between gender identity and sex.

This theme underscores the importance of perceived purpose and participant involvement in shaping identity disclosure, highlighting the need for a primer such as the one included in this study (see Measures).

Imposter Syndrome / Not 'Trans Enough'. A few participants described hesitation or discomfort with choosing a label due to feelings of inadequacy or perceived inauthenticity. These individuals expressed uncertainty about 'qualifying' for specific terms or worried that selecting a particular label would imply an identity they did not fully feel permitted to claim. For instance, Mahogany (44 years old, Woman, Transfemme) shared:

It made me doubt myself whether I was "woman enough yet" to answer solely "woman" to the question. I went back and forth twice on my answer and finally opted for "transfemme" instead, on the grounds that I am not stealth, I am openly trans in real life contexts. [...].

Easier for Others. Some participants indicated that they selected one identity label because it was easier for others to understand, even if not entirely accurate. For example, one participant stated: “[...] *I normally only select woman when filling things out because I see myself as a 'woman' more than a 'transfemme' or a 'transwoman'. I think it makes it easier for everyone if that's what I choose*” (Violet, 21 years old, Woman, Transfemme). These individuals emphasized simplification for clarity, communication, or social ease, suggesting a pragmatic, rather than identity-driven, rationale or understanding of data analysis needs.

Feelings Towards Labels. Participants expressed a range of personal feelings and reflections about gender identity labels for themselves. Some shared individual opinions on whether labels capture their identity experience. For instance, Lilac (32 years old, Non-binary, Transmasc, Genderqueer) stated: “[...] *I feel like non-binary encompasses the fact that I abstain from being defined by gender as it pertains to the socially constructed gender binary. That is enough information for me.*” Another participant noted that they were aware that others might find the request distressing or exclusionary:

I felt alright about choosing one specific label, as I align most with the identity of 'man' [...] I understand that this is not the case for many within the transgender community, and I acknowledge that I am privileged in this regard. (Crimson, 24 years old, Man, Transmasc)

This theme captures personal and community-oriented reflections on the limits and politics of labelling gender.

Discussion

The current study sought to improve data collection and recategorization processes for future research by developing a set of community-informed gender survey items and methods for obtaining participants' consent to recategorize gender identity data. In doing so, we gained

novel insights into the concordance between multi-select and forced-select gender categorization, including participants' experiences when responding to such items.

By obtaining participants' consent to recategorize them into a single gender category, the majority were at least willing to select a single gender category for data analysis. Despite the willingness of participants to make single-gender selections, our findings suggest that this approach oversimplifies gender identity and leads to a loss of gender diversity in data.

Specifically, over 70% of our participants used multiple labels to describe their gender identity, which corresponds to past research exploring the diversity of gender identities in TGD groups (Callander et al., 2021; Kuper et al., 2012). Although there was notable consistency for certain gender subgroups, particularly among individuals who initially selected a binary gender, there were also notable variations. These data highlight the importance of asking participants to select how they wish to be categorized, rather than allowing researchers to make such category decisions by assuming the most salient or relevant identities for an individual. Further, based on these findings, researchers who categorize participants after the fact should be explicit about gender diversity before and after these recategorizations and note the limitations of doing so (e.g., erroneous recategorization, oversimplifying gender).

The qualitative data revealed critical differences in how this research practice impacted participants. Most participants noted feelings of indifference. Similarly, some participants expressed that a single gender identity would suffice on a contextual basis for analytic purposes, recognizing the practical reasons behind the request. These responses may provide some reassurance for researchers who engage in gender recategorization processes. However, many participants (almost as many as those who were otherwise unimpacted) expressed discomfort or distress. Similar to previous research (Galupo et al., 2017), some participants noted that selecting a single gender misrepresented or inadequately represented

their gender. These responses emphasize the importance of consideration in how questions about gender are prefaced and highlight that gender assessment processes can negatively impact some TGD participants.

Our findings point towards a community-informed process for capturing and categorizing diverse gender identities for research purposes, while also highlighting important limitations of a categorical approach to representing gender identity, such as a two-step approach. Indeed, although the two-step approach allows researchers to differentiate SAAB from gender, it still requires participants to select only one option to represent their SAAB and gender and can overlook nuances surrounding intersex identity (Bauer et al., 2017; Fraser, 2018; Puckett et al., 2020). In contrast, our findings suggest that reliance on singular categories is inadequate in capturing gender identity - at least amongst gender diverse adults. Beyond the finding that most participants selected multiple categories to find the best fit for their identity, qualitative responses highlighted the dynamic nature of gender. For instance, we found that gender endorsement was sometimes context dependent (e.g., for research purposes, whether the audience is cisgender or transgender) rather than an exact reflection of their gender identity. Indeed, recent research has challenged the notion that the categorical, closed-ended components of the two-step approach accurately capture gender identity for TGD participants (Pao et al., 2025; Suen et al., 2020). In fact, some research suggests that multi-step (rather than two-step questions) may be more accurate in assessing gender identity, as they account for its dynamic and multifaceted nature (Bauer et al., 2017; Suen et al., 2020). Further, there are crucial analytic challenges with categorical predictor variables, such as achieving minimum cell sizes and sufficient statistical power, which often underlie the recategorizations in the first place (e.g., Rich-Edwards et al., 2018). Taken together, persisting with categorical gender identities when working with gender-diverse samples likely reduces construct validity, increases error, and is ultimately done to support

statistical analyses rather than to ensure accurate reflection of gender for the purposes of prediction.

Strengths and Limitations

This study builds on prior research examining different ways to assess gender identity in psychological research and the potential impacts of these questions on TGD participants (Galupo et al., 2017; Herman et al., 2024; Ker et al., 2025). In doing so, we identified novel gaps in current methods of assessing gender as a categorical construct, particularly as these methods pertain to TGD people who use multiple gender identity labels. These findings will be valuable for improving gender inclusivity in future psychological research; however, this study is not without limitations. First, our gender category list was designed to be inclusive, and some of our findings suggest we succeeded in this goal. For instance, some participants noted in their open-ended responses that the listed gender identities were pleasantly inclusive, including one participant who appreciated the inclusion of autigender⁴. However, a drawback of this approach was that it necessarily produced a long list that could be burdensome for participants to review and select from. Relatedly, there are ongoing questions about how to adequately accommodate Indigenous and other culturally specific gender identities (Bauer et al., 2017; Fraser, 2018; Hillier et al., 2010; Mills et al., 2019). Aside from Two-Spirit⁵, we did not include other culturally-specific identities in our gender label list, which may have limited our ability to capture such identities in the current study.

⁴ Autigender or anti-gender is an emerging term embraced by some individuals within the autistic community. The term is used to describe how the experience of gender can be heavily influenced by autism in a way that the two cannot be unlinked; it does not refer to autism as a gender identity (Laube, 2023).

⁵ Although Two-Spirit identity has existed since time in memorial, Two-Spirit is a contemporary and cultural umbrella term specific to (predominantly Canadian) First Nations peoples. It may be used interchangeably to refer to gender identity or sexual orientation, and has significant ties to community roles and knowledge. The term Two-Spirit is also regarded as an act of decolonial practice and reclamation of Indigenous existence in the face of ongoing colonial damage. Two-Spirit has become quite popularized as an all-in-compasing term, but words used to refer to Two-Spirit identity may differ between Nations and languages. As such, additional care, consultation and community-lead teams should be prioritized when conducting Indigenous-specific research (Bauer et al., 2017; Métis Nation of Ontario, 2022; Native Women's Association of Canada, 2025).

Second, resolving specific analytic challenges to more inclusive gender categorization fell beyond the scope of the current analyses. The gender identity questions and available response options would likely yield more categories than are practical for most statistical approaches, depending on one's methodology, sample characteristics, and sample size. Future research studies may consider different approaches to asking about gender identity, whether in combination with or apart from the current categorical approach. Finally, the current study was conducted as part of a larger study recruiting TGD adults who drink alcohol, which could affect generalizability. Similarly, though as intended, our community-based methodology may have encouraged more open responding to the gender identity questions, which means the response patterns may not generalize to studies that are not TGD-focused and/or involve community members. Furthermore, all participants resided in Canada, which is a relatively progressive country with respect to transgender rights (Flores et al., 2016) - although anti-transgender sentiment has, and continues to, rise in Canada at the time of writing this manuscript (Egale, 2024; Ellis, 2025; Johnstone, 2024; Khonina, & Salway, 2025; Ross, 2024). Therefore, it would be important to attempt to replicate the findings in other countries with varying levels of gender inclusivity.

Recommendations and Future Directions

Based on our findings, we offer the following recommendations for researchers aiming to ethically and accurately assess gender identity and incorporate gender-based analyses. These recommendations emphasize the critical need for careful consideration of data use and anticipated categorizations. We acknowledge that these recommendations may be challenging for some researchers to implement at the time of writing this manuscript. The rise in politicization of gender identity and anti-transgender rhetoric has resulted in the implementation of numerous policies that have stripped TGD people of various human rights at an alarming rate, not only across Canada (e.g., Egale Canada, 2024; Johnstone, 2024;

Khonina, & Salway, 2025), but in countries like the United States (Quinan, 2025; Reardon, 2025; Redfield, 2025; The White House, 2025) and the United Kingdom (Human Rights Watch, 2025). Notably, anti-transgender policy in the United States has subsequently impacted academic inquiry, with executive orders that prohibit and limit researchers from asking about and/or pursuing research questions related to gender identity (Fein et al., 2025; Reardon, 2025; Redfield, 2025). These bans harm research inquiry in both 2SLGBTQ+ and cisgender populations and disproportionately impact TGD populations. Not only do they further exacerbate TGD health disparities, but they also infringe on the ways in which researchers can meaningfully and ethically conduct studies about TGD populations and/or ask about gender identity, if at all (Barbee et al., 2022; Brightman et al., 2024; Cramer et al., 2025; Fein et al., 2025; Reardon, 2025; Redfield, 2025; Schanzle et al., 2023). Nevertheless, we encourage researchers to consider the following recommendations in ongoing and future research.

Gender for Demographic Purposes. As others have suggested (e.g., Veldhuis et al., 2024), researchers must first clearly define the specific information they aim to gather when designing demographic survey items. For example, if the goal is to differentiate between the number of transgender and cisgender participants, a direct question such as "Do you identify as transgender and/or gender diverse?" with an explanation of terms may be sufficient. This approach addresses some limitations of the SAAB two-step method, while also acknowledging that not all gender-diverse individuals self-identify as "transgender" (Bauer et al., 2017; Reisner et al., 2014; Ryan, 2019; Suen et al., 2020). In contrast, a two-step method that only asks "Do you identify as transgender?" (and omits other identities, e.g., nonbinary, genderfluid) may inaccurately omit gender diverse participants from relevant analysis. Additionally, Vivienne et al. (2023) suggest that providing a definition of umbrella terms

such as “transgender” and/or “gender diverse” may increase participation among TGD individuals.

Conversely, if the research question necessitates collecting data about participants' sex and gender identity, it is crucial to offer response options beyond the binary of male and female. This includes acknowledging intersex identity (e.g., intersex assigned at birth, intersex assigned after birth) and providing an option for participants, particularly TGD individuals, who may prefer not to disclose their sex assigned at birth (e.g., "prefer not to say") (Bauer et al., 2017; Conron et al., 2008; Fraser, 2018; Puckett et al., 2020; The GenIUSS Group, 2013). In our study, some participants responded positively to how our gender identity questions were prefaced. This suggests that questions should also be clear in their wording; the way questions are constructed can and does have implications for participants. For instance, asking “What is your *gender identity*,” accompanied with a definition of the term, is better aligned with contemporary inclusive terminology, rather than asking “What is your gender?” Although this clarification may seem marginal to some, consider that SAAB and gender have often been conflated within research (Bauer et al., 2017; Sullivan, 2020), and that harms perpetuated by researchers and medical professionals are frequently reported by TGD individuals (de Vries et al., 2020; Sullivan, 2020; Vincent, 2018).

These harms have and continue to pose concerns for TGD participants when answering questions about SAAB and/or gender identity (Conron et al., 2008; Suen et al., 2020). As our research suggests, these residual harms may also impact how TGD participants feel and ultimately respond to these questions. For assessing gender identity, best practice recommends using select-all-that-apply (SATA) checklists that also include a free-text self-describe option (Pao et al., 2025; Veldhuis et al., 2024). These self-described responses can further benefit researchers by informing and refining SATA options in future studies as gender terminology evolves (Moody et al., 2013; Pao et al., 2025). Moreover, it is essential to

include a diverse range of identities beyond common options such as man, woman, non-binary, and other. Researchers should avoid using "other" as a catch-all for TGD participants. Finally, it is vital to incorporate Indigenous and culturally specific gender identities relevant to the research context (e.g., Two-Spirit in Canadian and American research; Bauer et al., 2017; Fraser, 2018; Hillier et al., 2010; Mills et al., 2019).

In addition to these methodological considerations, researchers must recognize that asking TGD participants to disclose details about their identities and experiences requires establishing trust through transparency, community engagement, and a demonstrated commitment to participant safety and autonomy. Indeed, the current political climate may increase TGD people's concerns about safety and privacy when participating in research, particularly if multiple questions about gender identity are asked. To address these legitimate safety concerns and establish researcher-participant trust, researchers should be transparent about data security measures, clearly communicate how gender data will be used and protected, and explicitly state their commitment to the well-being of the TGD community in recruitment materials and consent forms. Additionally, researchers should actively consult with TGD individuals (e.g., hiring TGD researchers, establishing CABs, referencing 2SLGBTQ+ organization glossaries, etc.) to identify relevant terms, ideally through community-engaged research practices to build trust and rapport with the TGD community and participants (Collins et al., 2018; Dermody et al., 2025; Levitt et al., 2018).

Gender for Analytic Purposes. If researchers plan to include gender as a covariate or predictor, they should anticipate the need to recategorize participants and must obtain their informed consent to do so. Although recategorizing gender may at times be a requirement for certain analytic approaches, our findings suggest that doing so may also erase already marginalized gender identities. Researchers should aim to balance foreseeable analytic needs with fundamental ethical principles. For instance, the Belmont Report's (U.S. Department of

Health, Education, and Welfare, 1978) principle of “Respect for Persons” is often reflected in the process of informed consent, in which participants are provided with all necessary information to make an informed choice about whether to participate. In the spirit of this principle and consistent with our findings, we recommend that researchers obtain informed consent from participants prior to any gender recategorization processes to aid analyses. Cheung and colleagues (2020) applied a similar technique, using a pre-survey assessment asking participants to select the gender identity they would like to be grouped with, should the need arise. Providing contextual information while obtaining consent can enhance the accuracy of participant responses and overall findings (Glen & Hurrell, 2012; Tourangeau & Yan, 2007). Given our findings, we anticipate that most individuals will provide this consent, especially when researchers clearly describe the limitations of data analysis and explain the reasoning for recategorization. When participants do not provide consent (e.g., they refuse to answer or select a single gender), researchers must not penalize them for their decision. Erasing their data from analyses, which commonly occurs by default when a predictor variable is missing in most analytic models, constitutes a form of penalty. Alternatively, researchers could utilize sensitivity analyses, both with and without the gender predictor, to ensure the inclusion of all participants in the study, regardless of their choice regarding recategorization.

Often, researchers report frequencies or percentages for the gender categories included in their analytic models (Vivienne et al., 2023). Based on our study findings, these variables often do not accurately represent gender diversity because researchers frequently construct them and/or fail to account for the widespread selection of multiple gender identities in gender diverse samples. Therefore, researchers must describe the gender diversity of their sample using the original survey items (not just the analytic variables) and detail how recategorization took place. We therefore recommend that researchers present both

multi-select and single-select responses side by side when reporting findings that analyze gender variables.

A crucial future direction involves a fundamental re-evaluation of how researchers conceptualize and study gender as a construct (Suárez, 2024). Many of the persistent limitations inherent in categorical gender measurement could yield to a shift towards dimensional representations. For example, researchers have successfully employed dimensional approaches to constructs such as femininity and masculinity (e.g., Magliozzi et al., 2016). However, gender encompasses sub-constructs that extend beyond masculinity and femininity; indeed, our findings and those of others (e.g., Galupo et al., 2017; Marlar & Willcoxon, 2023) suggest that multiplicity is an oft-overlooked component of gender for TGD people. Therefore, future research must identify and incorporate additional dimensions to capture the broader spectrum of gender beyond these binary frameworks. This evolution in measurement would allow for a more nuanced and comprehensive understanding of individual gender experiences, moving beyond the inherent constraints of fixed categories.

Conclusions

As researchers more widely adopt gender-inclusive survey items and variables, it is critical to acknowledge and address the limitations of the approaches commonly used to collect and categorize this information. To this end, collaborating with TGD individuals to develop the survey items and consent process for gender categorization led to a practical and gender inclusive approach to capturing gender diversity in our sample. Nonetheless, there remain challenges in this practice, such as distress in some participants and concerns that these categories do not adequately reflect gender identities in the present day. Moving forward, it is important that researchers carefully consider how they assess gender, as well as how they pose and phrase gender-based questions, and how participants' responses will be presented and recognized in their study. Obtaining consent from participants about

recategorizing their gender for research purposes is one way to mitigate potential harms and gender erasure. Nonetheless, researchers must continue to carefully consider how they assess gender and use it in their analyses to ensure that their assessments most closely reflect reality, as well as the underlying diversity expected to relate to the outcomes of interest.

PROOF COPY

References

- Abramovich, A., & Michel, C.L. (2025). *Better sexual orientation and gender identity questions get better results*. Homeless Hub.
<https://homelesshub.ca/blog/2025/better-sexual-orientation-and-gender-identity-questions-get-better-results/>
- Barbee, H., Deal, C., & Gonzales, G. (2022). Anti-transgender legislation - A public health concern for transgender youth. *JAMA Pediatrics*, 176(2), 125-126.
<https://doi.org/10.1001/jamapediatrics.2021.4483>
- Bauer, G. R., Braimoh, J., Scheim, A. I., & Dharma, C. (2017). Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. *PLoS One*, 12(5). <https://doi.org/10.1371/journal.pone.0178043>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352.
<https://doi.org/10.1080/14780887.2020.1769238>
- Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and becoming a knowing researcher. *International Journal of Transgender Health*, 24(1), 1–6. <https://doi.org/10.1080/26895269.2022.2129597>
- Brightman, S., Lenning, E., Lurie, K. J., & DeJon, C. (2024). Anti-transgender ideology, laws, and homicide: An analysis of the trifecta of violence. *Homicide Studies*, 28(3), 251–269. <https://doi.org/10.1177/10887679231201803>
- Broussard, K. A., Warner, R. H., & Pope, A. R. (2018). Too many boxes, or not enough? preferences for how we ask about gender in cisgender, LGB, and gender-diverse samples. *Sex Roles*, 78(9–10), 606–624. <https://doi.org/10.1007/s11199-017-0823-2>

- Brunson, J. C. (2020). ggalluvial: Layered grammar for alluvial plots. *Journal of Open Source Software*, 5(49), 2017. <https://doi.org/10.21105/joss.02017>
- Brunson, J. C., & Read, Q. D. (2023). *ggalluvial: Alluvial plots in 'ggplot2'* (R package version 0.12.5). <http://corybrunson.github.io/ggalluvial/>
- Cahill, S., Singal, R., Grasso, C., King, D., Mayer, K., Baker, K., & Makadon, H. (2014). Do ask, do tell: High levels of acceptability by patients of routine collection of sexual orientation and gender identity data in four diverse American community health centers. *PLoS One*, 9(9). <https://doi.org/10.1371/journal.pone.0107104>
- Callander, D., Newman, C. E., Holt, M., Rosenberg, S., Duncan, D. T., Pony, M., Timmins, L., Cornelisse, V., Duck-Chong, L., Wang, B., & Cook, T. (2021). The complexities of categorizing gender: A hierarchical clustering analysis of data from the first Australian trans and gender diverse sexual health survey. *Transgender Health*, 6(2), 74–81. <https://doi.org/10.1089/trgh.2020.0050>
- Cameron, J. J., & Stinson, D. A. (2019). Gender (mis)measurement: Guidelines for respecting gender diversity in psychological research. *Social and Personality Psychology Compass*, 13(11). <https://doi.org/10.1111/spc3.12506>
- Cheung, A.S., Leemaqz, S.Y., Wong, J.W.P. et al. Non-Binary and Binary Gender Identity in Australian Trans and Gender Diverse Individuals. *Arch Sex Behav* 49, 2673–2681 (2020). <https://doi.org/10.1007/s10508-020-01689-9>
- Collins, S. E., Clifasefi, S. L., Stanton, J., The Leap Advisory Board, Straits, K. J. E., Gil-Kashiwabara, E., Rodriguez Espinosa, P., Nicasio, A. V., Andrasik, M. P., Hawes, S. M., Miller, K. A., Nelson, L. A., Orfaly, V. E., Duran, B. M., & Wallerstein, N. (2018). Community-based participatory research (CBPR): Towards equitable involvement of community in psychology research. *The American Psychologist*, 73(7), 884–898. <https://doi.org/10.1037/amp0000167>

- Conron, K. J., Scout, & Austin, S. B. (2008). "Everyone has a right to, like, check their box:" Findings on a measure of gender identity from a cognitive testing study with adolescents. *Journal of LGBT Health Research*, 4(1), 1–9.
<https://doi.org/10.1080/15574090802412572>
- Cramer, R. J., Fording, R. C., Kaniuka, A. R., Tucker, R. P., Diaz-Garelli, F., Hill, R. M., Brooks, B., & Hanson, B. (2025). Hate crime law associations with mental health and discrimination experiences among transgender and gender diverse adults. *Analyses of Social Issues and Public Policy*, 25(1). <https://doi.org/10.1111/asap.12447>
- de Vries, E., Kathard, H., & Müller, A. (2020). Debate: Why should gender-affirming health care be included in health science curricula? *BMC Medical Education*, 20(1).
<https://doi.org/10.1186/s12909-020-1963-6>
- Deutsch, M. B. (2016). Making it count: Improving estimates of the size of transgender and gender nonconforming populations. *LGBT Health*, 3(3), 181–185.
<https://doi.org/10.1089/lgbt.2016.0013>
- DeGagne, A. (2021). Protecting cisnormative private and public spheres: The Canadian conservative denunciation of transgender rights. *Studies in Social Justice*, 15(3), 497–517. <https://doi.org/10.26522/ssj.v15i3.2440>
- Dermody, S. S., Penta, S. M., Quinn, T. F., Uhrig, A., Wardell, J., Hart, T. A., Hendershot, C. S., Saewyc, E., & Abramovich, A. (2025). An ecological momentary assessment study of daily and momentary predictors of alcohol outcomes in transgender and gender diverse youth. *Alcohol: Clinical and Experimental Research*, 49(7), 1542-1553. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/acer.70081>
- Dermody, S. S., Quinn, T. F., Courtice, E. L., Hart, T. A., & Abramovich, A. (2025). Adopting an open science mindset to community-based research practices.

International Journal of Transgender Health, 1-3.

<https://doi.org/10.1080/26895269.2025.2596827>

Deutsch, M. B., Green, J., Keatley, J., Mayer, G., Hastings, J., Hall, A. M., Allison, R., Blumer, O., Brown, S., Cody, M. K., Fennie, K., Moscoe, G., St Claire, R., Stone, M. H., Wilson, A., & Wolf-Gould, C. (2013). Electronic medical records and the transgender patient: Recommendations from the World Professional Association for Transgender Health EMR Working Group. *Journal of the American Medical Informatics Association*, 20(4), 700–703.

<https://doi.org/10.1136/amiajnl-2012-001472>

Egale Canada. (2024, December). Brief on hate motivated violence targeting 2SLGBTQI communities.

<https://www.ourcommons.ca/Content/Committee/441/FEWO/Brief/BR13528939/br-external/EgaleCanada-e.pdf>

Ellis, T. (2025, April 22). *Queer Momentum responds to anti-trans policy in Conservative Party of Canada platform*. Momentum. <https://www.momentumcanada.net/cpc>

Fein, L. A., Tummarakota, S., Barnett, R., & Danker, S. (2025). The drastic toll of anti-trans legislation: Compromised access to health care and lost research opportunities. *Clinical Therapeutics*, 47(9), 665-666. <https://doi.org/10.1016/j.clinthera.2025.07.011>

Flores, A. R., Brown, T. N. T., & Park, A. (2017). *Public support for transgender rights: A twenty-three country survey*. The Williams Institute, UCLA School of Law. <https://williamsinstitute.law.ucla.edu/wp-content/uploads/Public-Opinion-Trans-23-Countries-Dec-2016.pdf>

Fraser, G. (2018). Evaluating inclusive gender identity measures for use in quantitative psychological research. *Psychology & Sexuality*, 9(4), 343–357.

<https://doi.org/10.1080/19419899.2018.1497693>

- Galupo, M.P., Mitchell, R.C., & Davis, K.S. (2015). Sexual minority self-identification: Multiple identities and complexity. *Psychology of Sexual Orientation and Gender Diversity*, 2(4), 355–364.
<https://doi-org.ezproxy.lib.torontomu.ca/10.1037/sgd0000131>
- Galupo, M. P., Pulice-Farrow, L., & Ramirez, J. L. (2017). “Like a constantly flowing river”: Gender identity flexibility among nonbinary transgender individuals. In J. Sinnott (Ed.), *Identity flexibility during adulthood* (pp. 163–177). Springer.
https://doi.org/10.1007/978-3-319-55658-1_10
- Glen, F., & Hurrell, K. (2012). *Measuring gender identity*. Equality and Human Rights Commission.
https://www.equalityhumanrights.com/sites/default/files/technical_note_final.pdf
- Glick, J. L., Theall, K., Andrinopoulos, K., & Kendall, C. (2018). For data’s sake: Dilemmas in the measurement of gender minorities. *Culture, Health & Sexuality*, 20(12), 1362–1377. <https://doi.org/10.1080/13691058.2018.1437220>
- Goldberg, A. E., Kuvalanka, K., & Dickey, L. (2019). Transgender graduate students’ experiences in higher education: A mixed-methods exploratory study. *Journal of Diversity in Higher Education*, 12(1), 38–51. <https://doi.org/10.1037/dhe0000074>
- Guyan, K. (2022). *Queer data*. Bloomsbury Publishing.
- Harrison-Quintana, J., Grant, J. M., & Herman, J. L. (2012). *A gender not listed here: Genderqueers, gender rebels, and otherwise in the national transgender discrimination survey*. The Williams Institute.
<http://williamsinstitute.law.ucla.edu/wp-content/uploads/Harrison-Herman-Grant-AGender-Apr-2012.pdf>
- Hillier, L., Jones, T., Monagle, M., Overton, N., Gahan, L., Blackman, J., & Mitchell, A. (2010). *Writing themselves in 3: The third national study on the sexual health and*

wellbeing of same sex attracted and gender questioning young people. Australian Research Centre in Sex, Health and Society, La Trobe University.

<https://www.acon.org.au/wp-content/uploads/2015/04/Writing-Themselves-In-3-2010.pdf>

Herman, R., Cavic, E., Puckett, J. A., Ran, D., Galupo, M. P., Garrett-Walker, J., & Veldhuis, C. B. (2024). Balancing community and research needs in gender measurement decisions. *Sexuality Research and Social Policy*, 22, 665–681.

<https://doi.org/10.1007/s13178-024-00994-5>

Human Rights Watch. (2025, May 9). *UK: Court ruling threatens trans people*.

<https://www.hrw.org/news/2025/05/09/uk-court-ruling-threatens-trans-people>

Jacobsen, K., Davis, C. E., Burchell, D., Rutherford, L., Lachowsky, N., Bauer, G., & Scheim, A. (2024). Misgendering and the health and wellbeing of nonbinary people in Canada. *International Journal of Transgender Health*, 25(4), 816–830.

<https://doi.org/10.1080/26895269.2023.2278064>

Johnson, K. C., LeBlanc, A. J., Deardorff, J., & Bockting, W. O. (2020). Invalidation experiences among non-binary adolescents. *The Journal of Sex Research*, 57(2), 222–233. <https://doi.org/10.1080/00224499.2019.1608422>

Johnstone, F. (2024, October 31). *Alberta Premier Danielle Smith betrays Albertan values, attacks freedom of trans youth and families*. Momentum.

<https://www.momentumcanada.net/ab-release>

Ker, A., Tan, K., Fraser, G., & Fenaughty, J. (2025). Evaluating an inclusive methodology for collecting and coding gender data in sexuality- and gender-diverse youth survey research. *Psychology of Sexual Orientation and Gender Diversity*. Advance online publication. <https://doi.org/10.1037/sgd0000883>

Khonina, M., & Salway, T. (2025). The rise of anti-trans laws and the role of public health advocacy. *Canadian Journal of Public Health*, 116, 97-99.

<https://doi.org/10.17269/s41997-024-00942-1>

Kidd, K. M., Sequeira, G. M., Rothenberger, S. D., Paglisotti, T., Kristjansson, A., Schweiberger, K., Miller, E., & Coulter, R. W. (2022). Operationalizing and analyzing 2-step gender identity questions: Methodological and ethical considerations. *Journal of the American Medical Informatics Association*, 29(2), 249–256.

<https://doi.org/10.1093/jamia/ocab137>

Kuper, L. E., Nussbaum, R., & Mustanski, B. (2012). Exploring the diversity of gender and sexual orientation identities in an online sample of transgender individuals. *The Journal of Sex Research*, 49(2–3), 244–254.

<https://doi.org/10.1080/00224499.2011.596954>

Lagos, D., & Compton, D. (2021). Evaluating the use of a two-step gender identity measure in the 2018 general social survey. *Demography*, 58(2), 763–772.

<https://doi.org/10.1215/00703370-8976151>

Laube, A. (2023, June 22). *Understanding anti-gender and anti-ace identities*. AutismBC.

<https://www.autismbc.ca/blog/advocacy/anti-gender-anti-ace-identities/>

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 26-46.

<https://doi.org/10.1037/amp0000151>

Lombardi, E., & Banik, S. (2016). The utility of the two-step gender measure within trans and cis populations. *Sexuality Research and Social Policy*, 13(3), 288–296.

<https://doi.org/10.1007/s13178-016-0220-6>

- MacEntee, K., Elkington, N., Segui, J., & Abramovich, A. (2024). Unveiling the pathways: Mapping and understanding hidden homelessness among 2SLGBTQ+ youth in Ontario. *Youth*, 4(3), 1244-1237. 10.3390/youth4030077
- Magliozzi, D., Saperstein, A., & Westbrook, L. (2016). Scaling up: Representing gender diversity in survey research. *Socius: Sociological Research for a Dynamic World*, 2. <https://doi.org/10.1177/2378023116664352>
- Marlar, J., & Willcoxon, N. (2023, August 22). Measuring diverse gender identities: Phase 2. *Gallup*. <https://news.gallup.com/opinion/methodology/509843/measuring-diverse-gender-identities-phase.aspx>
- Métis Nation of Ontario. (2022). *Two-Spirit Métis awareness resource*. https://www.metisnation.org/wp-content/uploads/2022/09/2S-Awareness-Resource_2022a.pdf
- Mills, S., Dion, M., Thompson-Blum, D., Borst, C., & Diemert, J. (2019). *Mapping the void: Two-Spirit and LGBTQ+ experiences in Hamilton*. Social Planning and Research Council of Hamilton. <https://macsphere.mcmaster.ca/bitstream/11375/28172/1/mappingthevoid-for-web.pdf>
- Moody, C., Obear, K., Gasser, H., Cheah, S., & Fechter, T. (2013). *ACPA standards for demographic questions*. American College Personnel Association. <https://archive.myacpa.org/sites/default/files/Proposal-Demographic-Questions-and-Responses-2.pdf>
- Muehlenhard, C. L., & Peterson, Z. D. (2011). Distinguishing between sex and gender: History, current conceptualizations, and implications. *Sex Roles*, 64, 791-803. <https://doi.org/10.1007/s11199-011-9932-5>

Native Women's Association of Canada. (2025). *Intersections: Indigenous and 2SLGBTQQIA+ Identities*.

https://nwac.ca/assets-knowledge-centre/29-June-2S_Intersections_Booklet_V2.pdf

Pao, C., Julian, C., Compton, D., Lagos, D., & Stacey, L. (2025). Demographic differences in responses to a two-step gender identity measure. *Sociological Science*, 12, 277–293.

<https://doi.org/10.15195/v12.a13>

Park, A. (2016, March). *Data collection methods for sexual orientation and gender identity*. The Williams Institute, UCLA School of Law.

<https://williamsinstitute.law.ucla.edu/publications/data-collection-sogi/>

Patterson, J. G., Jabson, J. M., & Bowen, D. J. (2017). Measuring sexual and gender minority populations in health surveillance. *LGBT Health*, 4(2), 82–105.

<https://doi.org/10.1089/lgbt.2016.0026>

Pinto, A. D., Aratangy, T., Abramovich, A., Devotta, K., Nisenbaum, R., Wang, R., & Kiran, T. (2019). Routine collection of sexual orientation and gender identity data: A mixed-methods study. *Canadian Medical Association Journal*, 191(3).

<https://doi.org/10.1503/cmaj.180839>

Poquiz, J. L., Coyne, C. A., Garofalo, R., & Chen, D. (2021). Comparison of gender minority stress and resilience among transmasculine, transfeminine, and nonbinary adolescents and young adults. *Journal of Adolescent Health*, 68(3), 615–618.

<https://doi.org/10.1016/j.jadohealth.2020.06.014>

Pradell, L., Parmenter, J. G., Galliher, R. V., Wong, E. G., Rowley, L., Huenemann, H., & South, S. (2024). The sexual objectification experiences of non-binary people: Embodied impacts and acts of resistance. *Sex Roles*, 90(2), 318–335.

<https://doi.org/10.1007/s11199-024-01447-4>

Puckett, J. A., Brown, N. C., Dunn, T., Mustanski, B., & Newcomb, M. E. (2020).

Perspectives from transgender and gender diverse people on how to ask about gender.

LGBT Health, 7(6), 305–311. <https://doi.org/10.1089/lgbt.2019.0295>

Quinan, C. (2025). From criminalization to erasure: Project 2025 and anti-trans legislation in the US. *Crime, Media, Culture: An International Journal*.

<https://doi.org/10.1177/17416590241312149>

Quinn, T. F., Uhrig, A., Dermody, S. S., Hart, T. A., & Abramovich, A. (2026). Exploring

alcohol use and harm reduction techniques among transgender and gender diverse

youth in Canada: A qualitative study. *International Journal of Transgender Health*,

27, 870-887. <https://doi.org/10.1080/15532739.2025.2471903>

Rankin, S., Weber, G. N., Blumenfeld, W. J., & Frazer, S. (2010). *2010 state of higher education for lesbian, gay, bisexual & transgender people*. Campus Pride.

<https://www.campuspride.org/wp-content/uploads/campuspride2010lgbtreportsummary.pdf>

Rankin, S., & Garvey, J. C. (2015). Identifying, quantifying, and operationalizing

queer-spectrum and trans-spectrum students: Assessment and research in student affairs. *New Directions for Student Services*, 2015(152), 73–84.

<https://doi.org/10.1002/ss.20146>

Reardon, S. (2025, March 6). New NIH grant terminations target transgender studies - even in mice. *Science*.

<https://www.science.org/content/article/new-nih-grant-terminations-target-transgender-studies-even-mice>

Redfield, E. (2025). *Impact of the ban on gender-affirming care on transgender minors*.

UCLA School of Law, Williams Institute.

<https://williamsinstitute.law.ucla.edu/wp-content/uploads/GAC-Access-EO-Jan-2025.pdf>

- Reisner, S. L., Biello, K., Rosenberger, J. G., Austin, S. B., Haneuse, S., Perez-Brumer, A., Novak, D. S., & Mimiaga, M. J. (2014). Using a two-step method to measure transgender identity in Latin America/the Caribbean, Portugal, and Spain. *Archives of Sexual Behavior*, 43(8), 1503–1514. <https://doi.org/10.1007/s10508-014-0314-2>
- Rich-Edwards, J. W., Kaiser, U. B., Chen, G. L., Manson, J. E., & Goldstein, J. M. (2018). Sex and gender differences research design for basic, clinical, and population studies: Essentials for investigators. *Endocrine Reviews*, 39(4), 424–439. <https://doi.org/10.1210/er.2017-00246>
- Ross, K. (2024, February 20). *CMA strongly opposes government efforts to restrict access to care*. Canadian Medical Association. <https://www.cma.ca/about-us/what-we-do/press-room/cma-strongly-opposes-government-efforts-restrict-access-care>
- Ryan, J. M. (2019). The problematics of assessing trans identity in survey research: A modest proposal for improving question design. *Societies*, 9(4), 85. <https://doi.org/10.3390/soc9040085>
- Schanzle, J., Kennedy, J., Rahman, F., & Hill, S. (2023). Anti-trans legislation in the US: Potential implications on self-reported victimization and suicidality among trans youth. *Journal of Adolescent Health*, 72(3), S17-S18. <https://doi.org/10.1016/j.jadohealth.2022.11.018>
- Scott, D., Derrett, S., Rupel, V. P., Jelsma, J., Gurung, G., Oduro, G. Y., & Withey-Rila, C. (2024). He/she/they - gender inclusivity in developing and using health-related questionnaires: A scoping review. *Quality of Life Research*, 34(1), 67–87. <https://doi.org/10.1007/s11136-024-03765-2>

- Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. In R. Z. Litten & J. P. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biochemical methods* (pp. 41–72). Humana Press/Springer Nature. https://doi.org/10.1007/978-1-4612-0357-5_3
- Springer, M. V., & Skolarus, L. E. (2019). Community-based participatory research. *Stroke* (1970), 50(3), e48. <https://doi.org/10.1161/STROKEAHA.118.024241>
- Statistics Canada. (2022, July 7). *Age, sex at birth and gender reference guide, census of population, 2021*.
<https://www12.statcan.gc.ca/census-recensement/2021/ref/98-500/014/98-500-x2021014-eng.cfm>
- Suárez, M. I. (2024). Problematizing the measurement of gender identity in K-12 education survey research: A systematic review. *Educational Studies*, 60(1), 68–88.
<https://doi.org/10.1080/00131946.2023.2248636>
- 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(7), 2301–2318.
<https://doi.org/10.1007/s10508-020-01810-y>
- Sullivan, A. (2020). Sex and the census: Why surveys should not conflate sex and gender identity. *International Journal of Social Research Methodology*, 23(5), 517–524.
<https://doi.org/10.1080/13645579.2020.1768346>
- Tate, C. C., Ledbetter, J. N., & Youssef, C. P. (2013). A two-question method for assessing gender categories in the social and medical sciences. *Journal of Sex Research*, 50(8), 767–776. <https://doi.org/10.1080/00224499.2012.690110>

The GenIUSS Group. (2013). *Gender-related measures overview*. The Williams Institute, UCLA School of Law.

<https://williamsinstitute.law.ucla.edu/wp-content/uploads/Gender-Related-Measures-Overview-Feb-2013.pdf>

The White House. (2025, January 28). Protecting children from chemical and surgical mutilation. The White House, Presidential Actions.

<https://www.whitehouse.gov/presidential-actions/2025/01/protecting-children-from-chemical-and-surgical-mutilation/>

Tordoff, D. M., Minalga, B., Ó Catháin, N., Fernandez, A., Gross, B., Glick, S. N., & Seattle Trans and Nonbinary Sexual Health (STARS) Advisory Board. (2025). Comparing two-step approaches to measuring gender identity: The reliability and applications of asking about sex assigned at birth vs transgender self-identification. *American Journal of Epidemiology*, 194(5), 1255–1263. <https://doi.org/10.1093/aje/kwae341>

Tordoff, D. M., Morgan, J., Dombrowski, J. C., Golden, M. R., & Barbee, L. A. (2019). Increased ascertainment of transgender and non-binary patients using a 2-step versus 1-step gender identity intake question in an STD clinic setting. *Sexually Transmitted Diseases*, 46(4), 254–259. <https://doi.org/10.1097/olq.0000000000000952>

Torgimson, B. N., & Minson, C. T. (2005). Sex and gender: what is the difference?. *Journal of Applied Physiology*, 99(3), 785–787. <https://doi.org/10.1152/japplphysiol.00376.2005>

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>

Trans Legislation Tracker. (2025). 2025 anti-trans bills. <https://translegislation.com/>

U.S. Department of Health, Education, and Welfare. (1978). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.

<https://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>

Urlacher, B. R. (2022). Gender identification and survey weighting: A shifting landscape. *PS: Political Science & Politics*, 56(1), 56–60.

<https://doi.org/10.1017/s1049096522001081>

Veldhuis, C. B., Cascalheira, C. J., Delucio, K., Budge, S. L., Matsuno, E., Huynh, K., Puckett, J. A., Balsam, K. F., Velez, B. L., & Galupo, M. P. (2024). Sexual orientation and gender diversity research manuscript writing guide. *Psychology of Sexual Orientation and Gender Diversity*, 11(3), 365–396.

<https://doi.org/10.1037/sgd0000722>

Vincent, B. W. (2018). Studying trans: Recommendations for ethical recruitment and collaboration with transgender participants in academic research. *Psychology & Sexuality*, 9(2), 102–116. <https://doi.org/10.1080/19419899.2018.1434558>

Vivienne, S., Hanckel, B., Byron, P., Robards, B., & Churchill, B. (2023). The social life of data: Strategies for categorizing fluid and multiple genders. *Journal of Gender Studies*, 32(5), 498–513. <https://doi.org/10.1080/09589236.2021.2000852>

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York. ISBN 978-3-319-24277-4. https://doi.org/10.1007/978-3-319-24277-4_9