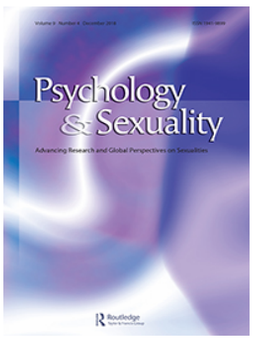


Psychology & Sexuality

ISSN: 1941-9899 (Print) 1941-9902 (Online) Journal homepage: http://www.tandfonline.com/loi/rpse20

Evaluating inclusive gender identity measures for use in quantitative psychological research

Gloria Fraser

To cite this article: Gloria Fraser (2018) Evaluating inclusive gender identity measures for use in quantitative psychological research, Psychology & Sexuality, 9:4, 343-357, DOI: 10.1080/19419899.2018.1497693

To link to this article: https://doi.org/10.1080/19419899.2018.1497693

Accepted author version posted online: 07 Jul 2018.
Published online: 01 Aug 2018.

PSYCHOLOGY & SEXUALITY
2018, VOL. 9, NO. 4, 343–357
https://doi.org/10.1080/19419899.2018.1497693

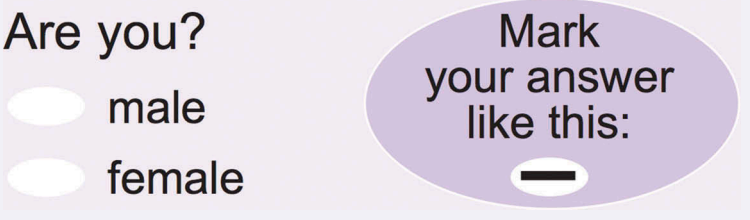
Evaluating inclusive gender identity measures for use in quantitative psychological research

Gloria Fraser

School of Psychology, Victoria University of Wellington, Wellington, New Zealand

ABSTRACT

Despite growing awareness of the discrimination, violence and health disparities experienced by transgender people, most researchers in psychology assess gender identity using a binary categorical measure: 'male' and 'female'. This standard generates imprecise gender identity data, which has significant implications for the social inclusion and well-being of the transgender community. Here, I argue that inclusive gender identity measures should be implemented across all fields of psychology, assess options for collecting gender identity data and offer specific advice for psychological researchers interested in the science of gender identity measurement.



KEYWORDS

Gender; gender identity; transgender; research methodology; survey design; data collection

Introduction

> Being countable – including our gender identity and gender expression – is an important step for having the resources we need to live healthy, safe, financially stable lives.
>
> (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014, p. xv)

Public awareness of gender diversity issues has heightened considerably in recent years, yet the question of how best to measure gender identity has been largely ignored in mainstream psychological research. The overwhelming majority of survey designs still opt for a single dichotomous measure with two possible response options: 'male' and 'female' (Magliozzi, Saperstein, & Westbrook, 2016). An emerging literature acknowledging the fluid and non-binary nature of gender identity highlights a multitude of problems with this traditional approach (see Deutsch, 2016; Frohard-Dourlent, Dobson, Clark, Doull, & Saewyc, 2016; GenIUSS Group, 2014; Reisner, Conron, et al., 2014). Notably, researchers using a single dichotomous item to measure gender cannot identify transgender participants (those whose gender identity differs from their assigned

CONTACT Gloria Fraser Gloria.Fraser@vuw.ac.nz School of Psychology, Victoria University of Wellington, Kelburn Parade, P.O. Box 600, Wellington 6012, New Zealand

This article was originally published with error. This version has been amended. Please see Erratum (http://doi.org/10.1080/19419899.2018.1515167)

sex at birth; Serano, 2007).[1] Moreover, some respondents with non-binary genders (those whose gender identity does not fit the common cultural binary of women and men; Barker, 2016) are unable to report their gender identity as they cannot select either gender option.[2]

Here, I recommend the use of inclusive gender identity measures in quantitative psychological research and assess options for collecting gender identity data. By 'inclusive measures', I refer to those that (1) recognise that not all respondents are cisgender (i.e. that not all respondents' gender identity is the same as their sex assigned at birth; Serano, 2007), and (2) capture data from people of all genders. I use 'transgender' as an umbrella term to refer to all people whose gender identity differs from their assigned sex at birth (Bauer et al., 2009).

The importance of inclusive gender identity measures

The last few decades have seen a dramatic increase in academic literature documenting the experiences of transgender people (MacCarthy, Reisner, Nunn, Perez-Brumer, & Operario, 2015; Winter, 2017). Over half of transgender people report having experienced some form of harassment or violence (James et al., 2016; Lombardi, Wilchins, Priesing, & Malouf, 2002). Results from the Community Health Centre Core Data Project revealed that 29% of transgender respondents reported a suicide attempt at some point in their lifetime (compared with 8.5% of cisgender participants; Reisner, White, Bradford, & Mimiaga, 2014), and 77% of Trans PULSE survey participants have seriously considered suicide (Bauer, Pyne, Francino, & Hammond, 2013). Rates of depression, anxiety and other mental health problems among transgender individuals far surpass those of the general population (Budge, Adelson, & Howard, 2013).[3]

Despite a booming growth of transgender research, substantial gaps in this body of literature remain. Because studies of transgender health and well-being typically use convenience or snowball samples, the validity of inferences that can be made about the needs of the wider transgender population is limited (Dean et al., 2000; Reisner et al., 2016). This lack of population-based data affects the ability to examine health disparities between transgender and cisgender people, and to develop effective strategies for addressing these disparities (Cahill & Makadon, 2014; Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014). Without inclusive gender identity data from national population-based studies, it is also impossible to accurately determine the size of the transgender population. Previous studies attempt to estimate the proportion of the population who are transgender by using data from specialist gender identity clinics (see Arcelus et al., 2015); however, the little data available from population-based studies implies that the number of transgender people is substantially higher than these studies suggest (Olyslager & Conway, 2007). Given that policy and funding decisions are determined, at least in part, by population size (Deutsch, 2016), underestimating the size of the transgender population has likely resulted in inadequate resourcing for transgender health and social services.

Extant literature on transgender health often treats transgender people as a homogenous group (Chang & Chung, 2015); however, researchers are increasingly recognising the importance of distinguishing the experiences of transfeminine, transmasculine and non-binary people (Vincent, 2018). The US Transgender Survey, for example, highlights that transgender women report higher rates of HIV than transgender men, while lifetime suicide attempt rates for transgender men are higher than those of transgender women and non-binary people (James et al., 2016). Understanding the differing needs and experiences within the transgender community is crucial in order to promote transgender health and well-being, but is impossible without the use of inclusive gender identity measures.

The use of inclusive gender identity measures is also key in developing our understanding of diverse sexualities. In proposing Sexual Configurations Theory (SCT) Van Anders (2015) emphasises the importance of modelling partnered sexuality in a way that is inclusive of diverse gender identities. A failure to acknowledge non-binary genders can, for example, 'serve to make the sexual orientations of some genderqueer (…) and trans individuals and their partners nonsensical' (p.

1181). By collecting inclusive gender identity data, researchers can empirically test emerging theories such as SCT, which presents a more nuanced model of partnered sexuality than traditional theories.

As Cahill and Makadon (2014) point out, lessons learnt from the collection of ethnicity data can inform the collection of gender identity data. Just as it would be inappropriate to offer only two response options on a measure of ethnic identification, restricting the collection of gender data to those who identify as male or female will, in many cases, result in incomplete or inaccurate gender identity data. As such, this methodological issue is not limited to the subfields of LGBT and transgender studies, but demands attention across *all* fields of social science research. Indeed, the considerations discussed here are of relevance to any study in which the gender of participants is recorded.

Finally, it is important to note that research is, itself, a form of knowledge production (Dick, 2004). With the use of single dichotomous measures, researchers are reinforcing the gender binary. By adopting inclusive measures, however, researchers challenge the assumption that gender is binary and communicate that there are more ways of being than just 'male' and 'female'. The use of inclusive gender measures within quantitative research may encourage the use of inclusive language around gender in other contexts, particularly for cisgender people who may be unfamiliar with terms such as 'transgender' or 'non-binary' before encountering them in a survey.

How should gender be measured?

Here, I outline four approaches to measuring gender identity that avoid the pitfalls of a single dichotomous measure. I discuss the advantages and disadvantages of adopting each approach, and describe the research settings in which each approach may be most useful. Because no 'gold standard' measure of gender identity currently exists, researchers must consider the advantages and disadvantages of using each method based on their research question, study design and participant demographics (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014). Figure 1 summarises this section in flow chart form; researchers can use this flow chart, together with the below recommendations, to guide their decision-making around selecting a gender identity measure.

Measuring gender identity using categorical lists

A popular alternative to a single dichotomous measure of gender identity is the use of a single categorical item with an expanded list of response options:

What is your gender identity? Select all that apply.

- Male
- Female
- Trans male/trans man
- Trans female/trans woman
- Non-binary
- Not listed (please state) _____

Research considerations

The list of response options offered by those who have employed this method varies considerably; Bockting, Robinson, Benner, and Scheltema (2004) added one additional response option of 'transgender' to the usual 'male' and 'female' response options, while the National Transgender Discrimination Survey offered a total of 15 possible identity terms, inviting participants to select to

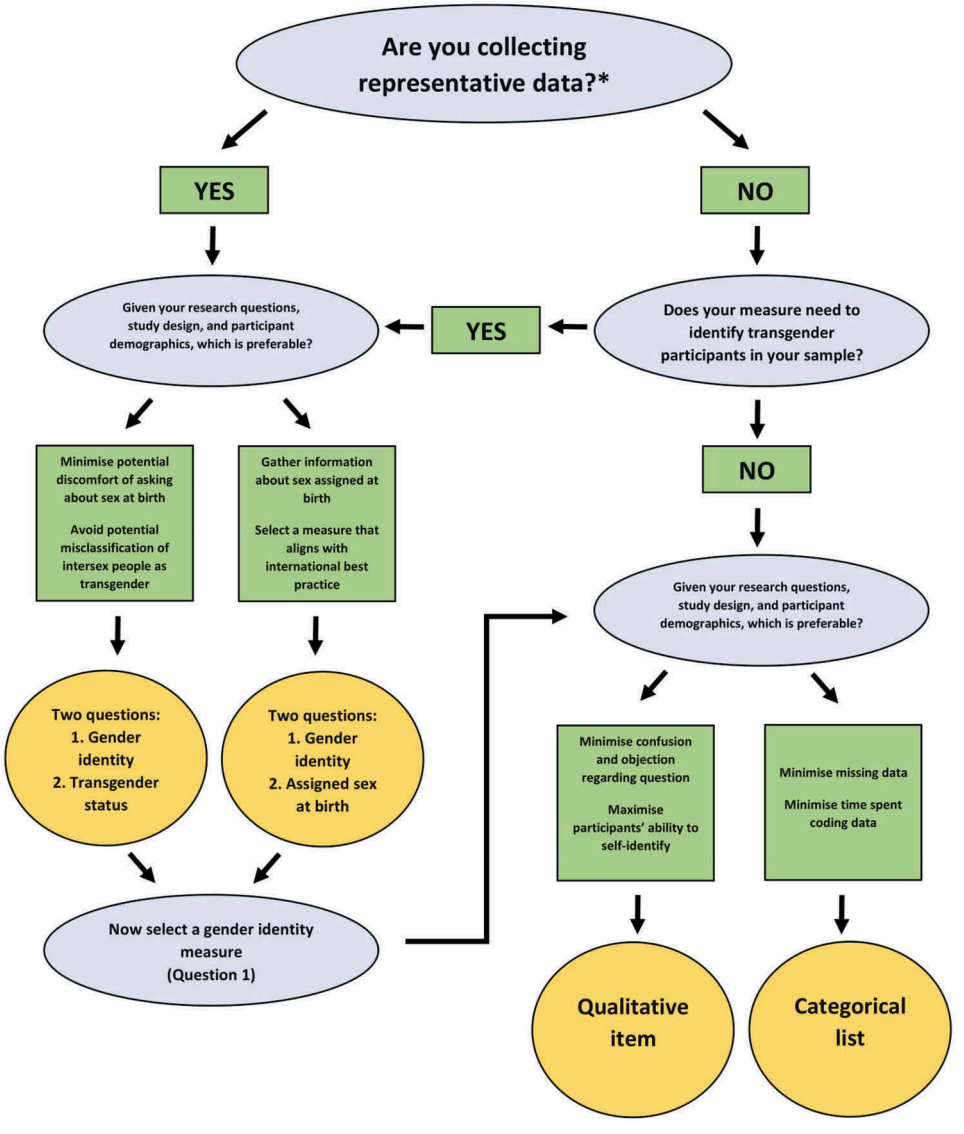


Figure 1. Selecting a gender identity measure: A handy guide.

what degree each term applied to them (Harrison-Quintana, Grant, & Rivera, 2015).[4] Researchers using this approach may attempt to list all possible terms for gender identity to ensure that participants are not excluded from reporting their gender or required to select a term with which they do not identify. Constructing an exhaustive list of genders, however, is an impossibility – the language used for gender identity changes at a rate that survey designers cannot hope to match (Frohard-Dourlent et al., 2016). For this reason, those using this option should always include an open-ended response option to ensure that all respondents can report their gender identity. Where possible, researchers using a categorical list measure should also allow participants to select all options that apply (e.g. 'male' and 'trans male'). This is best practice for collecting gender identity data with categorical lists as it ensures that respondents are not required to choose between identity labels (Ansara & Hegarty, 2014).

Because groups within the transgender community have diverse lived experiences and health-care needs (Cahill & Makadon, 2014), I do not recommend capturing transgender people in a third category of 'other' or 'transgender'. I also advise against this approach because 'transgender' is not in itself a gender, but is generally used as an adjective (Stryker, 2006), so must be accompanied by a noun such as 'male' or 'female' when collecting gender identity data. As such, researchers should include options of 'trans female/woman', 'trans male/man' and 'non-binary'.

When conducting research online, rather than with paper-and-pencil questionnaires, researchers have the option of using skip logic or branching when measuring gender identity. For example, researchers could include options of 'male', 'female' and 'not listed, please state'. Those that select the last option are then directed to an open box, expanded list of options or additional questions regarding their gender identity. Researchers should consider, however, that this could be perceived as othering[5] by transgender participants, as male and female are presented as default options (Figure 2).

While sex and gender researchers typically use the terms male/female in relation to sex and man/woman in relation to gender (e.g. West & Zimmerman, 1987), researchers routinely include 'male' and 'female' as response options to questions about gender identity rather than 'man' and 'woman' (e.g. Bauer, Braimoh, Scheim, & Dharma, 2017; Cahill et al., 2014; Reisner, Conron, et al., 2014). I recommend following this use of 'male' and 'female' in order to facilitate comparisons with existing gender identity data from other studies, and also because this is, for the most part, a technical distinction; colloquially, and in research outside the field of sex and gender, it is not unusual to use the terms 'male' and 'female' to describe gender identity.

Advantages

The categorical list option may minimise rates of missing data as a tick box measure is quick and easy for participants to complete. Data generated using this measure is also swift to code as a relatively small number of participants are likely to provide an open-ended response. As such, this may appeal to researchers working with large data sets (though, see below for a discussion around the feasibility of coding qualitative data in studies with tens of thousands of participants).

Disadvantages

A single item with expanded response options is not sensitive enough to identify all transgender participants as many transgender people do not self-identify as transgender (Bauer et al., 2017; Cahill & Makadon, 2014). For example, a transgender man may select 'male' rather than 'trans male'. Consequently, this approach would not be appropriate for researchers attempting to estimate the size of the transgender population or for whom it is important to identify all transgender people in their sample. This approach may, however, be suitable for researchers whose research questions are not focused on gender but rather wish to accurately describe their participant group.

Measuring gender identity using a single qualitative item

A second alternative measure of gender identity is the use of a single open-ended item that allows participants to self-identify their gender:

- What is your gender identity? _____

Research considerations

Although it is common to include an open-ended item as one of many possible response options in a measure of gender identity (e.g. Harrison-Quintana et al., 2015; Tate, Ledbetter, & Youssef, 2013), an open-ended item is rarely used as the *sole* measure of gender identity. Moreover, there is very little mention of this option in literature on the methodological issue of measuring gender.

Gender identity measure	Advantages	Disadvantages	Suitability
Categorical list	Minimises time spent coding data Might minimise rates of missing data, as a tick box measure is easy and quick for participants to complete	Unlikely to identify all transgender participants Might impair collaborative research if response options differ between studies	Suitable for studies in which it is not necessary to identify all transgender participants Suitable for very large studies where there is limited time to code qualitative data
Single qualitative item	Respondents can self-identify freely and can choose multiple identity terms Takes up little space Allows researchers to track changes in language use over time	Unlikely to identify all transgender participants Can be used in datasets with tens of thousands of participants, but might prove difficult to code in larger datasets	Suitable for studies in which it is not necessary to identify all transgender participants Suitable for studies in which there is a particular interest in self-identification of gender
Two-step approach	Likely to accurately identify a higher proportion of transgender participants than single item measures Aligns with international best practice	Asking about assigned sex at birth could be uncomfortable or distressing for some transgender participants Potential to misclassify some intersex participants as transgender	Suitable for studies in which it is important to identify transgender participants Suitable for representative studies
Single item transgender status	Likely to accurately identify a higher proportion of transgender participants than single item measures Avoid potential discomfort of asking about sex assigned at birth	Limited empirical research available to support the effectiveness of this approach	Suitable for studies in which it is important to identify transgender participants Suitable for representative studies

Figure 2. Advantages, disadvantages and suitability of four inclusive measures of gender identity.

This is somewhat puzzling, given that this option is recommended by a number of researchers (e.g. Ansara & Hegarty, 2014; Rankin & Garvey, 2015) and has proven popular among transgender participants in studies that offer it (Harrison-Quintana et al., 2015). It is possible that the lack of research about open-ended measures comes from the assumption that it is too difficult or time-consuming to measure gender identity using a qualitative measure. A recent study by Greaves and colleagues (2016), however, used an open-ended measure to collect sexual orientation data in the New Zealand Attitudes and Values Study (NZAVS; N = 18,261), suggesting that an open-ended gender identity measure is a feasible option for use in large-scale surveys.

While coding open-ended data may be more time-consuming than coding data from categorical measures, doing so allows researchers to conduct much-needed analyses of differences between subgroups within marginalised communities. The NZAVS team has, for example, published work on the health and well-being of self-identified asexual participants (Greaves et al., 2017), and has a forthcoming paper exploring differences between those who identify as pansexual and bisexual (Greaves, Sibley, Fraser, & Barlow, 2018). This work on the well-being of those outside the LGBT umbrella was made possible because participants of minority sexualities were not grouped into a single 'other' or 'non-heterosexual' category. The NZAVS team also collects gender identity data using an open-ended measure and has produced a coding scheme for use by other researchers coding open-ended gender identity data (see Fraser, 2017).

Advantages

There are numerous advantages to measuring gender using a single open-ended item. Firstly, participants can self-identify freely and can write as many terms as desired in the open-ended field. Secondly, the question takes up very little space on a questionnaire; an attractive feature, given that question space on surveys is often at a premium (Glen & Hurrell, 2012). Finally, this approach allows researchers to track changes in language use around gender identity over time and avoids the use of terminology that may be considered outdated or offensive by some transgender people (e.g. male-to-female and female-to-male).

Disadvantages

Like the single item with expanded response options, a single open-ended item is unlikely to identify all transgender participants. This method would be suitable for researchers who are not specifically studying gender, but wish to be inclusive, as well as researchers interested in how people self-identify their gender, and researchers who fear that some cisgender respondents would become confused or offended by a multitude of gender options (a potential issue I will discuss below).

Any open-ended measure may invite sarcastic or absurd responses from 'mischievous responders' (Robinson-Cimpian, 2014, p. 171) who provide inaccurate data in jest, or for ideological purposes (Jaroszewski, Lottridge, Haimson, & Quehl, 2018). For instance, a common mischievous response to questions around sexuality and gender identity is 'attack helicopter' (for discussion, see KnowYourMeme, 2015). Researchers may view potential mischievous responses as a reason to refrain from collecting open-ended gender data; however, recent research from Robinson-Cimpian (2014) shows that mischievous responders often provide untruthful responses to multiple survey questions. As such, researchers can use responses such as 'attack helicopter' to identify and remove data which may otherwise lead to inaccurate conclusions. If researchers choose to retain data from mischievous responders, they should code gender identity data as missing. In some cases, researchers may have to conduct an Internet search to confirm that the response is not genuine. A useful resource for doing this is www.gender.wikia.com, which explicitly identifies satirical genders.

Measuring gender identity using the two-step approach

The approach most commonly recommended by experts in gender-related measures is the two-step approach, which involves asking participants about both their current gender identity and their sex assigned at birth. This method was developed by the Transgender Health Advocacy Coalition in 1997 and was adopted in 2007 by the Centre of Excellence for Transgender Health:

(1) What is your gender identity? _____
(2) What sex was documented at birth on your original birth certificate?
- Male
- Female

Research considerations

Researchers can choose to measure gender identity using a single qualitative item or a categorical list. Response options for the sex assigned at birth question should be determined by the birth certificate options in the place of surveying. For example, in New Zealand a person's sex can be recorded as 'indeterminate' or 'not listed' if it cannot be determined at the time of birth (McDonald, 2015), while this option is not available on birth certificates issued in the US. As such, researchers in New Zealand should add 'indeterminate' and 'not listed' as response options to questions about sex assigned at birth.

Advantages

The primary advantage of the two-step approach is that it has high sensitivity when compared to single-item measures. Although many transgender people may not self-identify as such, respondents can be classified as transgender using this approach if their assigned sex differs from their gender identity (i.e. discordant responses). In a series of studies with 990 US adults, Tate and colleagues (2013) found that the percentage of transgender individuals identified in the sample was doubled when the two-step method was used to measure gender rather than a single item with four response options (female, male, transgender, other). As such, this is arguably the most appropriate measure of gender identity for use in population-based studies (which could be used to estimate the size of the transgender population). This measure is also suitable for researchers specifically interested in gender, or who wish to identify transgender participants.

Disadvantages

Although the two-step method appears promising for the collection of gender identity data in quantitative research, researchers must be aware that asking about an individual's assigned sex at birth may be uncomfortable for some transgender participants (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014). Because of this, there may be higher rates of missing data for this question when compared to single-item measures. Indeed, in a survey of experiences with answering gender identity questions in clinical settings, 7.9% of transgender respondents disagreed that they would answer such a question on a registration form at a health centre (Cahill et al., 2014). Another drawback of the two-step approach is the potential to incorrectly identify intersex people as transgender. An intersex person could, for example, be assigned male at birth and identify as a woman, but not identify as transgender. For further discussion of measures identifying intersex participants, see the section 'Future directions'.

Identifying transgender participants using a single item of transgender status

A final option for identifying transgender participants is the use of a single item that assesses transgender status:

Some people describe themselves as transgender when their gender identity is different from their sex that was documented at birth. For example, a person whose sex was documented as male on their original birth certificate, but who identifies as a woman. Do you consider yourself to be transgender?

- Yes
- No
- I am unsure if I am transgender
- I do not know what this question is asking

Research considerations

This approach has been used in the Massachusetts Behavioral Risk Factor Surveillance System (MA-BRFSS; Conron, Scott, Stowell, & Landers, 2012), the Survey of LGBT Americans (Taylor, 2013) and the Youth'12 study in New Zealand (Clark et al., 2014) to identify transgender participants. As this item does not measure gender identity, an additional item is required. As with the two-step approach, this can be either an open-ended item or a categorical list.

Advantages

Conron, Landers, Reisner, and Sell (2014) found that this single-item measure discriminated well between transgender and cisgender youths in the Massachusetts Gender Measures Project (sensitivity = 86%; specificity = 100%), suggesting that this option would be suitable for research settings in which it is important to identify transgender participants. Another key advantage of this option is that, unlike the two-step approach, it is unlikely to incorrectly identify intersex participants as transgender.

Disadvantages

Unlike the two-step method, there is little empirical research available to support the effectiveness of this approach. More research is needed, for example, to clarify whether this single-item measure correctly classifies a higher proportion of transgender respondents than the two-step approach.

Potential barriers to implementing inclusive measures

Small sample sizes

Perhaps the most commonly cited barrier to implementing inclusive gender identity measures is that the number of transgender participants in many studies is too small to allow for meaningful statistical analysis (Frohard-Dourlent et al., 2016; Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014; Westbrook & Saperstein, 2015). Although the size of the transgender population is exceedingly difficult to estimate, the few population-based studies available indicate that between 0.5% and 1% of adults are transgender (Conron et al., 2012; Olyslager & Conway, 2007). Consequently, many studies that include transgender participants lack statistical power, particularly when attempting to compare transgender and cisgender populations. Given the limited analytic potential of the data, some researchers argue that it is not justifiable to include questions that identify transgender respondents (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014).

Though small sample sizes are, indeed, a challenging aspect of research with minority groups, this does not warrant the use of dichotomous gender measures. High-quality data is routinely gathered on populations comparable in size to the transgender population; for example, the 2013 New Zealand Census form listed 'Niuean' as one of eight possible responses to a question about ethnicity, with results showing that 0.5% of people in New Zealand report Niuean as their ethnic group (Statistics NZ, 2013). Although the small number of Niuean people may prevent analyses of

differences between this and other ethnic groups, collecting this data is nonetheless viewed as important. Collecting data on small groups not only ensures that their (often marginalised) members are heard and respected, but is good methodological practice; every participant should be able to respond to every survey item, regardless of how small their group is. Conceptualising the collection of gender identity in a similar manner may help in the shift towards inclusive measures. Like ethnicity data, gender identity data could be dummy coded for use in regression analyses.

Frohard-Dourlent and colleagues (2016) emphasise that researchers are accountable to their participants. As such, researchers should not collect data from participants of all genders only to exclude transgender participants for ease of analysis. If data from transgender participants is too limited to allow statistical analyses, descriptive statistics can, and should, be reported in a table (e.g. Fraser et al., 2017, Table 1). This data could be used for meta-analyses and would allow for the possibility of conducting synthesis analyses using combined data sets (Perrino et al., 2015). Data from transgender participants in longitudinal studies could be aggregated across time to yield larger sample sizes, increasing statistical power (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014). Researchers should also take care not to use gender as a covariate (which often results in transgender participants being excluded from analysis) without theoretical rationale. If gender must be included researchers have the option of running analyses with and without gender as a covariate, then reporting whether this changed the results.

Lack of relevance to research

Some researchers may contend that their use of a single dichotomous measure of gender is appropriate on the basis that issues of gender are not relevant to their study. According to this perspective, collecting data about transgender people may be viewed as the responsibility of those producing research that will directly *benefit* transgender people; if research cannot be used to estimate the size of the transgender population, or to gain an understanding of their needs and well-being, implementing inclusive measures of gender identity may be viewed as unnecessary.

Regardless of whether gender is considered pertinent to a study, researchers have an ethical obligation to ensure that transgender people are included in their research. The ethical codes that guide social researchers typically contain reference to principles of non-discrimination, inclusion or respect for diversity and autonomy (e.g. Social Research Association, 2003). Researchers must also consider the impact of gender identity measures on participant response and retention – if transgender participants become frustrated by dichotomous measures of gender identity, this increases the likelihood that they will refuse to answer the gender identity question, or will withdraw from the research altogether. Indeed, in a qualitative study of inclusivity in sexual health research, one participant (a transgender man) commented that he was 'yet to encounter any survey on any topic at all that actually reflects [his] life' (Carrotte et al., 2016, p. 6).

Concerns that inclusive measures will confuse or offend cisgender participants

Although transgender participants may become frustrated by dichotomous measures of gender identity, the converse is also possible: that cisgender participants will be confused or offended by measures of gender identity that move beyond the male/female binary. Indeed, the introduction of Statistics New Zealand's 'gender diverse' category sparked a backlash from a conservative Christian lobby group, who claimed that Statistics New Zealand should retain male and female categories only (Radio New Zealand, 2015).

Empirical research using inclusive gender identity measures, however, provides evidence that the vast majority of cisgender respondents do not object to answering inclusive gender identity questions. Glen and Hurrell (2012) reported a dropout rate of just 0.2% across their suite of

questions about sex and gender, in an online sample of over 10,000 UK adults. Analyses of data from the MA-BRFSS (N = 28,540) showed that the non-response rate for the item assessing transgender status was 1.4%, which was lower than the non-response rate on measures assessing income in the same survey (Conron et al., 2012). Cognitive testing with participants in the Growing Up Today Study (a prospective cohort study of US young adults, N = 7,831), which used the two-step approach, identified no major problems in item comprehension or response among cisgender participants (Reisner, Conron, et al., 2014). The two-step approach has also been successfully used to assess gender identity in Latin America/the Caribbean, Portugal and Spain (Reisner, Biello, et al., 2014). Lastly, Broussard, Warner, and Pope (2017) assessed the attitudes of cisgender and gender diverse people in the US, finding that most participants preferred inclusive gender identity measures over a traditional dichotomous measure. They noted that, 'Among cisgender heterosexual participants, the binary format was also never ranked significantly higher than the expanded format' (p. 621).

Future directions

Future research should explore best practices for identifying intersex participants in quantitative research. While some researchers include 'intersex' as a response option in categorical list gender identity measures (e.g. Grant et al., 2011), many intersex people do not identify with the term 'intersex' as a gender identity (Gender Identity in U.S. Surveillance [GenIUSS] Group, 2014). As such, this method may not capture all intersex participants. Other researchers include 'intersex' as a response option in assigned sex at birth measures (e.g. Tate et al., 2013). This is also problematic as intersex infants are usually assigned male or female at birth (Bauer et al., 2017). The GenIUSS Group (2014) recommends asking about intersex identity or intersex status in a separate question either by defining what it means to be intersex and asking if this applies to participants or asking whether participants have been diagnosed by a medical doctor with an intersex condition. Unfortunately, there is a dearth of research examining the advantages and disadvantages of different measures of intersex identity. Future research could also reflect on the ways in which age, ethnicity, socio-economic status or other cultural factors may influence the understanding of gender identity measures.[6]

Concluding remarks

In this article I have outlined four inclusive measures of gender identity and have considered potential barriers to implementing these measures. I contend that the continued use of a single dichotomous measure of gender identity reflects poor methodological practice and may have negative implications for the social inclusion and wellbeing of transgender people (Kapusta, 2016). I wish to reiterate that there is, to date, no 'gold standard' measure of gender identity for use in quantitative research. Rather, the use of each measure of gender identity comes with its own advantages and disadvantages. As many feminist scholars contend, it is likely that quantitative research is inherently ill-equipped to truly capture the nuances and complexities of gender (for discussion, see Jayaratne & Stewart, 1991). This does not mean, however, that quantitative researchers should therefore abandon attempts to measure gender identity – as Westbrook and Saperstein (2015) point out, 'The resulting research is vital for expanding our understanding of the gender system, working to reduce gender inequalities, and providing entry into policy debates' (p. 537–538).

When deciding which inclusive measure of gender identity to use, researchers must consider their research questions, study design and participant demographics. Researchers collecting representative data arguably have a responsibility to use measures able to identify transgender participants as they have the ability to provide a much-needed estimate of the size of the transgender population. Those specifically interested in gender will, no doubt, also wish to use measures that

identify transgender participants; data produced from these measures may give researchers the ability to examine disparities between transgender and cisgender participants, or to explore the diversity of experiences within the transgender community. Finally, researchers collecting non-representative data without a specific interest in gender should use either a single qualitative item or a categorical list to measure gender identity. While gender may not be the focus of analysis, the use of an inclusive measure ensures all participants are able to report their gender and communicates that there are more ways of being than just 'male' and 'female'.

In sum, I encourage researchers in the social sciences to consider the measurement of gender identity in their own research in an effort not only to improve methodological practice but to combat the invisibility and erasure that comes with the continued use of dichotomous gender identity measures.

Notes

1. I draw on definitions of sex and gender provided by the Gender Identity in U.S. Surveillance [GenIUSS] Group (2014), who note that 'the term sex refers to biological differences among male, female, and intersex people' (p. x). Gender is described as 'a multidimensional construct that has psychological, social, and behavioural dimensions that include gender identity and gender expression' (p. ix).
2. Other non-binary people may select both options and be treated as missing data by researchers.
3. If gender was measured inclusively across all suicide and mental health studies, this would provide valuable insight as to whether empirically supported mental health interventions and suicide prevention strategies are effective in transgender communities(Haas et al., 2010).
4. Given that categorical lists differ between studies, the use of this approach may impair collaborative research – data sets using varying response options may be difficult to compare. Researchers using categorical lists should draw on previous literature to inform their response options in order to maintain consistency between studies.
5. Johnson and colleagues (2004) define othering as 'a process that identifies those that are thought to be different from oneself or the mainstream, and … can reinforce and reproduce positions of domination and subordination' (p.5253).
6. See the Gender Identity in U.S. Surveillance [GenIUSS] Group (2014) report for recommendations about data collection with a range of populations.

Notes on contributor

Gloria Fraser is a PhD and clinical psychology student at Victoria University of Wellington. She is interested in the methodology of gender measurement, as well as in the intersection of sex-sexuality-gender diversity and clinical psychology. Her doctoral research focusses on queer and transgender experiences of accessing mental health support in New Zealand.

Acknowledgments

Gloria Fraser was supported by a Victoria University of Wellington Summer Research Scholarship during the preparation of this article. I am grateful to Joseph Bulbulia, Kealagh Robinson and Jennifer Katherine Shields for their helpful comments on earlier drafts of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gloria Fraser http://orcid.org/0000-0002-7696-7280

References

Ansara, Y. G., & Hegarty, P. (2014). Methodologies of misgendering: Recommendations for reducing cisgenderism in psychological research. *Feminism & Psychology*, *24*(2), 259–270.

Arcelus, J., Bouman, W. P., Van Den Noortgate, W., Claes, L., Witcomb, G., & Fernandez-Aranda, F. (2015). Systematic review and meta-analysis of prevalence studies in transsexualism. *European Psychiatry*, *30*(6), 807–815.

Barker, M. J. (2016). Non-binary genders. In A. E. Goldberg (Ed.), *The SAGE Encyclopedia of LGBTQ Studies, Volume 2* (pp. 816–820). London: Sage.

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), e0178043.

Bauer, G. R., Hammond, R., Travers, R., Kaay, M., Hohenadel, K. M., & Boyce, M. (2009). "I don't think this is theoretical; this is our lives": How erasure impacts health care for transgender people. *Journal of the Association of Nurses in AIDS Care*, *20*(5), 348–361.

Bauer, G. R., Pyne, J., Francino, M. C., & Hammond, R. (2013). Suicidality among trans people in Ontario: Implications for social work and social justice. *Service Social*, *59*, 35–62.

Bockting, W., Robinson, B., Benner, A., & Scheltema, K. (2004). Patient satisfaction with transgender health services. *Journal of Sex & Marital Therapy*, *30*(4), 277–294.

Broussard, K. A., Warner, R. H., & Pope, A. R. (2017). Too many boxes, or not enough? Preferences for how we ask about gender in cisgender, LGB, and gender-diverse samples. *Sex Roles*, 1–19. doi:10.1007/s11199-017-0823-2

Budge, S. L., Adelson, J. L., & Howard, K. A. (2013). Anxiety and depression in transgender individuals: The roles of transition status, loss, social support, and coping. *Journal of Consulting and Clinical Psychology*, *81*(3), 545–557.

Cahill, S., & Makadon, H. J. (2014). Sexual orientation and gender identity data collection update: US government takes steps to promote sexual orientation and gender identity data collection through meaningful use guidelines. *LGBT Health*, *1*(3), 157–160.

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), e107104.

Carrotte, E. R., Vella, A. M., Bowring, A. L., Douglass, C., Hellard, M. E., & Lim, M. S. (2016). "I am yet to encounter any survey that actually reflects my life": A qualitative study of inclusivity in sexual health research. *BMC Medical Research Methodology*, *16*, 86.

Chang, T. K., & Chung, Y. B. (2015). Transgender microaggressions: Complexity of the heterogeneity of transgender identities. *Journal of LGBT Issues in Counseling*, *9*(3), 217–234.

Clark, T. C., Lucassen, M. F., Bullen, P., Denny, S. J., Fleming, T. M., Robinson, E. M., & Rossen, F. V. (2014). The health and well-being of transgender high school students: results from the new zealand adolescent health survey (Youth'12). *Journal of Adolescent Health*, *55*(1), 93-99. doi:10.1016/j.jadohealth.2013.11.008

Conron, K. J., Landers, S. J., Reisner, S. L., & Sell, R. L. (2014). Sex and gender in the US health surveillance system: A call to action. *American Journal of Public Health*, *104*(6), 970–976.

Conron, K. J., Scott, G., Stowell, G. S., & Landers, S. J. (2012). Transgender health in Massachusetts: Results from a household probability sample of adults. *American Journal of Public Health*, *102*(1), 118–122.

Dean, L., Meyer, I. H., Robinson, K., Sell, R. L., Sember, R., Silenzio, V. M., … Dunn, P. (2000). Lesbian, gay, bisexual, and transgender health: Findings and concerns. *Journal of the Gay and Lesbian Medical Association*, *4*(3), 102–151.

Deutsch, M. B. (2016). Making it count: Improving estimates of the size of transgender and gender nonconforming populations. *LGBT Health*, *3*(3), 181–185.

Dick, P. (2004). Resistance to diversity initiatives. In J. Helms Mills, A. J. Mills, & R. Thomas (Eds.), *Identity politics at work: Resisting gender and gendering resistance* (pp. 67–84). London: Routledge.

Fraser, G., Wilson, M. S., Garisch, J. A., Robinson, K., Brocklesby, M., Kingi, T., … Russell, L. (2017). Non-suicidal self-injury, sexuality concerns, and emotion regulation among sexually diverse adolescents: A multiple mediation analysis. *Archives of Suicide Research*, *22*(3), 432–452.

Fraser, G. E. (2017). *Statistical standard for coding gender identity* (NZAVS Technical Documents, e27). Retrieved from https://cdn.auckland.ac.nz/assets/psych/about/our-research/nzavs/NZAVSTechnicalDocuments/NZAVS-Statistical-Standard-for-Gender-Identity.xlsx

Frohard-Dourlent, H., Dobson, S., Clark, B. A., Doull, M., & Saewyc, E. M. (2016). "I would have preferred more options": Accounting for non-binary youth in health research. *Nursing Inquiry*, 1–9. doi:10.1111/nin.12150

Gender Identity in U.S. Surveillance [GenIUSS] Group. (2014). *Best practices for asking questions to identify transgender and other gender minority respondents on population-based surveys*. Los Angeles, CA: The Williams Institute.

Glen, F., & Hurrell, K. (2012). *Technical note: Measuring gender identity*. Manchester: Equality and Human Rights Commission.

Grant, J.M, Mottet, L.A, Tanis, J, Harrison, J, Herman, J.L, & Keisling, M. (2011). *Injustice at every turn*. Washington: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Greaves, L. M., Barlow, F. K., Huang, Y., Stronge, S., Fraser, G., & Sibley, C. G. (2017). Asexual identity in a New Zealand national sample: Demographics, well-being, and health. *Archives of Sexual Behavior*, *46*(8), 2417–2427.

Greaves, L. M., Barlow, F. K., Lee, C. H., Matika, C. M., Wang, W., Lindsay, C. J., … Stronge, S. (2016). The diversity and prevalence of sexual orientation self-labels in a New Zealand national sample. *Archives of Sexual Behavior*, *46*(5), 1325–1663.

Greaves, L. M., Sibley, C. G., Fraser, G., & Barlow, F. K. (2018). *Comparing pansexual and bisexual identified participants on demographics, psychological well-being, and politics in a New Zealand national sample*. Manuscript submitted for publication.

Haas, A. P., Eliason, M., Mays, V. M., Mathy, R. M., Cochran, S. D., D'Augelli, A. R., … Russell, S. T. (2010). Suicide and suicide risk in lesbian, gay, bisexual, and transgender populations: Review and recommendations. *Journal of Homosexuality*, *58*(1), 10–51.

Harrison-Quintana, J., Grant, J. M., & Rivera, I. G. (2015). Boxes of our own creation: A trans data collection wo/manifesto. *TSQ: Transgender Studies Quarterly*, *2*(1), 166–174.

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). *The Report of the 2015 U.S. Transgender Survey*. Washington, DC: National Center for Transgender Equality.

Jaroszewski, S., Lottridge, D., Haimson, O. L., & Quehl, K. (2018, April). *Genderfluid or attack helicopter: Responsible HCI research practice with non-binary gender variation in online communities*. Paper presented at the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, Canada. Retrieved from http://oliverhaimson.com/PDFs/JaroszewskiGenderfluidOrAttack.pdf

Jayaratne, T. E., & Stewart, A. J. (1991). Quantitative and qualitative methods in the social sciences. In M. M. Fonow & J. A. Cook (Eds.), *Beyond methodology: Feminist scholarship as lived research* (pp. 85–106). Bloomington: Indiana University Press.

Johnson, J. L., Bottorff, J. L., Browne, A. J., Grewal, S., Hilton, B. A., & Clarke, H. (2004). Othering and being othered in the context of health care services. *Health Communication*, *16*(2), 255–271.

Kapusta, S. J. (2016). Misgendering and its moral contestability. *Hypatia*, *31*, 502–519.

KnowYourMeme. (2015). *I sexually identify as an attack helicopter*. Retrieved from http://knowyourmeme.com/memes/i-sexually-identify-as-an-attack-helicopter

Lombardi, E. L., Wilchins, R. A., Priesing, D., & Malouf, D. (2002). Gender violence: Transgender experiences with violence and discrimination. *Journal of Homosexuality*, *42*(1), 89–101.

MacCarthy, S., Reisner, S. L., Nunn, A., Perez-Brumer, A., & Operario, D. (2015). The time is now: Attention increases to transgender health in the United States but scientific knowledge gaps remain. *LGBT Health*, *2*(4), 287–291.

Magliozzi, D., Saperstein, A., & Westbrook, L. (2016). Scaling up: Representing gender diversity in survey research. *Socius: Sociological Research for a Dynamic World*, *2*, 1–11.

McDonald, E. (2015). Intersex people in Aotearoa New Zealand: The challenges for law and social policy: Part I: Critiquing gender normalising surgery. *Victoria University of Wellington Law Revue*, *46*, 705–724.

Olyslager, F., & Conway, L. (2007, September). *On the calculation of the prevalence of transsexualism*. Paper presented at the World Professional Association for Transgender Health 20th International Symposium, Chicago, IL. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.692.8704&rep=rep1&type=pdf

Perrino, T., Beardslee, W., Bernal, G., Brincks, A., Cruden, G., Howe, G., … Brown, C. H. (2015). Toward scientific equity for the prevention of depression and depressive symptoms in vulnerable youth. *Prevention Science*, *16*(5), 642–651.

Radio New Zealand. (2015). *New gender category 'problematic'*. Retrieved from http://www.radionz.co.nz/news/national/279052/new-gender-category-'problematic'

Rankin, S. 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.

Reisner, S. L., Biello, K., Rosenberger, J. G., Austin, S. B., Haneuse, S., Perez-Brumer, A., … 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.

Reisner, S. L., Conron, K. J., Tardiff, L. A., Jarvi, S., Gordon, A. R., & Austin, S. B. (2014). Monitoring the health of transgender and other gender minority populations: Validity of natal sex and gender identity survey items in a US national cohort of young adults. *BMC Public Health*, *14*, 1224–1234.

Reisner, S. L., Deutsch, M. B., Bhasin, S., Bockting, W., Brown, G. R., Feldman, J., … Tangpricha, V. (2016). Advancing methods for US transgender health research. *Current Opinion in Endocrinology, Diabetes and Obesity*, *23*(2), 198–207.

Reisner, S. L., White, J. M., Bradford, J. B., & Mimiaga, M. J. (2014). Transgender health disparities: Comparing full cohort and nested matched-pair study designs in a community health center. *LGBT Health*, *1*(3), 177–1784.

Robinson-Cimpian, J. P. (2014). Inaccurate estimation of disparities due to mischievous responders: Several suggestions to assess conclusions. *Educational Researcher*, *43*(4), 171–185.

Serano, J. (2007). *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity*. Emeryville, CA: Seal.

Social Research Association. (2003). *Ethical guidelines*. Retrieved from http://the-sra.org.uk/wp-content/uploads/ethics03.pdf

Statistics New Zealand. (2013). *Ethnic group (detailed total responses), 2001, 2006, 2013 Censuses* (Excel File, 2013 Census QuickStats about Culture and Identity – Tables). Retrieved from http://www.stats.govt.nz/Census/2013-census/profile-and-summary-reports/quickstats-culture-identity.aspx

Stryker, S. (2006). *The transgender studies reader*. New York: Routledge.

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.

Taylor, P. (2013). *A survey of LGBT Americans: Attitudes, experiences and values in changing times*. Retrieved from: http://www.pewsocialtrends.org/files/2013/06/SDT_LGBT-Americans_06-2013.pdf

Van Anders, S. M. (2015). Beyond sexual orientation: Integrating gender/sex and diverse sexualities via sexual configurations theory. *Archives of Sexual Behavior*, *44*(5), 1177–1213.

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

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society*, *1*(2), 125–151.

Westbrook, L., & Saperstein, A. (2015). New categories are not enough: Rethinking the measurement of sex and gender in social surveys. *Gender & Society*, *29*(4), 534–560.

Winter, S. (2017, February). *The Australasian trans experience: First impressions fresh off the boat*. Paper presented at the Society for the Scientific Study of Sexuality Symposium, Sydney, Australia.