Moving Beyond Sex: Measuring Gender Identity in Telephone Surveys

Dan Cassino

1 Dept. of Social Sciences and History, Fairleigh Dickinson University

Keywords: political surveys, measurement, gender identity

https://doi.org/10.29115/SP-2020-0009

Survey Practice
Vol. 13, Issue 1, 2020

While researchers have long understood that sex and gender are theoretically distinct concepts, much survey research continues to rely on the sex of respondents as a proxy for their gender. Recent work on political and social attitudes has shown that gender has far-reaching consequences for attitudes that are not neatly summed up by measures of sex, but measures of gender identity have generally been used either in forms that are too lengthy for most survey research or have only been used in online modalities. Longer sets of gender identity items may be theoretically desirable, but a gender identity item embedded in a live-caller random digit dialing survey in New Jersey shows that just one gender identity item can account for a great deal of within-sex variance in political and social attitudes without leading to an overly high rate of refusals.

Despite the consensus that sex and gender are theoretically separable, most survey research continues to treat them as interchangeable, even measuring gender on telephone surveys by asking the interviewer whether the respondent sounds like a man or a woman (Bittner and Goodyear-Grant 2017a). Even discussions of the “gender gap” in American politics center on the difference between men and women (e.g., Norrander 1999)—a chasm that might better be described as the “sex gap.” No one is arguing that self-reported, or interviewer identified, sex is a bad proxy for gender, as sex and gender identity seem to line up in most cases. But the conflation of sex and gender can lead to errors (Lovenduski 1998), and if we are interested in gender, or the attitudes linked to it, perhaps we should just ask respondents about their gender identity directly.

While the theoretical case for measuring gender is strong, studies making use of gender identity scales have been carried out mostly in online samples, and the best population-based samples have been taken outside of the United States (US). In addition, most of the surveys incorporating gender identity scales have focused on general political views rather than measurement of specific attitudes, such as presidential approval. However, while they have not been widely used in political surveys, gender identity scales can be easily incorporated into telephone polls, and doing so offers a substantial improvement over using sex as a proxy for gender, while better matching the
theoretical consensus that sex and gender are different things. Moreover, they can be included in simplified forms that capture variance without intimidating or causing drop-off among respondents.

**Gender Gap or Sex Gap?**

There is little question about the long-term existence of the gap between the attitudes of men and women in the US (e.g., Box-Steffensmeier, De Boef, and Lin 2004, Gidengil et al. 2003; Kaufmann 2017; Lizotte and Sidman 2009), including significant gaps in political partisanship (e.g., Norrander 1999). Voters even apply gender to issues such as foreign policy, which have no obvious or necessary gender component (Deckman and Cassese 2019). But while there are clearly sex-based differences in political attitudes and behaviors, few scholars think that these differences are driven by biology. As Schneider and Bos (2019) put it: “Sex differences... require gendered explanations.”

In what seems to be the most comprehensive approach, McDermott (2016) measures gender identity through the use of the Bem Sex-Role Inventory (BSRI, Bem 1981), which measures agreement with gendered traits, embedded in a 2011 general political survey with 780 respondents from an online sample. She finds that masculine or feminine traits, much more than sex, divide the American public. People who score higher on masculine traits are more politically engaged and are more likely to affiliate with and vote for Republicans. The gaps between men and women are dwarfed by the gaps between individuals with high levels of masculinity and femininity: while McDermott finds a nine-point sex gap in party identification, she finds a 24-point difference between individuals with high and low levels of femininity, and an 18-point difference between those with high and low levels of masculinity. While the BSRI is well-validated, it requires more space on a survey than most researchers have available, with even a reduced version of the scale weighing in at 30 items.

The unwieldiness of such measures has led researchers to look for alternative ways of measuring gender identity that are more amenable to survey research. Bittner and Goodyear-Grant (2017a, 2017b), for instance, added self-placement scales of gender identity to a series of web-based surveys conducted during Canadian provincial elections in 2011 and 2012, with a total sample size of 6,073. Modeled on feeling thermometers, these scales ask participants to rate themselves on a 101-point unidimensional scale where one pole represents complete femininity, and the other complete masculinity.

The use of such scales is not entirely in line with best practice suggested by researchers, such as Westbrook and Saperstein (2015), who call for multidimensional scaling (as in Bem 1974) and multistage items asking about current sex and sex at birth, but arguably bring us closer to an ideal measurement with a smaller burden on survey time.
Bittner and Goodyear-Grant’s (2017a) scale gives somewhat different results from McDermott (2016), but like McDermott, they show significant variance in gender identity within sex categories. In contrast to the trait-based measurement of the BSRI, the self-report scales show a very strong negative correlation between masculinity and femininity, obviating the need for two separate scales in practice, even if they may be more desirable theoretically.

Lindqvist et al. (2020) review other options for creating better measures of gender identity and sex, such as the use of various “third” options on sex items, open-ended responses, and alternate gender identity scales such as the Multi-Gender Identity Questionnaire (Joel et al. 2014). They conclude that the appropriate measure of gender identity on a given survey depends on the research questions being addressed, and for political outcome variables, gender identity measures seem to fit the bill.

**Masculinity and Politics**

Especially in the wake of the 2016 US presidential election, there has been a surge in research on the impact of gender identity, especially among men, on political attitudes and behaviors. Some accounts, such as Katz 2016, argue that Donald Trump’s highly gendered campaign rhetoric led to increased support among marginalized working-class men, a constituency critical to his Electoral College victory. Cassino (2018) finds that an induced gender identity threat in the run up to the 2016 election led men to be more likely to support Trump, but only in a match-up with a female candidate. Valentino et al. (2018) notably found an important role for gender attitudes and sexism in 2016 vote choice, with the effects of sexism being magnified among angry voters, and anger being particularly potent in mobilizing voters with high levels of sexism. Cassese and Holman (2017) argue that masculinity, in conjunction with gendered religious views, drives the ideological gap between men and women. Religious men who hold masculine views of God and other authority figures hold much more conservative policy views and have more conservative political identities than similarly religious women. In essence, discussions of the effect of masculinity are arguing for an interactive effect between gender identity and sex among men—but the only way to test such claims is to measure gender identity in the first place.

**Measuring Gender Identity in a Live-Caller Telephone Interview**

Adding simple gender identity items to surveys allows for two important measures. First, it allows for the separate measurement of gender gaps and sex gaps and shows the areas in which they diverge. Second, it allows us to identify men with strongly masculine gender identities, a group that is expected to be different from both other men and from women who hold masculine gender identities.
The US data comes from a live caller random digit dialing (RDD) sample of New Jersey adults taken in October 2019. In addition to standard items about local and national politics, respondents were asked to describe their gender identity on two scales: a first-person gender identity scale (“How do you see yourself?”) and a third-person gender scale (“and how do other people see you?”). Both scales were unidimensional, with 6 points, ranging from “ Completely Masculine” to “ Mostly Masculine” to “ Slightly Masculine,” then on to “Completely Feminine,” with no neutral center point.

The exact wording of the item was, “The traits that we see as being masculine or feminine are largely determined by society, and have changed dramatically over time. As a result, everyone has some combination of masculine and feminine traits, which may or may not correspond with whether they are male or female. How do you see yourself? Would you say that you see yourself as Completely Masculine, Mostly Masculine, Slightly Masculine, Slightly Feminine, Mostly Feminine, or Completely Feminine?” Exact wording for other items can be found in the appendix.

While the data presented here made use of a 6-point scale, rather than Bittner and Goodyear-Grant’s 101-point scale, this is not because of any inadequacy in the 101-point scale. It is simply because survey researchers may be more comfortable with the use of smaller scales, rather than feeling thermometer type scales, and even these simpler scales can capture most of the variance found in the more complex scales. Similarly, there is no strong theoretical reason to prefer an even-numbered scale with no neutral center point to an odd-numbered scale with one, and researchers may well prefer to allow respondents to identify at a neutral point, especially in samples that may include higher proportions of non-binary respondents.

A total of 38% of respondents placed themselves as the extreme ends of the gender identity spectrum (18% on the masculine end and 20% on the feminine end). Results were almost identical for the third-person scale, so in line with previous researchers using such questions, we focus on the first-person scales because they are tightly correlated with the third-person scales asked immediately afterward (r=0.91) to make separate analyses redundant.

These gender identity placements mostly, but not entirely, match up with the respondent’s sex. More than 90% of men (identified by interviewers) and more than 80% of women placed themselves on the side of the gender identity scales that matched their sex, but there was tremendous variance within sex groups. Only 36% of men said that they were entirely masculine, with the largest group (50%) saying that they were “mostly” masculine, and 5% saying that they were at least “slightly” feminine. A total of 35% of women placed themselves as completely feminine and 42% as “mostly” feminine, but 17% said that they saw themselves as masculine. The assumption that men are entirely masculine, and women are entirely feminine is correct for about 35% of respondents but would overstate the connection between gender and sex for most respondents.
Figure 1: Gender identity by sex

and completely mischaracterize about 11% of the sample. To the extent that we are interested in gender, the sex proxy is doing a good, but not great, job. It works pretty well for most respondents, while badly mischaracterizing a minority.

These results match up well, but not perfectly, with Bittner and Goodyear-Grant’s (2017a, 2017b) 101-point scale. They found almost the same proportion of both men (39%) and women (36%) identifying with the extreme gender category congruent with their sex, and found about the same proportion of respondents in the gender categories conventionally associated with the opposite sex (8% in their sample, 11% here, mostly women in both). Given the differences in mode (online versus live caller) and population (Canada versus the US), these similarities are striking and support the idea that the 6-point scale is generally capturing the same construct as the larger scale.

About 5% of the sample (45 respondents out of 801) declined to answer the first-person gender identity item, with slightly more respondents (58) declining to answer the third-person item. Most of those missing responses (28/45) were from respondents who said that they did not know the answer; only 17 (2%) refused to answer the question. In chi-square testing, missing data on the gender identity scale was significantly related to Republican affiliation, higher income and whiteness, but not age or sex. On the whole, the refusal rate is similar to the refusal rate for the income item (55 respondents, or 7%). Medeiros et al. (2020) similarly find no evidence of negative reactions to non-binary categorical sex items on surveys.
Gender Identity and Political Attitudes

The fact that these gender identity measures are measuring something distinguishable from sex and are not causing a decline in data quality, does not necessarily mean that they are useful to researchers studying public opinion. Space on surveys is limited, so any item included should work to explain attitudes in a way that existing items, such as sex, do not. For instance, on a 5-point party identification scale (1 is Republican, 2 is Lean Republican, 3 is Independent/no lean, and so on), males, overall, have a mean score of 2.8, with females at 2.4 (mean and margin of error calculations for all questions and groups are presented in Table 1). However, there is a significant difference between men who identify as “completely” masculine (mean of 3.3) and those who do not (mean of 2.5), and the gap between those men who identify as “completely” masculine (mean of 3.3) and women who are “completely” feminine (mean of 2.7), at 0.6 points, is substantially larger than the gap between male and female respondents. The biggest observed gap is between respondents, male and female, who “completely” identify with their sex (mean of 3.0), and the (much smaller) group that identifies with the gender identity of the other sex (mean of 2.0). Even on a trait as normally stable as partisanship, the use of sex as a proxy for gender hides enormous within-sex variation by gender identity.

While previous work, such as Bittner and Goodyear-Grant (2017a) show that gender identity predicts general political attitudes, researchers may well be interested in more narrow questions. The results presented here show that gender identity scales, and their interaction with sex do just that.

For instance, 44% of men approve of the job Trump is doing as president, compared with 32% of women. But this hides the extent of the gender gap: 54% of men who identify as completely masculine approve of him (54%), significantly more than those who do not identify as completely masculine (36%), with a similar difference between completely feminine women (39%) and other women (28% approval). Again, the gap between men and women (12 points) is dwarfed by the gap between those whose gender identity matches their sex (47%) and those that do not (22% approval).

These types of differences are present on all of the issues studied, ranging from scores on a hostile sexism scale, an abortion rights scale, support for an assault weapon ban, gubernatorial approval, and support for the impeachment of President Trump.

A pattern emerges across all of these issues. In general, individuals who strongly identify with their gender identity corresponding with their sex hold more conservative views than those who do so “mostly,” “slightly,” or not at all, replicating a finding from Bittner and Goodyear-Grant (2017a). This effect is exacerbated among men, with “completely masculine” identifying men...
holding more conservative views than other men. This seems to support the idea that masculinity is an important element in political views, with gender identity impacting men’s political behaviors differently than women’s.

These results, however, raise the question of causality. Is gender identity pushing political identity, or is political identity pushing gender identity? The answer, according to experiments previously carried out by Cassino and Besen-Cassino (2020) is yes. Increasing the salience of gender identity changes the expression of partisanship, and increasing the salience of partisanship changes expressed gender identity on similar scales. So, while it may be the case that more conservative men express more traditional gender identities, it is also the case that these identities lead men to hold more conservative political views.

**Recommendations**

On nearly every issue studied and party identification, gender identity accounts for variance in political attitudes not explained by sex. This is especially true among men, as those who identify as strongly masculine are different in their attitudes from almost every other group. Even if we were to ignore that and continue to use sex as a proxy, adherence to the gender identity commonly associated with the respondent’s sex also has a substantial effect on political attitudes and party identification which cannot be accounted for without the use of a gender identity item.

None of this is to say that sex is a bad proxy for gender identity. On average, men are masculine and women are feminine. But, as shown in the issue areas covered here, it works better on some issues than others, and leaving out gender identity means that researchers are likely to ascribe to party identification some variance that is better explained by gender, especially among very masculine men. There is little disagreement that sex and partisanship are important drivers of contemporary U.S. politics, but it is impossible to accurately measure their effects without measuring the correlated effects of gender identity, especially among men. As Lindqvist et al. (2020) point out, we could measure age as “young” or “old,” and capture most of its effects, but no one would seriously countenance doing so.

These results demonstrate that while more onerous measures such as the BSRI and multi-dimensional scales may be ideal from a theoretical perspective, a single first-person gender identity item, even in a shortened form, is sufficient to capture a great deal of gender-based variance. It also has a relatively low refusal/do not know rate, comparable to the income items used almost universally in political surveys. Also, these measures are empirically interesting: the strong correlation between gender identity and partisanship implies that the use of partisanship measures as predictors in the absence of gender identity measures biases the estimated effects of partisanship.
Figure 2a: Selected issue positions by party ID, sex and gender identity

Error bars represent 95% confidence intervals.
Figure 2b: Selected issue positions by party ID, sex and gender identity

Error bars represent 95% confidence intervals.
To put it bluntly, there is no reason not to use a gender identity item on political or social surveys. Doing so not only matches our practice with our theoretical understanding, it accounts for otherwise unexplained variance and may even serve to signal our inclusiveness to respondents from gender identity minority groups, helping us better capture their views.

Contact Information
M-MS3-02, 285 Madison Ave, Madison, NJ 07940
dcassino@fdu.edu
973-443-8799

Appendix. Exact wording of other items mentioned in text

Trump Approval: Do you approve or disapprove [rotate] of the way Donald Trump is handling his job as President?

Impeachment: Based on what you have heard or read about the allegations surrounding the President and his phone call with a foreign president, do you support or oppose his impeachment in the House of Representatives?

Gubernatorial Approval: And turning to New Jersey, do you approve or disapprove [ROTATE] of the job Phil Murphy is doing as governor?

Abortion Scale: Turning to reproductive rights, do you think abortion should be legal in all cases, legal in most cases, illegal in most cases or illegal in all cases?
Sexism Scale: Please indicate the extent to which you agree or disagree with the following statements: (Strongly agree/ Agree/ Disagree/ Strongly disagree).

Most women interpret innocent remarks or actions as being sexist

Most women fail to appreciate fully all that men do for them.

Women seek to gain power by getting control over men.

Women are too easily offended.

Assault Weapons Ban: In order to help prevent gun violence and mass shootings in this country, please tell me if you would support or oppose [rotate] a nationwide ban on so-called assault weapons.

Party Identification: In politics today, do you consider yourself a Democrat, Republican, Independent, or something else? [If Independent] Which way do you lean?
REFERENCES

Bem, Sandra L. 1974. “The Measurement of Psychological Androgyny.” Journal of Consulting and Clinical Psychology 42 (2): 155–62. https://doi.org/10.1037/h0036215.

———. 1981. “Gender Schema Theory: A Cognitive Account of Sex Typing.” Psychological Review 88 (4): 354–64. https://doi.org/10.1037/0033-295x.88.4.354.

Bittner, Amanda, and Elizabeth Goodyear-Grant. 2017a. “Sex Isn’t Gender: Reforming Concepts and Measurements in the Study of Public Opinion.” Political Behavior 39 (4): 1019–41. https://doi.org/10.1007/s11109-017-9391-y.

———. 2017b. “Digging Deeper into the Gender Gap: Gender Salience as a Moderating Factor in Political Attitudes.” Canadian Journal of Political Science 50 (2): 559–78. https://doi.org/10.1017/s0008423917000270.

Box-Steppensmeier, Janet M., Suzanna De Boef, and Tse-Min Lin. 2004. “The Dynamics of the Partisan Gender Gap.” American Political Science Review 98 (3): 515–28. https://doi.org/10.1017/s0003055404001315.

Cassese, Erin, and Mirya R. Holman. 2017. “Religion, Gendered Authority, and Identity in American Politics.” Politics and Religion 10 (1): 31–56. https://doi.org/10.1017/s1755048316000407.

Cassino, Dan. 2018. “Emasculation, Conservatism, and the 2016 Election.” Contexts 17 (1): 48–53. https://doi.org/10.1177/1536504218766551.

Cassino, Dan, and Yasemin Besen-Cassino. 2020. “Political Identity, Gender Identity or Both? The Political Effects of Sexual Orientation and Gender Identity Items in Survey Research.” European Journal of Politics and Gender. https://doi.org/10.1332/251510820x15912551895078.

Deckman, Melissa, and Erin Cassese. 2019. “Gendered Nationalism and the 2016 US Presidential Election: How Party, Class, and Beliefs about Masculinity Shaped Voting Behavior.” Politics & Gender, October, 1–24. https://doi.org/10.1017/s1743923x19000485.

Joel, Daphna, Ricardo Tarrasch, Zohar Berman, Maya Mukamel, and Effi Ziv. 2014. “Queering Gender: Studying Gender Identity in ‘Normative’ Individuals.” Psychology & Sexuality 5 (4): 291–321. https://doi.org/10.1080/19419899.2013.830640.

Kaufmann, K.M. 2017. “The Gender Gap.” In Beyond Red State and Blue State, 104–20. Oxfordshire, UK: Routledge.

Lindqvist, Anna, Marie Gustafsson Sendén, and Emma A. Renström. 2020. “What Is Gender, Anyway: A Review of the Options for Operationalising Gender.” Psychology & Sexuality, February, 1–13. https://doi.org/10.1080/19419899.2020.1729844.

Lizotte, Mary-Kate, and Andrew H. Sidman. 2009. “Explaining the Gender Gap in Political Knowledge.” Politics & Gender 5 (2): 127–51. https://doi.org/10.1017/s1743923x09000130.

Lovenduski, Joni. 1998. “Gendering Research in Political Science.” Annual Review of Political Science 1 (1): 333–56. https://doi.org/10.1146/annurev.polisci.1.1.333.

McDermott, Monika L. 2016. Masculinity, Femininity, and American Political Behavior. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190462802.001.0001.

Medeiros, Mike, Benjamin Forest, and Patrik Öhberg. 2020. “The Case for Non-Binary Gender Questions in Surveys.” PS: Political Science & Politics 53 (1): 128–35. https://doi.org/10.1017/s1049096519001203.
Norrander, Barbara. 1999. “The Evolution of the Gender Gap.” *The Public Opinion Quarterly* 63 (4): 566–76. https://doi.org/10.1086/297871.

Schneider, Monica C., and Angela L. Bos. 2019. “The Application of Social Role Theory to the Study of Gender in Politics.” *Advances in Political Psychology* 40 (1): 173–213. https://doi.org/10.1111/pops.12573.

Valentino, Nicholas A, Carly Wayne, and Marzia Oceno. 2018. “Mobilizing Sexism: The Interaction of Emotion and Gender Attitudes in the 2016 US Presidential Election.” *Public Opinion Quarterly* 82 (S1): 799–821. https://doi.org/10.1093/poq/nfy003.

Westbrook, Laurel, and Aliya Saperstein. 2015. “New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys.” *Gender & Society* 29 (4): 534–60. https://doi.org/10.1177/0891243215584758.