Measuring Gender Identity in Latin America and the Caribbean

Oscar Castorena and Valerie Schweizer-Robinson
Vanderbilt University

September 19, 2023

Key Findings:

- LAPOP Lab developed a revised gender identity question that prioritizes the accurate and inclusive measurement of respondent gender as well as the comfort of respondents and interviewers with the survey.

- Results from cognitive pretesting and online survey experiments in Guatemala informed the adoption of the revised gender instrument.

- For social science public opinion surveys in the LAC context, a single open-ended gender identity question, with options for the interviewer to select in the field, is optimal.

- Sensitivity concerns with a gender identity question can be minimized, but not eliminated, by using language such as “for statistical purposes” and asking people to “confirm” their gender.
Measures of gender are expanding past the gender binary. In recent years, survey projects have developed approaches to the accurate and inclusive recording of people's gender. However, these efforts have emerged predominantly in developed Western contexts. LAPOP Lab is engaged in efforts to incorporate more diverse gender options in surveys in the Latin America and Caribbean (LAC) region. This Methodological Note offers perspective and provides recommendations for measuring gender identity in social science public opinion surveys in the LAC region.

We first describe the general context of asking about gender in LAC, then delve into a review of current best practices and recent literature. We then present results from pretesting and a web survey that LAPOP Lab conducted in Guatemala. Finally, we provide recommendations for the creation of gender identity measurements by public opinion surveys in the region.

Overall, this note recommends a single, open-ended gender identity question where the answer categories are visible to the interviewer but not the respondent. The question-and-answer categories should be adapted to where the survey is taking place, should prioritize both interviewer and interviewee comfort, and should include options outside the gender binary. These recommendations are informed by our review of the state of the field, as well as results from pretesting and an experiment carried out by LAPOP Lab, which are described in this note.

Gender Identity in the Latin America and Caribbean Context

In North America and Europe, many surveys are adopting more inclusive gender identity modules, including open-ended questions, more exhaustive lists of gender options, and including separate questions on sex and gender. Surveys in Latin America have begun including more diverse gender options as well. The cultural context around discussing gender identity is very different across these regions.

In the LAC region, asking about gender is often a sensitive topic, because many assume their gender should be visible based on how they present themselves. As can be seen from LAPOP pretesting results discussed later in this note, asking someone about their gender has the potential to offend and make individuals uncomfortable.

Further, those who do identify outside the gender binary may be cautious to share their identity. Many non-binary individuals in the region experience discrimination and violence. In Latin America, the life expectancy of transgender people may be almost half that of their cisgender counterparts. Even though laws in certain countries allow self-reporting of gender on identification cards and despite increased legislation to protect the rights of sexual and gender minorities, the LAC region has the highest rate of violence against sexual and gender minorities in the world. This context may leave survey respondents hesitant to share their true identity.
When respondents are made uncomfortable by questions, survey quality diminishes as satisficing increases, cooperation decreases, and/or interviewer effects are exacerbated. Given the sensitive nature of asking about gender identity in the LAC region, including instruments with response options beyond the binary requires careful consideration of ways to minimize discomfort and maximize survey quality. A starting point is to take stock of the state of the field with respect to gender identity questions.

State of the Field

In developed Western contexts, an open-ended gender question with a write-in option is often regarded as the best practice in surveys. Such surveys may be executed by web and this approach (open-ended, self-administered) gives the respondent freedom in their answer without having to navigate through a list of options and potentially not find a particular identity as an option. However, this creates a data cleaning problem, as then every respondent’s written answer needs to be cleaned and coded. For large comparative survey projects, such as LAPOP Lab’s AmericasBarometer, this presents several challenges. Importantly, face-to-face surveys are not self-administered, handing over a tablet to offer a write-in option can present security risks, and some individuals may not feel comfortable or capable of typing in their response. Even when the lab uses self-administered approaches, coding open-ended responses requires staff time and other resources. Given these constraints, an approach that is similar but comparatively more practical in interviewer-administered surveys is to ask an open-ended question and to provide answer options for the interviewer to select from.

In 2022, the National Academies of Science Engineering and Medicine (NASEM) released a report recommending the standardization of sex, gender identity, and sexual orientation questions in various survey modes. The NASEM team tested their approach with English and Spanish speakers, but only in a U.S. context. This recommendation was specifically made to the U.S. National Institutes of Health (NIH), but since it is the first all-encompassing, expert-led report on asking these types of questions, we consider it an important resource for those seeking to craft updated gender identity questions. Regarding gender identity, the NASEM recommendation is to first ask about sex, then follow up with a question about respondents’ gender. They recommend this in an effort to include people with transgender experience. A two-step question is also recommended by a study conducted in the LAC region – however, this recommendation is made only for health surveys, similar to the NASEM study. It is important to note that these reports found many respondents conflated gender terms (e.g., man, woman) with sex terms (e.g., male, female). In order to provide answer options that reflect what people understand, one must consider the semantic context in which the survey is being fielded.
Using data from the Canadian Election Studies, researchers compared a 2021 two-step gender question – asking about gender first, then asking about transgender status – to a 2019 single gender question with three answer categories – “man”, “woman,” and “other” (“other” including “transgender” as an example). They found that the single-step question inflated the percentage of non-binary respondents because transgender individuals often chose the “other” category rather than “man” or “woman.” This measurement error altered the relationships between non-binary identity and a number of other variables in the study. The authors argue that having transgender as a separate, mutually exclusive category in a gender question not only produces measurement error but is ethically problematic. The authors recommend that if researchers wish to identify transgender respondents, it is best to ask about transgender identity in a standalone question, rather than asking about assigned sex at birth and re-assigning respondents whose sex and gender differ.

Other research backs up the notion that a two-step question can run into implementation challenges. Pew Research Center has also tested a two-step question approach in the U.S. – first asking about biological sex, then asking about gender identity. They found that, among English and Spanish speakers who mentioned the sex and gender questions in their comments, 30% expressed concerns about the questions, some specifically mentioning that having multiple questions was an issue. The authors mention that if gathering data on sex is not necessary, there is no need to ask about it. Asking only about gender saves respondent time and minimizes privacy concerns and discomfort, as pointed out by other experts.

If researchers are interested in examining attitudes by gender, then surveys should ask respondents about their gender identity as opposed to their sex. Interestingly, a single gender identity measure has been found to explain more variance in political attitudes than measures of sex. In brief, a single-item gender identity measure has the potential to generate higher-quality data without sacrificing crucial information. Of course, what constitutes the most important measure of individual identity depends on the type of study. However, since social science researchers are seldom interested in the anatomy of a person, and are more interested in how respondents identify themselves, it is often appropriate to focus solely on gender identity.
To summarize, recent efforts to consider inclusive approaches to gender identity point to the benefits of using an open-ended question when possible. While in some cases it may be feasible to permit respondents to write in their identity, this approach is difficult to implement in large survey projects, in particular within a face-to-face mode. A reasonable solution for a survey instrument is to ask an open-ended question that contains options for the interviewer to select upon hearing the response. Additionally, in linguistic environments where it is common to conflate gender and sex terms, researchers should add language to the answer options that account for this (e.g., "man/male"). Further, including "transgender" as an answer category to a gender question should be avoided, as it can lead to discomfort and/or measurement error. Rather, separately asking about transgender identity is an option for researchers who want to identify transgender respondents. Finally, because social science research often focuses on gender rather than someone’s sex at birth (in contrast, for example, to health surveys), asking only about gender identity is advised. With these ideas as background, we now turn to a discussion of our lab’s efforts to test alternative gender identity questions in the field.

Testing Alternatives

Pretesting

Pretesting can be informative in research on gender identity questions, revealing nuances in how respondents understand sex and gender. LAPOP Lab has done extensive cognitive pretesting of gender identity questions in multiple LAC countries. Here, we present a set of takeaways from our lab’s experiences in pretesting a survey conducted in Guatemala in 2022. These findings led to important revisions to our approach to asking about gender identity.

A core insight is that asking about gender identity in a direct fashion (For statistical purposes, could you please tell me your gender?) resulted in respondent discomfort. Round 1 of pretesting interviewed 15 respondents, ranging from 19 to 64 years old. Some interviewees became visibly uncomfortable after this question was asked. Another respondent said it was "strange" to be asked this question. Other respondents indicated that other people would likely be "offended" or "upset" by this question. Notably, interviewers also reported being uncomfortable asking the question.

To minimize discomfort, the LAPOP Lab team tested revisions to the instrument that frame the question as a matter of bureaucratic procedure. To that end, the lab assessed responses to wording that continued using the phrase “for statistical purposes” before the question and refined the wording to ask people to “confirm” their gender. The revised question is as follows: For statistical purposes, could you please confirm your gender?
A second round of pretesting was conducted with 10 new respondents, ranging from 19 to 58 years old. Respondents still showed discomfort with the gender identity question, but there was marked improvement from the version in round 1. Some interviewees mentioned it was a little uncomfortable, with a few people saying that you should be able to tell "just by looking at [them]." A number of respondents indicated that they were surprised by the question, in some instances even laughing or calling it strange, but they said they did not mind answering it. Overall, interviewers noted that asking this version of the gender question, with both the “for statistical purposes” and “confirm” language, was not too burdensome for them or the respondents.

The cognitive pre-testing in Guatemala affirmed that a carefully worded open-ended gender identity question is feasible to ask in the field. The lab did not, however, ask a two-step question in these face-to-face pretests given the high levels of sensitivity around these questions. Rather, the lab tested a two-step question in a web survey, which was also conducted in Guatemala in 2022. We present the results from this experiment in the next section.

**Experiment**

To test a two-step gender identity question against the question that emerged from face-to-face cognitive pretests, LAPOP Lab conducted an experiment within an online Computer-Assisted Web Interviewing (CAWI) survey. The survey was designed by LAPOP and implemented via Offerwise; a total of 2,634 respondents were sampled from the firm’s opt-in panel, with quotas to increase the extent to which the sample approaches (though does not fully meet) national benchmarks. The goal of the experiment was to determine whether the single-step gender question has any discernable effects on data collection compared to a two-step sex and gender identity question. Specifically, we assess treatment effects on respondents’ survey satisfaction, breakoff rates, and non-response.

Table 1 identifies the treatment groups and shows the questions and answer categories. Treatment 1 is the single-step gender identity question. Treatment 2 is the two-step question.
### Table 1: Treatment Groups

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Question(s)</th>
<th>Answer Categories</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment 1</td>
<td>Q1TC: Para fines estadísticos, ¿me podría por favor confirmar su género?</td>
<td>1 - Hombre/ masculino (Man/male)</td>
<td>577</td>
<td>43.9%</td>
</tr>
<tr>
<td></td>
<td><em>(For statistical purposes, can you please confirm your gender?)</em></td>
<td>2 - Mujer/femenino (Woman/female)</td>
<td>731</td>
<td>55.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 - No se identifica como hombre ni como mujer <em>(Does not identify as a man or woman)</em></td>
<td>6</td>
<td>0.5%</td>
</tr>
<tr>
<td>Treatment 2</td>
<td>Q1BIO: Al nacer, ¿cuál sexo fue anotado en su partida/acta/certificado de nacimiento?</td>
<td>1 - Masculino (Male)</td>
<td>613</td>
<td>46.6%</td>
</tr>
<tr>
<td></td>
<td><em>(At birth, what sex was written on your birth certificate?)</em></td>
<td>2 - Femenino (Female)</td>
<td>703</td>
<td>53.4%</td>
</tr>
<tr>
<td></td>
<td>Q1ID1: Y más allá de lo que diga partida/acta/certificado de nacimiento, ¿cómo describe su género actualmente? <em>[Mark all that apply]</em></td>
<td>1 - Hombre/ masculino (Man/male)</td>
<td>611</td>
<td>46.4%</td>
</tr>
<tr>
<td></td>
<td><em>(And regardless of what the birth certificate says, how do you currently describe your gender? [Mark all that apply]</em>)</td>
<td>2 - Mujer/femenino (Woman/female)</td>
<td>698</td>
<td>53.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 - Transgénero (Transgender)</td>
<td>6</td>
<td>0.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 - No binario (Non-binary)</td>
<td>6</td>
<td>0.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 - Uso un término diferente <em>(Uses a different term)</em></td>
<td>9</td>
<td>0.7%</td>
</tr>
</tbody>
</table>
A two-step sex and gender identity question was found, in this experiment, to have no effect on overall survey satisfaction, overall comfort, and reported confusion with the gender question compared to the single-step question. Moreover, the two-step treatment had a similar breakoff rate to the single-question treatment. However, when respondents were asked how comfortable they felt answering the question about their gender specifically, the two-step sex and gender identity question resulted in significantly lower levels of reported comfort compared to the single-question format. We were unable to test differences in non-response across treatments because there was little to no item non-response on these questions.

Table 2 shows the survey satisfaction questions asked in the CAWI Guatemala study. The first two focus on general survey satisfaction/comfort (SVYRATE, SVYRATE2), and the last two focus specifically on the gender identity question(s) (Q1SVYRATE1, Q1SVYRATE2). These four satisfaction questions were asked of all respondents at the end of the survey.

The first thing to note is that there are no statistically significant differences for several of the satisfaction/comfort questions. Specifically, Table 2 shows mean responses on SVYRATE, SVYRATE2, and Q1SVYRATE2 are not statistically different across treatment groups. Those in Treatment 2 (the two-step sex and gender question) report similar survey satisfaction (SVYRATE) and comfort (SVYRATE2) compared to those in Treatment 1. Additionally, those in Treatment 2 report similar levels of confusion with the gender question (Q1SVYRATE2) compared to those in Treatment 1. However, it should be noted that of the five people who chose to write something in the open-text option in Treatment 2, four people wrote sexualities instead of genders, meaning they misunderstood the question.30

The second thing to note is that there is a statistically significant difference between treatment groups for reported comfort with the gender question (Q1SVYRATE1). Those in Treatment 2 (the two-step gender question) reported feeling less comfortable with the question about gender compared to those in Treatment 1, by a difference of .18 on the 11-point scale. Figure 1 plots this effect along with the null effects found for the other three attitudinal outcome measures.
### Table 2: Mean Responses to Survey Satisfaction Questions

<table>
<thead>
<tr>
<th>Survey Satisfaction Questions</th>
<th>Treatment 1 Mean</th>
<th>Treatment 2 Mean</th>
<th>Two-Sided P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVYRATE: En una escala de 0 a 10, donde 0 significa que a usted no le gustó nada, y 10 significa que le gustó mucho, ¿que tanto le gustó a usted responder a esta encuesta? (On a scale of 0 to 10, where 0 means you did not like it at all and 10 means who liked it a lot, how much did you like responding to this survey?)</td>
<td>8.77</td>
<td>8.82</td>
<td>0.51</td>
</tr>
<tr>
<td>SVYRATE2: En una escala de 0 a 10, donde 0 significa muy incómodo y 10 muy cómodo, ¿qué tan cómodo se sintió en general con las preguntas que le hicimos? (On a scale of 0 to 10, where 0 means very uncomfortable and 10 means very comfortable, how comfortable were you with the questions we asked you?)</td>
<td>8.40</td>
<td>8.49</td>
<td>0.33</td>
</tr>
<tr>
<td>Q1SVYRATE1: Pensando en la pregunta que hicimos sobre tu género, en una escala de 0 a 10, donde 0 significa muy incómodo y 10 muy cómodo, ¿qué tan cómodo se sintió con esta pregunta? (Thinking back to the question we asked about your gender, on a scale of 0 to 10, where 0 means very uncomfortable and 10 means very comfortable, how comfortable were you with this question?)</td>
<td>9.11</td>
<td>8.93</td>
<td>0.03</td>
</tr>
<tr>
<td>Q1SVYRATE2: Pensando en la pregunta que hicimos sobre tu género, en una escala de 0 a 10, donde 0 significa no confundido y 10 muy confundido, ¿qué tan confundido se sintió con esta pregunta? (Thinking back to the question we asked about your gender, on a scale from 0 to 10, where 0 means very confused and 10 means very confused, how confused were you with this question?)</td>
<td>2.26</td>
<td>2.20</td>
<td>0.70</td>
</tr>
</tbody>
</table>
This experiment provides evidence that asking a two-step gender question that includes sex at birth, even in an anonymous online survey, results in higher levels of discomfort with the gender question compared to the single-step question. It is important to note that the magnitude of these differences, although significant, is relatively small, and further testing in other contexts is needed. Nonetheless, these results support the recommendation that, for social science opinion surveys in a Latin American context, asking a single-step gender question is a best practice for maximizing quality in the data collection process.
Conclusion: Our Recommended Gender Identity Measurement Strategy for Social Science Public Opinion Surveys in the LAC Region

Asking about gender identity can be a sensitive topic, in particular in the LAC region. Our pretesting results serve as evidence of the sensitive nature of the topic from the perspective of respondents and interviewers. In a context marked by comparatively high levels of discrimination and violence against gender minorities, asking about gender identity can generate discomfort that is normatively concerning and that may affect survey data quality. In such a context, it is critical that researchers take extra care in crafting a gender identity measure that empowers respondents of all genders to provide accurate answers and that reduces incentives for interviewers to skip or fabricate an answer to the question.

Drawing on recent research by others and from our own investigations, we offer two core recommendations:

- For social science public opinion surveys in the LAC context, a single open-ended gender identity question, with options for the interviewer to select in the field, is optimal

- Sensitivity concerns with that question can be minimized, but not eliminated, by using language such as “for statistical purposes” and asking people to “confirm” their gender

This Note discussed several considerations in crafting gender identity questions in a social science context. First, current best practices in the field hold that an open-ended question is optimal; it offers respondents the agency to provide an accurate, self-determined answer. While some recommend a write-in approach, this is only feasible on self-administered surveys that are limited in scope. For large survey projects and especially for face-to-face modes, we recommend an approach that is open-ended but provides outcome categories only visible to the interviewer who codes the oral response into the appropriate category. Second, when respondents often conflate gender and sex terms, it is sensible to have the answer options (seen by the interviewer) include “male” and “female,” in case respondents use a sex term instead of a gender one. Third, including “transgender” as an answer category to a gender question should be avoided, as it is normatively problematic and can lead to measurement error. Those who wish to report on transgender individuals specifically should ask a separate question to gauge transgender identity. Finally, to capitalize on respondents’ time, and because social science studies tend to focus on gender rather than sex, we recommend asking only one question on gender (unless a goal of the study is to report on transgender individuals, as mentioned). Our pretests identify a way to ask this question with wording that minimizes discomfort and an experiment in Guatemala in 2022 provides further evidence that a single-step gender question should be preferred over a two-step question asking about sex at birth.
These recommendations have limitations. By not having a “true” open-ended question with a write-in, there are identities that are not captured in the listed options provided to the interviewer. In these cases, the interviewer selects “other,” which loses information and fails to meet the normative ideal of being as inclusive as possible. This includes cases in which an individual identifies only as transgender; unless a separate question about transgender identity is asked, those who state their gender identity as transgender lack a corresponding answer option in the single gender identity question. We also note that our lab’s research on this, while informed by experiences across the Americas, has only systematically considered the topic within the context of pretesting and an experiment carried out in Guatemala in 2022. We hope future research will extend systematic investigations into these issues across time and to other countries of the region.

There is not one perfect way to measure gender in surveys; rather, what approach is best will vary across context, mode, and study purpose. For face-to-face general population social science survey projects in the LAC region, our lab’s research efforts have identified a strong contender for a single-item gender identity question that minimizes respondent discomfort and increases the quality of the data collection process.
Notes

1. The United States Census’ Household Pulse Survey, the European Union Agency for Fundamental Right’s LGBTI survey, the United States National Crime Victimization Survey, and the UK Office for National Statistics Census 2021 are just a few examples.

2. An example is the National Survey on Sexual and Gender Diversity (ENDISEG) 2021 in Mexico. See https://www.inegi.org.mx/programas/endiseg/2021/.

3. We recognize and acknowledge that cultural norms regarding gender identity vary within these regions as well; our discussion focuses on broad patterns in and across regions.

4. Respondents during LAPOP’s cognitive pretesting indicated that their gender should be obvious “just by looking at [them].”


7. See Barton 2016; Castillo 2012.


10. Andreenkova and Javeline 2018.


17. See both Miller et al. 2021 and NASEM 2022.


20. See https://www.vanderbilt.edu/lgbtqi/resources/how-to-ask-about-sexuality-gender


24. For more on LAPOP’s pretesting procedures, see Boidi and Salles Kobilanski 2022. LAPOP’s Director of Fieldwork & Regional Partnerships, Fernanda Boidi, led the pretesting effort in Guatemala.

25. Since these pretests took place in Guatemala, this question was only ever asked in Spanish. Here we provide an English translation.

26. LAPOP would like to thank USAID and particularly Todd Anderson for suggesting the use of “confirm.”

27. 2,634 respondents reached the portion of the survey where they were assigned to an experimental treatment group. Of those, 2,008 completed the entire survey.

28. For a discussion of online sampling in the LAC region, see Castorena et al. 2023.

29. The percentages for Q1D1 do not necessarily sum to 100 since it is a mark all that apply question.

30. Nine people chose the “I use a different term” option. Only 5 people wrote in a term. Write-ins included “gay” (3 respondents), “bisexual” (1 respondent), and “hombre” (1 respondent).

31. We take this recommendation from Albaugh et al. 2023.
References


©2023, LAPOP Lab www.vanderbilt.edu/lapop


Oscar Castorena (oscar.castorena@vanderbilt.edu) is Senior Statistician and Online Research Lead at LAPOP Lab.

Valerie Schweizer-Robinson (valerie.j.schweizer-robinson@vanderbilt.edu) is Statistician and Student Programs Lead at LAPOP Lab.

This report was edited by Liz Zechmeister and Laura Sellers. This report was translated by Camilo Plata and J. Daniel Montalvo. Formatting, production, copy editing, graphics, and report distribution were handled by Mariana Rodríguez and Laura Sellers. Our data and reports are available for free download on the project website. Please follow us on Twitter to stay in touch.

As a charter member of the American Association for Public Opinion Research (AAPOR) Transparency Initiative, LAPOP is committed to routine disclosure of our data collection and reporting processes. More information about the AmericasBarometer sample designs can be found at vanderbilt.edu/lapop/core-surveys.

This Methodological Note is made possible by the support of the American People through the United States Agency for International Development (USAID) and Vanderbilt University. The contents of this Methodological Note report are the sole responsibility of its authors and LAPOP and do not necessarily reflect the views of USAID, the United States Government or any other supporting organization. LAPOP’s AmericasBarometer surveys are supported predominantly by USAID and Vanderbilt University.