

Measuring Gender and Sex in Surveys: Lessons Learned from 50 Years of Cross-National Survey Data and Nonresponse Patterns

Ilona Wysmulek¹ 

¹Institute of Philosophy and Sociology, Polish Academy of Sciences (IFIS PAN)

While gender and sex play a pivotal role in public opinion surveys, their measurement often draws criticism for being imprecise and incomplete. Despite increasing attention, the scope, challenges and changes in the gap between survey practices and contemporary gender theory remain unclear. This paper examines the problems that arise when aligning gender/sex research developments with 50 years of cross-national survey data. Specifically, it analyses survey documentation on respondent's sex and/or gender as well as nonresponse trends across 23 international survey projects, 174 waves and 3329 national surveys, covering the period from 1966 to 2017 and including 156 countries. The results of this large-scale historical analysis of sex and gender survey measures uncover persistent challenges of inadequate documentation, limited sex/gender self-identification options, and lack of nonbinary categories. It also shows increasing popularity of the term “gender” since it appeared in survey documentation in the 1990s. Nonresponse patterns reveal country effect in prevalence of missing codes for sex/gender items, with an overall decreasing trend. The paper discusses observed challenges and addresses survey practitioners' possible counterarguments against introducing change. It contributes to the literature on empirical data collection on sex and gender and offers practical insights to survey producers and survey data users.

Keywords: gender; sex; public opinion survey; comparative; cross-national; survey nonresponse

1 Introduction

Gender and sex are important organizing forces in social life and play a pivotal role in public opinion research. However, measuring them in surveys poses challenges, confusion, and increasing criticism. Scholars note that the gap between theory developments and survey practice is larger for sex and gender than for many other core concepts in the social sci-

ences (Cartwright and Nancarrow, 2022). At the same time, the scope of this gap, its key challenges, and how it has changed over time are not fully understood. Whereas trends in how surveys represent sex and gender have been explored in the United States (Westbrook & Saperstein, 2015), no study has examined these trends across nations and time. This paper builds on the Westbrook and Saperstein's (2015) trailblazing study and contributes to the discussion on measuring sex and gender by offering a systematic, longitudinal and cross-national analysis of surveys and their patterns in reporting nonresponse. Specifically, this article examines how sex and gender are measured in 142 countries across the 23 largest cross-national survey projects, covering a total of 3715 national surveys from 1966 to 2017. It seeks to reveal the trends, problems, and lessons learned from 50 years of longitudinal, worldwide survey research.

Survey measures of sex and gender have been criticized as being imprecise, incomplete, and even misleading (Bit-

Supplementary Information The online version of this article (<https://doi.org/10.18148/srm/2025.v19i3.8464>) contains supplementary material.

Corresponding author: Ilona Wysmulek, Institute of Philosophy and Sociology, Polish Academy of Sciences (IFIS PAN), Nowy Świat 72, 00-330 Warsaw, Poland (Email: ilona.wysmulek@ifispan.edu.pl)

tner & Goodyear-Grant, 2017; Magliozzi, Saperstein, & Westbrook, 2016; Westbrook & Saperstein, 2015). At the very minimum, gender scholars call for more precision in distinguishing the *sex* and *gender identity* of a person at the time of the survey. Although there is an ongoing debate on how and to what extent sex and gender are interwoven, there seems to be a consensus that sex and gender are distinct concepts that may change over time and cannot be used interchangeably (Westbrook & Saperstein, 2015). Gender that is *self-identified* or *otherwise determined* is another important distinction increasingly recognized by scholars (Westbrook & Saperstein, 2015). In addition, there is a heated debate over how many response categories gender measures should include (Bauer et al., 2017; Magliozzi, Saperstein, & Westbrook, 2016). Gender scholars and activists call for surveyors to recognize the diversity of gendered lives, which should be reflected in *nonbinary response categories*. Binary gender measures directly or indirectly hide this diversity and practically violate principles of inclusiveness, respect, and recognition of minority groups (Medeiros, Forest, & Ohberg, 2020). Growing empirical evidence demonstrates that male and female response categories are neither exhaustive nor mutually exclusive (e.g., Magliozzi, Saperstein, & Westbrook, 2016). However, it remains unclear to what extent gender theory has influenced the survey practice.

Previous studies show that contemporary gender scholars and activists advocate for changes in the survey measurement of sex and gender, while survey producers are often resistant to implementing these changes (Bittner & Goodyear-Grant, 2017; Westbrook & Saperstein, 2015). However, the arguments for and against revising sex and gender items in surveys are rarely presented in a balanced and practical way, often creating pressure and confusion for both survey producers and gender researchers. Despite recognizing that survey items may not only reflect but also reinforce categorizations, survey producers may, for different reasons, resist changes in measurement (Durand, 2016). In particular, cross-national survey projects, characterized by complex decision-making processes and comparability concerns, may be careful in introducing innovations in measurement. At the same time, cross-national surveys aim for conceptual clarity across participating countries and often develop standardized documentation practices, which may, in turn, foster discussions on conceptual tensions. Moreover, the extent of the pressure that survey producers face to revise their approach to measuring sex and gender in surveys varies across countries and over time (Cartwright & Nancarrow, 2022). Based on previous research, we may expect changes in terminology, reporting, and documentation practices of sex and gender, which reflect theoretical shifts in sex and gender research. However, we know relatively little about whether and to what extent survey practice

in measuring sex and gender changes across countries and over time, especially in the collaborative contexts of cross-national survey projects.

This paper contributes to the discussion on survey measurement of sex and gender by identifying central themes of contemporary gender theory that are both consequential and potentially applicable in survey practice. Its goal is to trace and systematically assess changes in how cross-national surveys define and operationalize sex and gender, while highlighting the main challenges that arise when aligning contemporary gender theory with existing cross-national survey data.

The key research questions in this paper are as follows:

- How has the approach to collecting survey data on sex and gender changed in cross-national citizen surveys over the past 50 years, from 1966 to 2017?
- What problems arise when aligning current gender theory developments with survey practices?
- To what extent are there period and country effects in the approach to collecting data on sex and gender?
- Are there identifiable patterns on how national surveys report missing data on sex and gender?

2 Defining Sex and Gender: Theory Developments

Discussions on how to define sex and gender are heated, ongoing, and highly politicized. They originated in the early 1950s, when the term “gender” first appears as a necessary conceptual tool to talk about or even to separate sociocultural features from biological ones that are embedded in the sexes (see Simone de Beauvoir as quoted in Stock 2021, p. 14; Muehlenhard & Peterson, 2011). In many of the studies that followed, using two different terms allowed to conceptually differentiate physical and physiological dimensions (chromosomes, genes, hormones, and anatomy), which are typically referred to as “sex,” from self-identity, norms, and relations, which are typically referred to as “gender” (Kennedy et al., 2022; Muehlenhard & Peterson, 2011). “Distinguishing sex from gender was a very important step in recognizing that biology is not destiny—that many of the apparent differences between women and men might be societally imposed rather than natural or inevitable” (Crawford, 2012, p. 26; see also Muehlenhard & Peterson, 2011).

In current theory developments, sex and gender are seen as *two distinct theoretical concepts*, although the relation and determinism embedded in them is still debated. As Bittner and Goodyear-Grant put it: “The separateness of sex and gender is practically orthodoxy in many corners of the sex and gender literature, although this is not to deny the different meanings feminist scholars assign to sex and gen-

der or to the parts of the literature that see no meaningful distinction between the two” (2017, p. 1026). For example, recent studies suggest using the new term “gender/sex” to underline that bodily and sociocultural features are inseparable (Fausto-Sterling, 2019; Hyde et al., 2018; van Anders, 2015; Morgenroth & Ryan, 2021). Using the terms “sex at birth,” “gender identity,” and “gender/sex” becomes indicative of a researcher’s epistemological school of thought, with fundamental differences in standpoints existing among researchers. Despite many of the disagreements, it is clear that *neither “sex” nor “gender” are self-explanatory*. They should be *defined with care and not used interchangeably* when operationalized into survey measurement to serve specific research goals.

Recent research has also shown that *neither sex nor gender is static* (Rosiecka, 2021). Recognizing that gender/sex may change over time makes researchers *call for precision* if we ask in surveys about gender/sex as registered at birth, recorded on birth certificates, or recorded on official documents (Rosiecka, 2021). A possibly dynamic nature of gender/sex is even more important if we are interested in gender identity, gender roles, and living/presenting gender, which are concepts that are strongly embedded in social contexts and less constrained by legal procedures.

Recognizing that gender/sex *determined by others* does not always “match” with *self-identified* gender/sex also has direct consequences for the survey setting. Usually gender/sex in legal documents is “determined by others” as it is “registered” or “recorded” at the moment when a person has little or no impact on it. Countries differ and evolve their approaches and flexibility as to when and how a person can change gender/sex in their official documents. However, it is not clear to what extent and how survey practices within those countries adjust to these changes. In surveys, recorded gender/sex may come not only from participants (respondents) but also from official documents, the interviewer’s assessment, or the response of a household member. Conceptually, these all are ways to capture gender/sex as *perceived or determined “by others.”* Contemporary research and “doing gender” perspectives call for distinguishing this method of categorization from self-identified gender/sex, which is the measure that captures a category that the respondent feels and acts like a member of (Hyde et al., 2018). Recognizing that self-identified and otherwise-determined gender/sex are two different concepts has direct implications for cross-national survey research. Mixing these concepts shakes up the comparability and common theory apparatus that researchers apply at the data analysis stage.

Another key consideration in defining sex and gender is the debate over the *number of categories* these concepts include. Are these *binary* (“dummy”) concepts? The view that humans comprise two distinct categories, women and men or female and male, appeared in the 1930s and has been se-

riously challenged since then by empirical evidence (Hyde et al., 2018). Individuals with physiological features that are not typically male or female are often termed “intersex” or “sex diverse” (Ainsworth, 2015; Hyde et al., 2018). Individuals of different sex/gender from the one assigned at birth are labeled with the umbrella term “transgender.” Individuals identifying themselves in possibly multiple ways outside female and male categories (e.g., “agender,” “gender fluid,” “bigender”) are called by a joint term “nonbinary” (Hyde et al., 2018).

It is not clear *how many* people identify as intersex, transgender, and nonbinary. The estimates vary from about 0.5 % to as much as 2 % of population (Bauer et al., 2017; Hyde et al., 2018; Spizzirri et al., 2021). In a paper published in *Nature*, biologist Claire Ainsworth (2015) underlines that there is growing evidence on the sex spectrum beyond binary. “Biologists may have been building a more nuanced view of sex, but society has yet to catch up” (Ainsworth, 2015, p. 291). Social scientists, however, are divided into camps. Some emphasize that it is necessary to recognize this spectrum, both in sex and gender. “Minorities do not count until they are counted” (see Winters & Conway as quoted in Meier & Labuski [2013, p. 290]). Others, however, defend binary division as a necessary simplification for social functioning (Morgenroth et al., 2021). There is sufficient empirical evidence to make us aware that sex and gender are complex measures and their data collection should be carefully planned and well-documented.

Although the amount of criticism on sex and gender identity measures continues to grow, previous research in the United States has shown that surveys are slow to adapt to requested changes in how they code sex and gender (Westbrook & Saperstein, 2015). At the same time, survey methodologists and survey researchers seem to almost unquestionably perceive sex and gender as important demographic attributes (Magliozzi, Saperstein, & Westbrook, 2016). Sex and gender are thus complex measures, and their data collection should be carefully planned and well-documented. The extent to which cross-national survey data producers have put these ideas into practice is an open empirical question.

3 Data and Methods

This study on trends in measuring sex and gender in cross-national surveys stems from the Survey Data Recycling (SDR) project and uses the SDR v2.0 database (SDR2). The SDR project is a large-scale ex post harmonization initiative that developed an analytical framework and tools for reprocessing information from multiple survey sources and applied them to construct SDR2 (Tomescu et al., 2024; Słomczynski & Tomescu-Dubrow, 2018).

SDR2 is a publicly available multicountry, multiyear database for research on political participation, social capital, and well-being (Slomczynski et al., 2023). SDR2 includes projects that are multiwave, intended to be representative of the adult population, noncommercial, documented in English, and freely available for academic use. In total, it contains harmonized information from 23 international survey projects for 4,402,489 respondents interviewed across 3329 national surveys that, taken together, span the period from 1966 to 2017 and 156 countries (see Table 1). For details regarding SDR2, including criteria for source data selection, harmonization methodology, and documentation, see Tomescu-Dubrow et al. (2024).

This paper follows the survey selection criteria of SDR2 for two main reasons. First, SDR2 is indicative of cross-national surveys that are among the most influential and well-regarded for comparative and longitudinal analysis. Second,

SDR2's selection criteria are also broad enough to include a variety of substantive topics, which makes these surveys highly relevant to study of sex and gender. At the same time, it is important to underline that SDR2 covers only cross-national survey projects. Participating countries may adopt changes to survey measures differently in cross-national survey setting than in their national-level surveys. As a result, the analysis in this paper is limited to trends observed within the context of large collaborative projects and does not capture changes that may be occurring independently at the national level.

To see how cross-national survey projects define and measure gender/sex, I first analyze all documentation sources—questionnaires, codebooks, technical reports, and data dictionaries—available for each wave of the 23 cross-national survey projects listed in Table 1. This extensive documentation was gathered in the SDR project to facil-

Table 1

Cross-National Citizen Surveys in SDR2 and the Scope of Studied Documentation, 1966–2017

Abbreviation	Project name	Project waves	National surveys	People surveyed	Analyzed documents (PDF, TXT, DOC, XLS)
ABS	Asian Barometer Survey	4	43	63,002	51
AFB	Afrobarometer	6	136	204,464	274
AMB	AmericasBarometer	7	164	273,259	546
ARB	Arab Barometer	4	36	43,928	10
ASES	Asia Europe Survey	1	18	18,253	2
CB	Caucasus Barometer	7	20	40,557	26
CDCEE	Consolidation of Democracy in Central and Eastern Europe	2	24	26,354	1
CNEP	Comparative National Elections Project	2	14	26,736	73
EB	Eurobarometer	48	872	930,560	1236
EQLS	European Quality of Life Survey	3	93	105,527	102
ESS	European Social Survey	8	194	371,801	437
EVS	European Values Study	4	120	164,997	227
ISJP	International Social Justice Project	2	19	24,668	5
ISSP	International Social Survey Programme	31	792	1,132,073	2083
LB	Latinobarómetro	19	338	390,744	735
LITS	Life in Transition Survey	3	98	119,072	21
NBB	New Baltic Barometer	6	18	21,601	13
NEB	New Europe Barometer	7	8	70,415	37
PA1	Political Action I (An Eight Nation Study)	1	8	12,588	1
PA2	Political Action II	1	3	4057	1
PPE7N	Political Participation and Equality in Seven Nations	1	7	16,522	8
VPCPCE	Values and Political Change in Postcommunist Europe	1	5	4723	3
WVS	World Values Survey	6	239	336,588	897
Total	–	174	3329	4,402,489	6789

itate the creation of SDR2 target variables (i.e., variables constructed via ex post harmonization). The last column in Table 1 sums up the corpus of survey documents corresponding to each project wave by document type and total number of pages. Project waves differ significantly as to the types and sizes of their master documentation and other accompanying documents available to secondary users.

I examine the corpus of survey documents in search of any text on data collection of respondents' sex and/or gender. This paper concentrates on the respondent because a measure of their gender/sex is available in all national surveys of a given project wave. Information on the sex/gender of other household members, items on attitudes, and beliefs about sex and gender, as well as other sex-/gender-related data are unequally covered in the national surveys of given project waves. Hence, this paper does not analyze them. Also, in some countries, language itself is gendered or may be gendered to a different extent, meaning that it has masculine and feminine nouns and uses gender-related nouns and pronouns in questions. As I rely on English documentation, I do not analyze the gendered-language aspect of survey questionnaires in this paper. Throughout the paper, I follow the definition that "sex" and "gender" capture different theoretical constructs (i.a. Bittner and Goodyear-Grant, 2017). I use the term sex/gender when it is not clear which theoretical construct a study captures and when these constructs are not separated.

For a systematic search, I use keywords: "sex", "gender", "male", "female", "woman", "man", "women", "men". In addition, SDR2 documentation identified the names of the gender/sex variables in each source data file selected for ex post harmonization, and I use these variable names as additional keywords to collect relevant text from the source documentation. In total, the scope of this paper is 215 source variables on sex/gender. Specifically, I used T_GENDER_CWT_SDR2.xlsx, T_GENDER_DVR_SDR2.xlsx and T_GENDER_GVR_SDR2.pdf files published on Harvard Dataverse (Slomczynski et al., 2023).¹

To organize relevant text from source documentation on the project-wave level, I use the SDR2 Detailed Source Variables Report (DVR) tool (Tomescu-Dubrow et al., 2024; Wysmulek, 2019). The SDR2 DVR for respondent's sex/gender (T_GENDER_DVR_SDR2.xlsx in Slomczynski et al., 2023) provides the basis to which I add the information necessary for the analyses in this paper. For example, to show trends, I assign the year of data collection

to project waves and national surveys. The correspondence between project wave and data collection year is high, but not complete: in a few instances, data collection for the national surveys that form a project wave spans a few years. In such situations, I assign to the project wave the data collection year in which the majority of its national surveys were fielded.

I conduct analysis of trends in reporting nonresponse on sex/gender variables at the national survey level of SDR2. For this purpose, I used the SDR2 microlevel datafile to create two new variables for 3329 national surveys: First, a binary flag indicating whether any missing data on sex/gender is present at the national survey level; Second, a categorical variable capturing the type of nonresponse. For the nonresponse type, I coded answer categories such as "don't know," "no answer," "refused," and their combination, like "don't know/refused", as "don't know"-type responses, assuming they come from a respondent or an interviewer. I coded as "other missing" the undocumented missing codes for sex/gender and answer categories like "gender not registered," "not ascertained," "interviewer error," and "NA"—in cases when it is not clear what the NA abbreviation refers to. I assume that the "other missing" category includes both actual responses as well as processing errors or label omissions. I then merged the binary variable for reporting nonresponse and the categorical variable on nonresponse type with the national survey characteristics coded in SDR2, which allowed me to analyze (a) whether the number of surveys reporting missing data on sex/gender increases over time (b) whether the composition of missing data types changes over time and (c) whether the percentage of surveys with missing data is higher in certain countries, and how this relates to their overall participation in cross-national surveys.

4 Results

I identify three key problems in sex and gender measures throughout the 174 project waves and 3329 national surveys of the 23 international projects: (1) the overall poor documentation of sex/gender items, (2) predominant lack of an option to self-identify sex/gender, and (3) lack of a non-binary option for sex/gender with a tendency for a hidden nonbinary option through item nonresponse instead. I discuss each below.

4.1 Weak Documentation of "Gender" and "Sex" Survey Items

The source documentation and source survey data that inform SDR2 carry information on respondents' "sex" and/or

¹ SDR2 harmonized 215 source data files. Generally, a source data file includes all national surveys of a given project wave. However, repositories sometimes provide separate files for each national survey of a project wave, or one single data file for multiple project waves. Therefore, the number of source project waves (174) in SDR2 is different from the number of source data files (215).

“gender.” The term “sex” appears in the documentation of 104 (62%) of the 174 project waves, while “gender” appears in the documentation of 52 waves (29.9%). The documentation of 14 project waves (8%) carries both terms. Thus, we may expect to find information on the following key dimensions related to the sex/gender source measures and the underlying concepts:

1. Clarity on how a survey defines sex and gender: Are these concepts seen as distinct, and if so, which of them does a survey intend to capture?
2. Clarity about how the data on respondents’ sex/gender were collected: Does the information come from the respondent (self-identified sex/gender) or from others (e.g., the interviewer)?
3. Clarity on whether gender/sex is understood as a static or dynamic characteristic. If dynamic, which time point is of interest for the survey (e.g., as registered at birth or as identified at the time of the survey)?
4. In the case of binary sex/gender measures, information on how respondents who did not fit a binary sex/gender category were coded. For example, were they assigned to one of the main categories, added to “missing,” assigned randomly to the main categories, filled in based on other sources during data cleaning, or was yet another approach used?

The analysis of all available source documentation revealed that 51 of the 174 project waves lack any information on how the sex/gender data were collected and coded. Put differently, 29% of the source waves include only the variable label and response options in the data files themselves but no mention of the concept and its operationalization in the source documentation. In one instance, even variable labels were not present, but only a variable name for respondent’s gender, which seems to be a processing error (LITS 1; see Table 1 for definitions of abbreviations).

When documentation sources describe the sex/gender measure, the information provided is generally not clear across many of the identified dimensions of current sex and gender conceptualizations. The analyzed corpus of survey documentation (codebooks, data dictionaries, questionnaires, technical reports) does not include any direct *definitions* of sex and gender, or clarification of these terms, for example, in the notes to interviewers or in the available interviewer training documents.

Although current theory developments suggest that the terms “sex” and “gender” capture different theoretical constructs (e.g., Bittner & Goodyear-Grant, 2017), the analyzed survey documentation does not provide a clear stance on whether terms “sex” and “gender” have been treated as distinct or as synonyms. For example, it is not clear if translation from national languages to English language master documentation has been sensitive to conceptual differences

between terms “sex” and “gender”. As mentioned earlier, 8% of the 174 project waves use both terms within the set of documents describing the same wave. Some projects use these terms interchangeably across different documentation sources of the same wave, or even within the same master documentation source (e.g., AMB 2004). None of the source questionnaires include separate questions on sex and gender. These findings suggest that the terms may be treated synonymously. However, if the translation process is not sensitive to conceptual differences between sex and gender, translation could be a different reason for the presence of both terms in the documentation of a single project wave.

As Fig. 1 shows, the use of term “sex” in cross-national survey documentation has gradually decreased over the years. The term “gender” first appeared in survey documentation of cross-national projects in the 1990s, and by the beginning of 2000s, at least half of all documentation sources began using the term “gender”, either exclusively or alongside “sex”. By 2014–2017, the term “sex” alone appeared in only 35% of cross-national documentation sources, indicating a substantial shift in terminology over time.

Regarding *self-identified versus ascribed sex/gender*, only a few documentation sources directly specify the origin of this information (e.g., AFB). For most of the documentation, to trace the source of the sex/gender data it is necessary to sift through information, for example, by identifying mentions of the procedure in technical reports, sampling design documents, or examining the instructions to interviewers given in questionnaires.

Finally, there is no single survey documentation source that explains whether the sex/gender measure is seen as static or dynamic. Nor is there any description of how to code respondents who do not fit a binary measure, if relevant.

4.2 The Option to Self-Report Sex/Gender Is Largely Absent

In most instances, the respondents’ sex/gender is identified by the interviewer. As Fig. 2 illustrates, “Just write, don’t ask” is a common instruction that the interviewer receives in survey documentation.

From the well-documented surveys, we learn that data on sex/gender may come from the respondent (self-identified), the interviewer (observation), or a household member (as in sampling procedures using a Kish grid and a contact sheet), or they may be administrative data from national registers, also used in stratified sampling designs. Based on theory, we may expect a substantive difference between a group

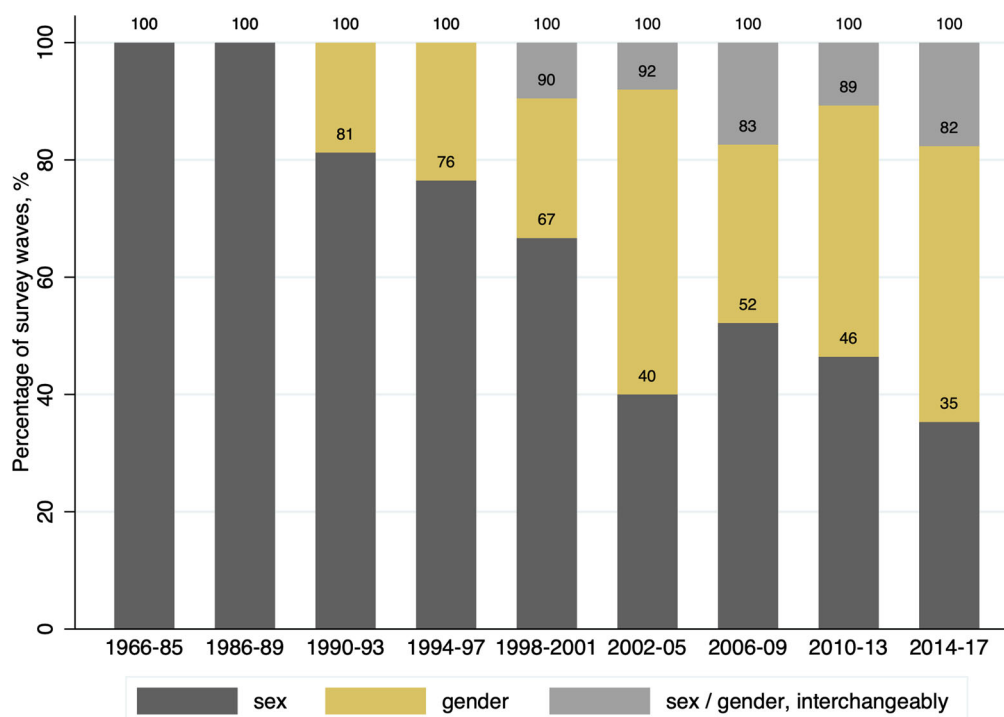


Fig. 1

Changes in Using Terms "Sex" and "Gender" in Documentation of the Survey Project Waves. The total number of analyzed surveys waves of cross-national projects is 174.

S10. Gender of the interviewee (JUST WRITE DON'T ASK)
Male.....1 Female.....2

VARIABLES DEMOGRAFICAS

32.- Sexo del entrevistado (anotar lo sin preguntar)
1) Masculino 2) Femenino

ATENCIÓN: ES UN REQUISITO LEER SIEMPRE LA HOJA DE CONSENTIMIENTO INFORMADO ANTES DE COMENZAR LA ENTREVISTA

Q1. [Anotar, no preguntar] Género: (1) Hombre (2) Mujer

Examples from: AMB_2012_SV, ISSP_2012_MX, LB_2013

Fig. 2

Documentation Screenshots Illustrating "Just Write, Don't Ask" Sex/Gender Items

an individual identifies with and a group others assign an individual to.

As Fig. 3 illustrates, differences in collecting information about respondents' sex/gender occur also among national surveys of the same project wave. Some countries include a question on sex/gender in the questionnaire, which suggests that respondents self-identify their sex/gender. In other countries of the same wave, the documentation states

explicitly that the respondent's sex/gender is coded by interviewers or comes from the contact sheet, sample database, or national registers. It also illustrates country-level differences in using the terms "sex" and "gender" or even omit the terms by asking "Are you a man or a woman?" (NL, Netherlands, ISSP 2011).

I checked whether there is a tendency for survey items using the term "gender" to appear as a question in a master

SEX - Sex of Respondent

BVQ_01. SEX

AU BE-WAL DK FI FR NO: Are you ...?

BE-FLA: What is your sex? Note whether the person is a male or a female.

CH: Information automatically taken from sample database

CL CN CZ DE ES GB-GBN RU SK TR US: Sex coded directly by interviewer (without asking respondent)

HR PH PT: Respondent sex

IT: You are:

KR: Now I would like to ask about the members of your household. Please include all members who are living in your house and who are temporarily living somewhere else. Answer the following question beginning with yourself: Is [Person] male or female?

LT SI: Your sex

NL: Are you a man or a woman?

SE: Not asked in survey, taken from National Register.

TW: Gender

ZA: Copy from contact sheet

1 Male

2 Female

9 No answer

Fig. 3

*Varying Sex/Gender Questions and Modes of Data Collection within a Single Survey Wave:
Excerpt from the ISSP 2011 Codebook*

questionnaire and/or master codebook (in English) and for those using the term “sex” to come from other sources than a questionnaire (administrative data, contact sheets etc.). I did not find consistent patterns across project waves. It has been also challenging to assess this due to inconsistencies and lack of precision in documentation.

Secondary data users often seem to be not aware of this variation and assume comparability of this measure across countries, even if the concept (sex vs. gender) or measurement approach (self-identified vs. otherwise determined) differs.

4.3 Lack of Nonbinary Answer Options for Sex/Gender

With one exception, AMB Canada, 2017, none of the 3329 national surveys from 1960 till 2017 included in SDR2 provide the option to identify sex/gender other than as male or female. EVS 1981 features the three answer options: (1) “Male,” (2) “Female, housewife,” and (3) “Female, not housewife.” Technically the EVS 1981 question is not binary, but conceptually it is. Although troubling in itself, this question does not provide an answer category in addition to the binary male/female divide.

If we take the perspective of a minority group, then it seems respectful and appropriate to give members of this group space to answer a survey question. From the basic standards of questionnaire development, we know that

response options should be exhaustive of all possibilities. For example, survey methodologists Krosnick and Presser (2010) call it a “convention wisdom” of question design that response options should be “exhaustive and mutually exclusive” (p. 264). Binary sex/gender items violate this standard (see also Magliozzi, Saperstein, & Westbrook, 2016).

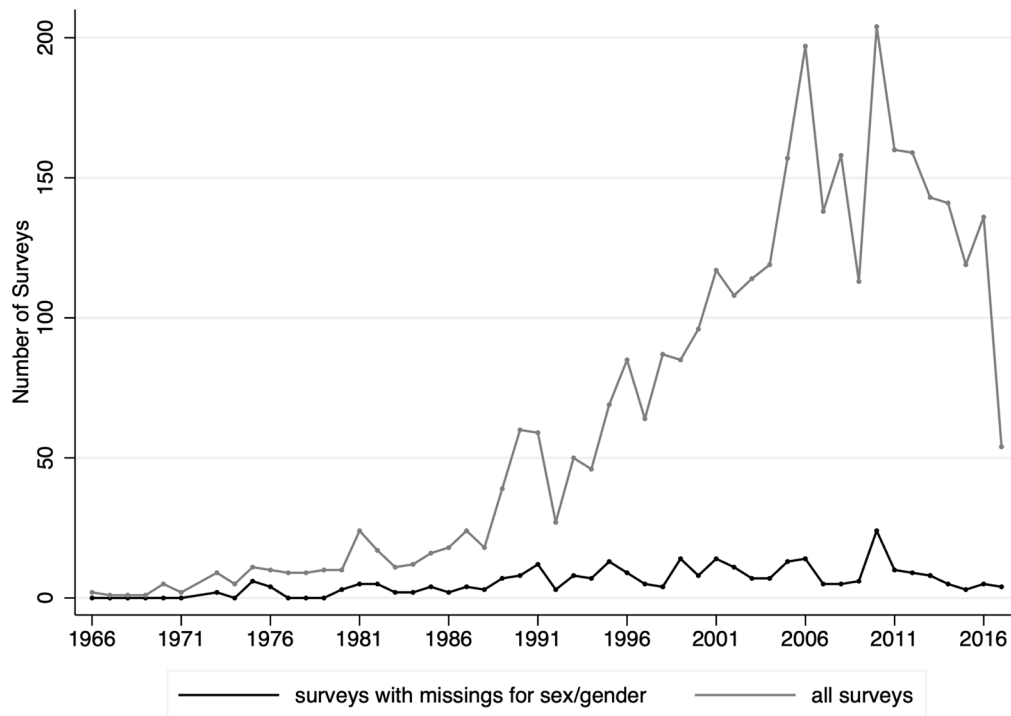
Nonresponse Patterns to Sex/Gender Items: A “Hidden” Nonbinary Option?

For the binary sex/gender measure, we can envision that a missing value (e.g., “don’t know” “refuse to answer,” “other missing”) may indicate instances when the “truthful” response did not fit into the available dichotomous sex/gender answer option. To investigate this possibility, I systematically checked all missing codes for sex/gender items in SDR2 (for details, see Data and Methods).

Missing data for sex and/or gender appear in 285 of 3329 national surveys.² The “don’t know” category is twice as common as “other missing” (189 and 96 national surveys, respectively).³ In terms of frequencies at the respondent level, in our analyzed sample of 4,396,012 respondents,

² Five national surveys (out of 3329 surveys) have no data on sex/gender, or, in other words, there is a full item nonresponse. These are Serbia and Montenegro in EB 77.3 (2012) and South Africa, Mexico, and Finland in WVS 1 (1980s). These cases are coded as QNA (i.e., question not asked or .r) in the SDR2 Stata datafile. I exclude these five countries from further nonresponse trend analysis.

³ Although technically possible, we do not have situations when both “don’t know” and “other missing” appear for the same sex/gender measure in a given national survey.

**Fig. 4**

Changes in Reporting “Missing” for Sex/Gender Items in Surveys of Cross-National Projects. The total number of analyzed surveys is 3329, and among them the total of 285 surveys have missing data on sex/gender.

as few as 0.05 % (2204 respondents) received the SDR2 code “don’t know,” and 0.02 % (661 respondents) “other missing” for their sex/gender.

The first observation is that in ASES, EQLS, PA2, PPE7N, and VPCPCE, all respondents are categorized into the groups of either male or female (i.e., no missing values on respondents’ sex/gender). Other than that, there are no clear signs of survey project patterns. In the vast majority of analyzed survey projects, missing on sex/gender appear in some countries, while in other countries of the same wave they are not present. Larger projects—like EB, ISSP, WVS, and ESS—have a greater variation in approach to missing data for sex/gender on the national level than in smaller surveys like CB, NBB, and NEB.

There are some observable changes over time. Although generally the amount of survey data has been rapidly growing since 1966, the relative number of surveys that report missing on sex/gender items is decreasing (Fig. 4). It may be a sign of standardization in survey practices throughout time. It may also be due to a perception among survey providers that missing on sex/gender is an indicator of a poor(er) survey quality and thus should be cleaned, corrected, and/or avoided. It may be that the whole records

with missing on sex/gender are removed from the data⁴. In the analyzed documentation, the strategy of how missing on sex/gender has been treated is not well described.

Despite the general decreasing trend, the missing data that do appear are better labeled and increasingly refer to “don’t know” but not “other missing” (see Fig. 5). This is especially visible in the trend across ISSP waves—one of the largest survey projects in our sample with relatively clear documentation on sex/gender.

Country-level effects in allowing for missing codes for sex/gender items are the most pronounced in the analyzed data. Some countries tend to have missing codes for sex/gender items no matter what survey project they participate in (see Figs. 5 and 6). The leading country with missing codes for sex/gender is Canada. Out of the total 26 surveys from Canada, 81 % have nonbinary measures: 20 times with missing codes and once with the third response option “other,” as mentioned above. Other leading countries in which the missing codes for sex/gender items appear most frequently are Australia (in 76 % of surveys), New Zealand (75 %), and Thailand (71 %), as illustrated in

⁴ I thank the reviewer for adding that this may also be the practice, based on first-hand experience.

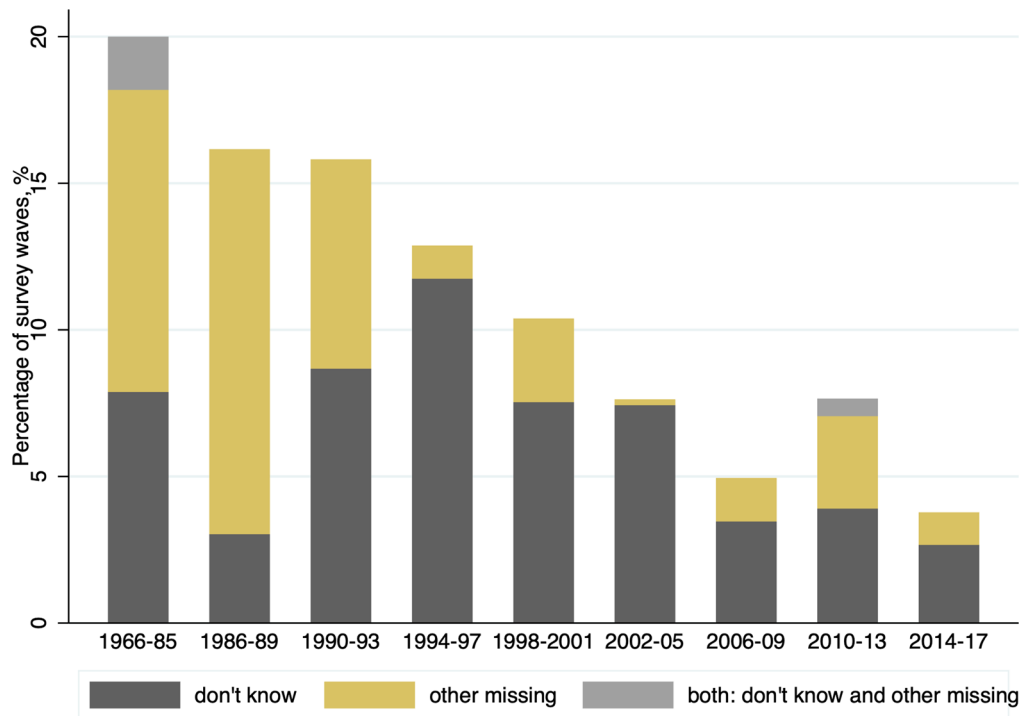


Fig. 5

Change in composition of missing data types over years. The total number of analyzed surveys is 3329.

Fig. 6. Additionally, Fig. 7 provides a closer view to trends in reporting nonresponse across countries with the most frequent participation rate (20 surveys or more per country) in the period from 1966 till 2017. It shows that Canada, Australia and New Zealand relatively frequently participate in international survey projects and have a consistently different approach to measuring sex/gender compared to other countries⁵.

5 Discussion

How is it possible that a variable largely recognized as important receives such severe criticism and seems to be resistant to change in line with new developments in theory? One possible answer is that surveyors perceive sex and gender as relatively easy concepts within an otherwise complex instrument. Questions on, for example, how to measure precarity or attitudes toward migration may seem more difficult

and urgent to address than those related to sex and gender. The problems with the sex and gender measure then would have been just an oversight, which is relatively easy to fix. However, it may also be possible that the survey research community has its own counterarguments and reasons for not changing sex and gender items.

One counterargument against the change in gender/sex items is the potential threat of putting a respondent and an interviewer into a *socially awkward* situation (Westbrook & Saperstein, 2015). How would an interviewer ask about one's gender/sex? How would this question be taken by the respondent? Why do we need to complicate a relatively easy matter?⁶ In survey practice, the change in how we ask about sex and gender is perceived by some as *high risks, questionable gains*. Survey researchers recognize that interviewers are *not trained* to ask about sex and gender and fear that respondents may be offended or discouraged by these questions (Medeiros, Forest, & Ohberg, 2020).

Because a survey question is a communication tool between a researcher and hundreds or even thousands of people, it relies on the ability to formulate questions that are "acceptable and easy to answer" to "as many people as pos-

⁵ Thailand does not appear in Fig. 7, because the total number of surveys for Thailand in SDR2 is below 20. Specifically, Thailand participated a total of seven times in three survey projects (Asia Europe Survey, Asian Barometer survey and World Value Survey) in the studied period from 1960 till 2017.

⁶ I refer here to a reaction to the presentation of this paper that we got at a conference from one of the survey managers.

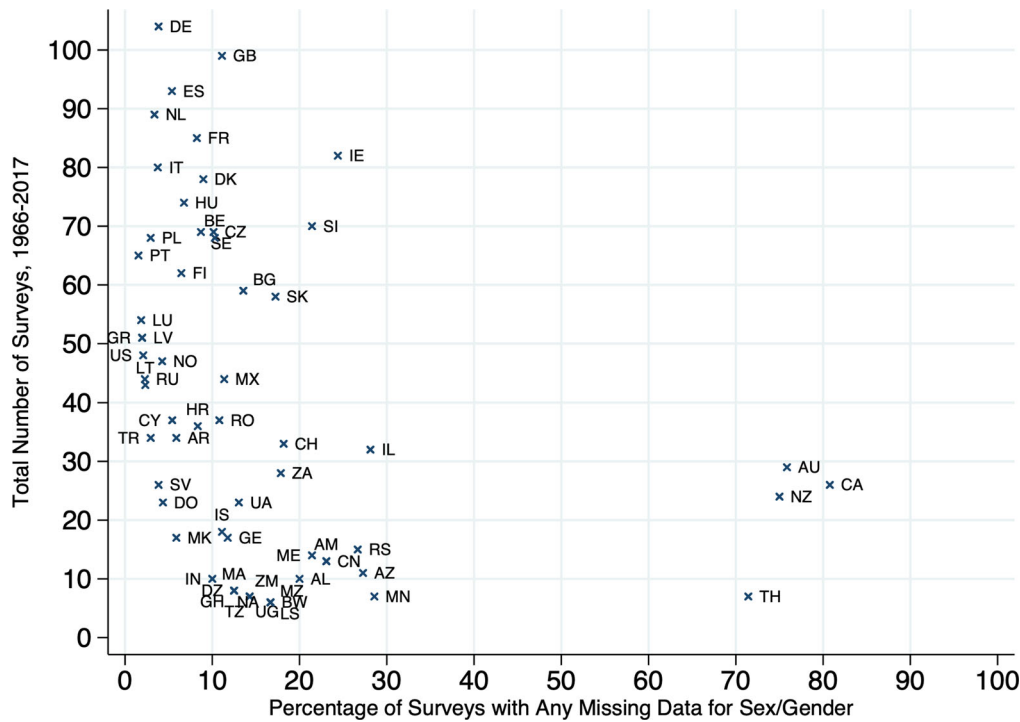


Fig. 6

Country-level effects: Percentage of Surveys with Any Missing Data for Sex/Gender in Relation to the Total number of Surveys by Country. The total number of analyzed surveys is 3329.

sible across all population groups” (Rosiecka, 2021, p. 5). Consequently, it takes the *perspective of the majority* (Durand, 2016). In this trade-off, a possible assumption is that the potential group for which the conventional sex/gender question is problematic is smaller than the group that may be offended or discouraged by the “new” sex/gender question, and as a consequence it may cause bigger problems for the overall survey quality. The risk of losing neutrality and clarity in the sex/gender question is multiplied in cross-national research, with varying cultural and political contexts around the sex/gender topics.

In practice, cross-national survey developers are under constant pressure from *survey length, costs, and comparability* issues (Groves, 2004; Johnson et al., 2018). Not only for sex and gender items, but for any survey items, changing an instrument leads to a set of consequences that methodologists are increasingly aware of. Cross-national surveys come through a decision-making path where each step is prone to different error types. This path can be described with the survey life cycle model and the total survey error paradigm, which aim to raise awareness about decision trade-offs and minimize the sum of all errors in care for data quality (Smith, 2011; Survey Research Center, 2016). Technically, it brings us to the point that both the lack of change as well as the change in the sex/gender instrument

will lead to “survey errors,” albeit of different kinds. It may suggest that despite the growing awareness of possible “errors” that come with repeating the “traditional” sex/gender measure, these errors are still perceived by surveyors as less challenging for research than introducing a new instrument. However, this possible trade-off argument does not explain the poor documentation of the sex/gender measure.

Another possible source of reluctance may stem from the awareness of the complexity inherent in the cross-national survey data collection process. This complexity, involving various layers of decision-making and different decision-makers (actors), is detailed in the survey life cycle model (Survey Research Center, 2016). For instance, a team of researchers at the headquarters of a large international survey might introduce a third answer category to the question on gender, instructing that only the question, but not the answer options, should be read out by an interviewer. Countries participating in this hypothetical survey may have different reactions to this change, leading to increased training time, consultations, and overall project costs. Assuming that all countries agree to implement the change in their national-level questionnaires, they would face further challenges related to translation and comparability issues stemming from

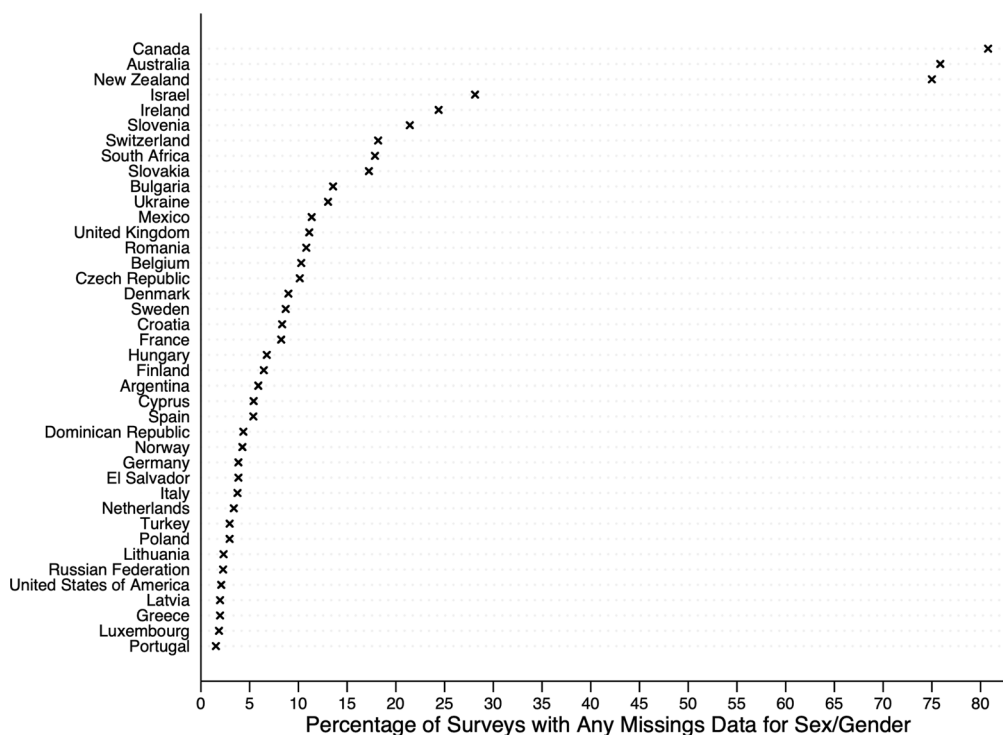


Fig. 7

Percentage of Surveys with Any Missing Data for Sex/Gender for Countries with Most Frequent Participation in Cross-national Projects in 1966–2017. Figure includes only countries with the total of 20 surveys or more during the period 1966–2017.

*linguistic differences.*⁷ Once these challenges are dealt with, there is the next step of the actual fieldwork, where interviewers, their coordinators, and, of course, their respondents may adopt different and inconsistently documented ways to deal with the modified sex/gender question. Possible differences in fieldwork practices are particularly important when considering the “allowance” for nonresponse on sex/gender questions. This example showcases possible challenges associated with introducing change to the sex/gender instrument that are particular to large-scale, cross-national projects. They may be an additional source of reluctance to introduce the change, on top of other arguments. The risk of a flawed concept-realization link holds for every concept in cross-national survey projects, but it is especially pronounced for culturally embedded and highly politicized concepts, which gender and sex are.

On top of the data collection challenges, there is a risk that it would still be not quite clear how to best use the nonbinary questions on sex and gender (Kennedy et al., 2022). In addition to the possible problem of the small

number of cases, some researchers underline challenges with weighting procedures that nonbinary sex/gender items bring (Kennedy et al., 2022). These problems are not unmanageable, although they are indicative of the extent to which a binary sex and gender division is embedded in our everyday life, not only in surveys.

6 Conclusions

This paper offers a large-scale historical analysis of how our approach to survey measures of sex and gender developed over time, since the first publicly available cross-national survey datasets. This paper does not provide answers on how to best measure sex and gender, but it identifies key problems and trends in measuring sex and gender in light of ongoing theoretical developments, aiming to foster discussion and decision-making across areas of conceptual tensions.

A literature review reveals that there has been a growing recognition of sex and gender diversity over the years. There are also changes in survey guidelines and postulates of gender activists on coding sex and gender in surveys. With careful attention to these changes, I checked 3329

⁷ Recent studies show that there are languages in which the English term “gender” is increasingly borrowed and incorporated in everyday speech, with an often derogative connotation (Mazzuca et al., 2023).

cross-national surveys from the 23 largest cross-national survey projects to analyze the main trends, problems, and nonresponse patterns in cross-national survey data on sex and gender.

I found that, despite developments in gender theory and postulates of gender activists, cross-national survey research has been slow to implement changes in the measurement of sex and gender. The approach to coding sex and gender has not changed much in 50 years of cross-national survey research. As a result, cross-national data for 1966–2017 do not reflect any noticeable developments of gender theory.

Based on this analysis, I identified three key problems in how respondents' sex and gender are captured in major international survey projects. First, respondents' sex and gender are poorly documented in cross-national surveys. Specifically, there is a predominant lack of clarity on how a survey defines sex and gender. Connected to clarity on definition, sex/gender items are documented as static characteristics and the time point of interest for the measure is often not clear (for example, registered at birth or currently). There is also a lack of precision as to data collection mode—self-identified versus otherwise-determined sex/gender. Furthermore, the analyzed survey documentation lacks clarity on a survey strategy of how to code non-binary respondents in the case of a binary measure (for example, if they were assigned to one of the categories, added to “missing,” assigned randomly to two categories, or filled in based on other sources during data cleaning or if yet another approach was used).

In addition to this umbrella problem of poor documentation, I illustrate the second key problem with sex/gender survey data, which is a predominant lack of self-identified sex/gender and the often-mixed data collection modes within the same survey waves. In most cases for which I was able to deduce the data collection mode from the documentation, sex/gender items were coded by the interviewer as “Just write, don't ask.”

The problem is not trivial. In face-to-face surveys it may seem awkward to ask a person in front of you about their sex or gender. Interviewers are not trained to do so, and survey providers also do not seem convinced that this is necessary. It is hard to assess the scale of the problem. To what extent do we miss something important? To what extent would self-identified sex/gender responses differ from sex/gender identified by others? It is hard to judge as we have no data about it. We may assume that methodological bias is not large. However, despite a methodological bias, we may think about the other social and ethical responsibilities that those survey data collectors have. We may think about people whose sex/gender is hard to identify based on physical appearance or who self-identify differently than their visual appearance suggests. In this case we put both

respondents and interviewers in an uncomfortable position, as interviewers do not have proper training on how to ask about sex/gender in a way that is not socially awkward. Most likely, they guess or assume what sex/gender a person is. While this issue is especially pronounced in face-to-face interviews, the growing use of self-administered survey modes such as online questionnaires may help reduce reliance on interviewer assumptions.

However, the problem becomes even more complex if we think about languages in which gender identification is necessary for the grammatical construction of a question. These languages are often called “gendered languages” by linguists, and their nouns, pronouns, adjectives, and verbs are categorized as masculine, feminine, or neuter. In the instances of such “gendered language,” the mistake of wrong identification of gender is high. In face-to-face interviews, respondents may feel uncomfortable throughout the interview. Would they not correct an interviewer right away? Most probably some would, those who are open and comfortable with their sex/gender identity and self-confident enough not to take the interviewer mistake to heart. However, in some cases such a mistake of wrongly identified gender can put a shadow on the whole questionnaire and may have longer consequences of deepening stress, psychological discomfort, and insecurity connected to the problem that others are misidentifying one's gender based on appearance. These implications are also relevant for self-administered survey modes, as questionnaires may follow different scenarios regarding whether and how to account for self-identified gender in their design. This includes decisions about the placement of gender questions—whether they appear early in the questionnaire or later—and whether respondents are redirected to different questionnaire versions.

The last and third problem that I identify is that sex/gender measures in cross-national surveys of 1960–2017 are binary only. I checked 3329 national surveys and found only one instance of the nonbinary survey measure—AMB Canada 2017, which codes “male”, “female”, and “other”. I checked whether there are nonresponse patterns in sex/gender, which I propose to interpret as a possible “hidden” nonbinary option. Analysis of reporting nonresponse by national surveys reveals that there are some observable changes over time and some country effects that are stronger than survey project effects. The relative number of surveys that report missing on sex/gender is decreasing. At the same time, missing on sex/gender are better labeled with time and increasingly refer to “don't know” rather than undocumented missing. There are four leading countries in which sex/gender measures are nonbinary or, in other words, allow for some flexibility through missing codes. These are Canada, Australia, New Zealand, and Thailand. It is necessary to underline once again that these are trends

for countries participating in cross-national survey projects. The pace of change and tendencies in measuring sex/gender may be different for national surveys within studied countries. National-level developments, as well as the reasons why some countries measure sex/gender differently than others, even within international survey contexts, seem both interesting and important to explore in future research.

This paper demonstrates that gender theory has little impact on survey practice for cross-national survey projects, but it remains unclear why. Is it an oversight, neglect, or interdisciplinary communication problem? Is it related to cultural context with variation across countries, reflecting the conservative and liberal spirit of a given time and place? Can gender scholars simply ignore survey research as a useless instrument for their purposes?

This paper pinpoints key problems and difficulties with survey measures of sex/gender with respect to the risk of convincing the already convinced and not reaching those who are resistant to change. What if we change nothing? What can go wrong? What if citizen surveys continue to use the concepts of sex and gender interchangeably? This paper shows that it would result in a mixture of answers on sex at birth, legal sex, and gender identity. Although we do not know for sure, we can conservatively assume that these concepts are aligned for a majority of our respondents and would lead to no “significant” difference in collected data. However, three groups would bear the main consequences of inaction: respondents with misaligned sex and gender (no matter how small the group is, they deserve recognition and respect), interviewers (no proper training if the problem appears), and researchers (difficulties in interpretation, limiting analytical possibilities to test some theories). The main problem seems to lie not in the data that we end up collecting but in the process of their collection and in their quality, which is understood as the interpretative power of data.

What are the possible options as to how to react to the problem? The least controversial and easiest to apply is to *choose*, on substantive grounds, which concept the survey intends to measure. This leads to three practical steps: (1) to *design* survey instruments with consistent terminology, (2) to include an explanation of the concept in *interviewer trainings* and instructions, and (3) to develop documentation for *secondary users* of survey data that is *consistent* and clear as to what concept they measure.

The second option is to include a set of questions, for example, a question on sex at birth and another one on current gender identity. They would provide new analytical pathways, even if there was minimal change in the measurement approach that the survey used so far; two separate variables would then be captured through data collection: sex and gender identity. Even if the way to collect data on sex and gender is more experimental but not traditional, the

need for clarity as to the sex/gender definition operationalized in the survey remains. Similarly, key practical steps in applying this strategy involve clarity and consistency during instrument development and interviewer training stages and in published documents for secondary users.

From the radical social-constructivist stance, there is an option to resign from collecting data on sex and gender all together as any categorization is the means of reinforcing “‘hierarchies’ of dominance and subordination” (see Judith Butler as quoted in Stock, 2021, p. 20). Although not many seem to support this radical standpoint, researchers seem to increasingly recognize that surveys not only reflect but also shape reality.

In this paper, the main claim is not that coding practices should change dramatically. However, based on analysis of a large pool of survey data across the world, it calls for not neglecting careful instrument development and documentation of sex and gender items.

Acknowledgements I cordially thank Irina Tomescu-Dubrow, Kazimierz M. Slomczynski, Weronia Boruc, Nika Palaguta and Przemek Powalko and the Survey Data Recycling project team for their feedback and their immense support in data collection for this paper. I am also thankful to Joshua K. Dubrow and Jakub Wysmulek for their valuable comments.

Funding No funding was received to assist with the preparation of this manuscript.

Data Availability This paper uses Survey Data Recycling SDR database and its accompanying documentation on gender/sex published at Harvard Dataverse: <https://doi.org/10.7910/DVN/YOCX0M>. Additional research data collected and used in the paper has been added as a supplementary document.

References

- Ainsworth, C. (2015). Sex redefined. *Nature*, 518(7539), 288–291. <https://doi.org/10.1038/518288A>.
- Bauer, G.R., Braimoh, J., Scheim, A.I., & Dharma, Ch (2017). Transgender-inclusive measures of sex/gender for population surveys: mixed-methods evaluation and recommendations. *PLoS ONE*, 12(5), 1–28. <https://doi.org/10.12759/hsr.39.2014.2.257-291>.
- Bittner, A., & Goodyear-Grant, E. (2017). Sex isn’t gender: reforming concepts and measurements in the study of public opinion. *Political Behavior*, 39(4), 1019–1041. <https://doi.org/10.1007/s11109-017-9391-y>.
- Cartwright, T., & Nancarrow, C. (2022). A question of gender: gender classification in international research. *International Journal of Market Research*, 64(5), 575–593. <https://doi.org/10.1177/14707853221108663>.

- Crawford, M. (2012). *Transformations: women, gender and psychology* (2nd edn.). McGraw-Hill.
- Durand, C. (2016). Surveys and society. In C. Wolf, D. Joye, T. Smith & F. Y.-Ch (Eds.), *The SAGE Handbook of survey methodology* (pp. 57–66). SAGE.
- Fausto-Sterling, A. (2019). Gender/sex, sexual orientation, and identity are in the body: how did they get there? *Journal of Sex Research*, 56(4–5), 529–555. <https://doi.org/10.1080/00224499.2019.1581883>.
- Groves, R. M. (2004). *Survey errors and survey costs*. Wiley Series in Survey Methodology. Wiley. Reprint
- Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2018). The future of sex and gender in psychology: five challenges to the gender binary. *American Psychologist*, 74(2), 171–193. <https://doi.org/10.1037/amp0000307>.
- Johnson, T. P., Pennel, B.-E., Stoop, I. A. L., & Dorer, B. (2018). The promise and challenge of 3MC research. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multi-cultural contexts* (pp. 1–12). John Wiley & Sons. 3MC.
- Kennedy, L., Khanna, K., Simpson, D., Gelman, A., Jia, Y., & Teitler, Y. (2022). *He, she, they: Using sex and gender in survey adjustment*. Vol. 3. New York: Department of Statistics, Columbia University. Unpublished manuscript. June 3, 2024, from stat.columbia.edu/~gelman/research/unpublished/Using_gender_in_surveys.pdf
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd edn., pp. 264–313). Emerald.
- Magliozzi, D., Saperstein, A., & Westbrook, L. (2016). Scaling up: representing gender diversity in survey research. *Socius*, 2, 1–11. <https://doi.org/10.1177/2378023116664352>.
- Mazzuca, C., Borghi, A. M., van Putten, S., Lugli, L., Nicoletti, R., & Majid, A. (2023). Gender is conceptualized in different ways across cultures. *Language and Cognition*. <https://doi.org/10.1017/langcog.2023.40>.
- Medeiros, M., Forest, B., & Öhberg, P. (2020). The case for non-binary gender questions in surveys. *PS—Political Science and Politics*, 53(1), 128–135. <https://doi.org/10.1017/S1049096519001203>.
- Meier, S. C., & Labuski, C. M. (2013). The demographics of the transgender population. In A. K. Baumle (Ed.), *International handbooks of population* (Vol. 5, pp. 289–330). Springer.
- Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: an integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, 16(6), 1113–1142. <https://doi.org/10.1177/1745691620902442>.
- Morgenroth, T., Sendén, M. G., Lindqvist, A., Renström, E. A., Ryan, M. K., & Morton, T. A. (2021). Defending the sex/gender binary: the role of gender identification and need for closure. *Social Psychological and Personality Science*, 12(5), 731–740. <https://doi.org/10.1177/1948550620937188>.
- Muehlenhard, C. L., & Peterson, Z. D. (2011). Distinguishing between sex and gender: history, current conceptualizations, and implications. *Sex Roles*, 64(11–12), 791–803. <https://doi.org/10.1007/s11199-011-9932-5>.
- Rosiecka, H. (2021). Methodology assurance review panel (MARP). Methodology for decision making on the 2021 Census sex question concept and associated guidance. <https://uksa.statisticsauthority.gov.uk/wp-content/uploads/2021/04/Methodology-for-decision-making-on-the-2021-Census-sex-question-concept-and-associated-guidance-1.pdf>. Accessed 3 June 2024.
- Slomczynski, K. M., & Tomescu-Dubrow, I. (2018). Basic principles of survey data recycling. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts* (pp. 937–962). Wiley. 3MC.
- Slomczynski, K. M., Tomescu-Dubrow, I., Wismulek, I., Powalko, P., Jenkins, C. J., Ślarzyński, M., Zieliński, M. W., Skora, Z., Li, O., & Lavryk, D. (2023). SDR2 database. Harvard Dataverse. <https://doi.org/10.7910/DVN/YOCX0M>
- Smith, T. W. (2011). Refining the total survey error perspective. *International Journal of Public Opinion Research*, 23(4), 464–484. <https://doi.org/10.1093/ijpor/edq052>.
- Spizzirri, G., Eufrásio, R., Lima, M. C. P., de Carvalho Nunes, H. R., Kreukels, B. P. C., Steensma, Th D., & Abdo, C. H. N. (2021). Proportion of people identified as transgender and non-binary gender in Brazil. *Scientific Reports*. <https://doi.org/10.1038/s41598-021-81411-4>.
- Stock, K. (2021). *Material girls: why reality matters for feminism*. London: Fleet.
- Survey Research Center (2016). Guidelines for best practice in cross-cultural surveys. <http://ccsg.isr.umich.edu>. Accessed 3 June 2024.
- Tomescu-Dubrow, I., Slomczynski, K. M., Wismulek, I., Powalko, P., Li, O., Tu, Y., Ślarzynski, M., Zieliński, M. W., & Lavryk, D. (2024). Harmonization for cross-national secondary analysis: survey data recycling. In I. Tomescu-Dubrow, C. Wolf, K. M. Slom-

- czynski & J.C. Jenkins (Eds.), *Survey data harmonization in the social sciences* (pp. 147–167). Wiley.
- 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. <https://doi.org/10.1007/s10508-015-0490-8>.
- 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. <https://doi.org/10.1177/0891243215584758>.
- Wysmulek, I. (2019). From source to target: Harmonization workflow, procedures and tools. Building multi-source databases. <https://wp.asc.ohio-state.edu/dataharmonization/about/events/building-multi-source-databases-december-2019/>. Accessed 3 June 2024.