



An Adaptation of Gradational Gender Scales in a Japanese Sample: Nonresponse Rates and Response Distributions for Self-Rated Masculinity and Femininity

Christina Pao^{1,2} , Fumiya Uchikoshi³ 

[1] Department of Sociology and the Office of Population Research, Princeton University, Princeton, NJ, USA.

[2] Department of Sociology, University of South Carolina, Columbia, SC, USA. [3] The Harvard Academy for International and Area Studies, Weatherhead Center for International Affairs, Harvard University, Cambridge, MA, USA.

Measurement Instruments for the Social Sciences, 2026, Vol. 8, Article e18677, <https://doi.org/10.5964/miss.18677>

Received: 2025-07-02 • **Accepted:** 2026-01-16 • **Published (VoR):** 2026-04-16

Handling Editor: Dorothee Behr, GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

Corresponding Author: Christina Pao, Sloan College, 911 Pickens St, Columbia, SC 29208, USA. E-mail: christina.pao@sc.edu

Supplementary Materials: Code [see [Index of Supplementary Materials](#)]



Abstract

Although scholars have long depicted gender as multidimensional, many surveys have historically captured gender coarsely. A recent methodological advancement has been gradational gender scales—unipolar scales that ask about self-rated femininity and masculinity. Nonetheless, these scales have yet to receive wider attention outside of the U.S. and Europe. Using data collected by the Social Science Japan Data Archive (SSJDA) in October 2023 and February 2024, we evaluate response patterns to gradational gender scales among Japanese respondents. We find high levels of nonresponse for these questions, and responses are less bimodally distributed and less extreme in value than in other country contexts. Respondents are particularly less likely to provide answers to scales that are “gender-atypical”—e.g., masculinity for women, and femininity for men. These findings reveal large differences in response patterns in a Japanese sample, indicating the need and promise for future research when administering gradational gender scales in different contexts.

Keywords

gender expression, Japan, survey methods, masculine, feminine



This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Theoretical Background

A large body of literature has shown that gender is an act to be “done” or “performed” (Butler, 1988; West & Zimmerman, 1987), resulting in a gendered organization of society maintained by social interactions (Goffman, 1977; Ridgeway & Correll, 2004). Thus, the way that gender is *expressed* in these interactions, e.g., via femininity and masculinity, offers differential social power and status (Connell, 2020; Edwards, 2005; Hamilton et al., 2019). Despite this nuanced theoretical understanding of gender, measures of gender in empirical data have historically been coarse—often proxying the complexity of gender as a binary sex variable (male/female) (Ridgeway & Saperstein, 2024; Westbrook & Saperstein, 2015). This mismatch between theoretical understanding and empirical measurement has come to the fore as there have been changes to social meanings of gender. For instance, transgender, nonbinary, and otherwise noncisgender identification has become more prevalent (at least via disclosure on surveys) (Lagos, 2022); and, among cisgender individuals, there are some settings where women receive social advantages undertaking traditionally masculine roles and men undertaking traditionally feminine roles (e.g., Evans, 1997; Rosette & Tost, 2010). Despite these changes to the distribution and social meaning of gender, many censuses and surveys did not have the tools to capture gender more holistically (Bates et al., 2022; Pao et al., 2025).

Many studies have attempted to ameliorate this gap between measurement and construct by creating new measures of gender *identity* (Budnick et al., 2025); however, studies of gender *expression* have been less common (Magliozzi et al., 2016). Gender expression has been defined as a dimension of gender that is “how an individual signals their gender to others through behavior and appearance” (Bates et al., 2022, p. 4). This has often been operationalized as masculinity and femininity, wherein gender identities of “man” and “woman” often correlate with higher “masculinity” and “femininity,” respectively (Westbrook & Saperstein, 2015). To capture dimensions of masculinity and femininity, scholars have begun using gradational gender scales, which ask survey respondents to mark their self-rated masculinity and femininity, separately, on unipolar scales (Alexander et al., 2021; Cassino, 2020; Garbarski, 2023; Magliozzi et al., 2016). These scales have been taken up on population-level surveys like the U.S. General Social Survey and European Social Survey and targeted sub-population surveys like the U.S. Trans Survey.

At present, gradational gender scales have been most popular in the U.S. and Europe; however, this tool to study gender expression may be most useful in some of the places that currently do not have them—such as Japan. Japan is a particularly interesting case due to its persistent gender stratification, despite continued women’s workplace participation (Brinton & Oh, 2019; Holbrow, 2025; Nemoto, 2016), and many institutional conditions (such as educational systems) that hinder gender equality (Uchikoshi et al., 2025). Nonetheless, Japan is also a place of innovation for gender measurement—with recent studies adapting new measures (such as a three-step gender measure) to a Japa-

nese local context (Hiramori & Kamano, 2020). Though there have been advances in measuring gender more holistically in Japan, to our knowledge, gradational gender scales have not been tested yet with a Japanese sample.

In the following, we translate gradational gender scales to Japanese and use them in two waves (October 2023 and February 2024) of an online longitudinal survey fielded by the Social Science Japan Data Archive (SSJDA) hosted at the Institute of Social Science, the University of Tokyo. We evaluate item nonresponse rates, response distributions, and within-respondent response (in)stability over time. These metrics highlight potential country-specific differences which pattern gradational gender scale responses and indicate the need for further research within Japan and other non-Western contexts when using these scales internationally.

Instrument

The instrument that we used was based off Magliozzi, Saperstein, and Westbrook's (2016) gradational gender scales. There are two sets of gradational scales proposed by Magliozzi et al.: the first are answered in response to the question, "In general, how do you see yourself?". The second are answered in response to the question, "In general, how do most people see you?". For both questions, there are two Likert scales, respectively labeled "Feminine" and "Masculine", each containing 7-points (marked 0 to 6) with 0 being "not at all" and "very." These scales have been validated and further tested in English with a U.S. sample (Garbarski, 2023) and have been translated internationally on the European Social Survey, Round 11 data. Other studies have previously used these scales to indicate differences by other sociodemographic characteristics or social outcomes, like health and political attitudes, showing the utility of these measures to understanding gender more holistically (Choi & Merlo, 2021; Hart et al., 2019; Pao, 2023; Pao & Roundy, 2025).

For international comparability, one of the authors, a native Japanese speaker, translated Magliozzi et al.'s instrument into Japanese—maintaining the scale points, question ordering, and format of the original English scales. As with U.S. and European precedents, all respondents, regardless of gender identity, would see both masculinity and femininity scales. Importantly, this translation of gradational gender scales was exploratory so diverged from some other known best practices, such as using double translation and team discussion processes (Behr & Shishido, 2016; Harkness, 2003). This translation used for gradational gender scales was the one that was approved by SSJDA for greater respondent clarity, but its slight divergences from standard gradational gender scales (namely, through the inclusion of the words "masculine" and "feminine" in the question prompt) may itself cause differences in response patterns. The translated version of the gradational gender scales can be seen in Figure 1. Note that these two questions were asked consecutively. Responses stemming from the first two scales are referred to as

“self-rated femininity” and “self-rated masculinity,” since respondents are rating how they see themselves. Responses stemming from the latter two scales are called “reflected appraisal femininity” and “reflected appraisal masculinity,” since they are what respondents believe others see. Though we believe that these slight deviations in translation from the original scales increased clarity for native Japanese speakers, we encourage future studies that use different translation processes, and possibly other translations, to replicate our findings.

Figure 1

Gradational Gender Scales Translated Into Japanese

一般的にいて、あなたは自分のことをどれだけ男らしい（女らしい）と思いますか。以下から当てはまるものを選んでください。

| | 全くそう思わない | | | | | | とてもそう思う |
|------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 男らしい | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 女らしい | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

一般的にいて、あなたは周囲からどれだけ男らしい（女らしい）と思われていると思いますか。以下から当てはまるものを選んでください。

| | 全くそう思わない | | | | | | とてもそう思う |
|------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 男らしい | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 女らしい | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Note. An English back-translation for the first instrument is, “Generally speaking, how masculine (feminine) do you consider yourself to be? Please choose the option that best applies,” and for the second instrument, “Generally speaking, how masculine (feminine) do you think you are perceived to be by others? Please choose the option that best applies.”

Data and Scale Development

We draw on data from an online longitudinal survey conducted by the SSJDA at the Institute of Social Science, the University of Tokyo. These data and codebooks are available via application from the SSJDA website at the Institute of Social Science: <https://ssjda.iss.u-tokyo.ac.jp/Direct/>. All code used to analyze these data are available via the Open Science Foundation: <https://osf.io/t7s98>.

The Wave 1 of the SSJDA Panel survey was conducted in February and March 2021, targeting men and women aged 20 to 39 residing in Japan as of the end of December 2020. Participants were selected using a stratified two-stage random sampling method, where the country was first divided into 11 regions, and then each region was further stratified into five categories based on population size. The number of survey locations allocated to each stratum was proportional to its population size. After a total of 100 locations were randomly selected, 50 individuals from each location were randomly chosen based on the Basic Resident Register (*Jūmin Kihon Daichō*). Invitation letters were mailed to 5,000 respondents, and responses were collected via the web using an online

survey service called Lime Survey. The number of valid responses to the SSJDA Panel Wave 1 was 1,329, resulting in a response rate of 26.6%. Respondents who completed the survey received a 500 Japanese yen gift card. In February 2022 when Wave 3 of the survey was fielded, additional survey participants were newly recruited. The target population consisted of men and women aged 21 to 40 residing in Japan as of the end of December 2021 to make their birth years equivalent to the Wave 1 participants. Using the same sampling method with an increased sample size (6,600), the number of valid respondents was 1,576, resulting in a response rate of 23.9%.

The authors received permission to field questions in the SSJDA Panel after applying through a call for proposals. The questions proposed by the authors were fielded in the SSJDA Panel in Waves 6 and 7, October 2023 and February 2024 respectively. Survey invitation letters were sent to 2,152 participants for the Wave 6 survey and to 2,112 participants for the Wave 7 survey after excluding individuals who had declined to participate in the follow-up survey. The number of valid responses was 1,528 and 1,603, respectively, resulting in response rates of 71.0% and 75.9%.

Two thirds of the respondents (1,024 respondents in Wave 6 and 1,091 respondents in Wave 7) were randomly assigned to the gradational gender scale questions. All respondents were given the option to skip questions. Respondents who saw the gradational gender scales were randomly assigned to see the measure at either the beginning or the end of the survey. The original purpose of this design was to examine the relationship between gradational scale measures and various life course outcomes (e.g., subjective health), and to test whether this relationship differs depending on whether the gradational gender questions are asked before or after the outcome questions. Because the allocation to the beginning and end conditions was random and balanced in both waves, we believe that the overall results—and the differences between waves 6 and 7—are not affected by our use of different survey arms. Out of the respondents who were asked the gradational gender responses, 992 respondents (1,046 respondents in Wave 7) remain after restricting the sample to those with complete information on key sociodemographic characteristics: gender,¹ educational attainment,² marital status, employment status, and individual annual income.³ These characteristics had nonresponse rates ranging from 0.2% for (marital status) to 1.9% for (annual income) in the raw data. The descriptive statistics of the Waves 6 and 7 are shown below, separately by gender.

1) A small number of respondents selected “other” or did not provide a response to the gender question. Due to the limited sample size, this study focuses on respondents who identified as either men or women.

2) Educational attainment was not asked in Waves 6 or 7. Therefore, we use information from earlier waves (Waves 1 to 4), prioritizing the most recent data. If educational information is missing in the latest available wave for a given respondent (e.g., Wave 4), we impute it from an earlier wave (e.g., Wave 3), if available.

3) While age is another key trait, the nonresponse rate of this variable is 0%.

Broadly, the sample also looks similar between survey waves.⁴ Educational attainment, employment status, and annual income distributions is also identical for both men and women. Compared to Wave 6, respondents in Wave 7 are slightly older than Wave 6 respondents and slightly more likely to be married. This reflects the longitudinal nature of the survey.

Table 1*Descriptive Statistics of SSJDA Sample by Sample Wave*

| | Wave 6 | | Wave 7 | |
|---------------------------------|-----------|-----------|-----------|-----------|
| | Men | Women | Men | Women |
| Age | | | | |
| 29 years old or less | 85 (22%) | 167 (28%) | 78 (18%) | 125 (21%) |
| 30–34 years old | 85 (22%) | 137 (23%) | 94 (21%) | 137 (23%) |
| 35–39 years old | 124 (31%) | 144 (24%) | 136 (31%) | 145 (24%) |
| 40 years old or over | 100 (25%) | 150 (25%) | 134 (30%) | 197 (33%) |
| Educational attainment | | | | |
| BA+ | 261 (66%) | 270 (45%) | 276 (62%) | 279 (46%) |
| non-BA | 133 (34%) | 328 (55%) | 166 (38%) | 325 (54%) |
| Employment status | | | | |
| Employed | 356 (90%) | 515 (86%) | 410 (93%) | 512 (85%) |
| Non-employed | 38 (9.6%) | 83 (14%) | 32 (7.2%) | 92 (15%) |
| Individual annual income | | | | |
| Income <= 5M JPY | 253 (64%) | 545 (91%) | 277 (63%) | 536 (89%) |
| Income > 5M JPY | 141 (36%) | 53 (8.9%) | 165 (37%) | 68 (11%) |
| Marital status | | | | |
| Married | 187 (47%) | 290 (48%) | 224 (51%) | 311 (51%) |
| Non-married | 207 (53%) | 308 (52%) | 218 (49%) | 293 (49%) |
| <i>n</i> | 394 | 598 | 442 | 604 |

Quality Criteria

The response rate on gradational gender scales from the overall sample was relatively low across the overall sample: 91.0% (92.7%) for self-rated femininity in Wave 6 (Wave 7); 86.6% (94.2%) for self-rated masculinity in Wave 6 (Wave 7); 90.4% (92.3%) for reflected appraisal femininity in Wave 6 (Wave 7); and 86.4% (93.8%) for reflected appraisal mascu-

4) Of the 1,024 respondents asked to complete the gradational gender scale questions in Wave 6, 8.2% did not participate in the Wave 7 survey. Thus, while the attrition rate is not negligible, it remains relatively small.

linity in Wave 6 (Wave 7). This is notable compared to the relatively higher response rates for other questions in the survey, such as labor force participation with a 99.1% response rate and parenthood status with a 99.8% response rate, both in Wave 6. This compares to nonresponse rates of more sensitive characteristics like annual income in the raw data (98.1%). Although the response rates for the gradational gender scales are not high, they are noticeably higher in Wave 7 than in Wave 6. This pattern may suggest that respondents become more comfortable with these items after repeated exposure and may better understand what the questions imply.

Item nonresponse was particularly high for the gender “atypical” scales—namely, femininity for men and masculinity for women. Table 2 shows item nonresponse for gradational gender scales by wave and by gender. Around 22 percent of the men (women) in the sample did not respond to the *self-rated* femininity (masculinity) in Wave 6, and around 23 percent of the men (women) in the sample did not respond to the *reflected appraisal* femininity (masculinity). In Wave 7, nonresponse decreased for men and women for both self-rated and reflected appraisal scales, but the decrease in nonresponse was steeper for women than for men. By Wave 7, 17 percent of men (10 percent of women) did not respond to their “gender-atypical” self-rated scale, and 18 percent of men (11 percent of women) did not respond to their “gender-atypical” reflected appraisal scale.

Table 2

Item Nonresponse for Gradational Gender Scales (SSJDA)

| | Wave 6 | | | | Wave 7 | | | |
|---------------------------------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|
| | Men | | Women | | Men | | Women | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Self-rated femininity | .22 | 0.42 | .00 | 0.06 | .17 | .38 | .00 | 0.00 |
| Self-rated masculinity | .00 | 0.05 | .22 | 0.42 | .00 | .00 | .10 | 0.30 |
| Reflected appraisal femininity | .23 | 0.42 | .01 | 0.09 | .18 | .39 | .00 | 0.04 |
| Reflected appraisal masculinity | .00 | 0.00 | .23 | 0.42 | .00 | .00 | .11 | 0.31 |

Note. $n = 394$ (men) and 598 (women) for Wave 6, 442 (men) and 604 (women) for Wave 7.

Notably, those who did and did not respond to the gradational gender scales varied by other sociodemographic characteristics. Table 3 uses Wave 6 data to show differences in responders and nonresponders. Men who responded to all gradational scales were 76.4% of the sample; men who responded to masculinity scales only were 21.1% of the sample; and other combinations (femininity only or nonresponse on both scales) were 2.5% of the sample. Women who responded to all gradational scales were 75.6% of the sample; women who responded to femininity scales only were 20.4% of the sample; and other combinations of scale nonresponse were 4.0% of the sample. Men (women) who

responded to all gradational gender scales were substantially different than those who replied to only masculinity (femininity) scales.

Table 3

Balance Table of Gradational Gender (Non)Respondents in Wave 6 (SSJDA)

| | Men | | | Women | | |
|---------------------------------|------------|-------------------------|--------------------------------|------------|------------------------|--------------------------------|
| | All scales | Masculinity scales only | No scales or other combination | All scales | Femininity scales only | No scales or other combination |
| All | 76.4% | 21.1% | 2.5% | 75.6% | 20.4% | 4.0% |
| Age groups | | | | | | |
| 29 years old or less | 84.7% | 14.1% | 1.2% | 82.6% | 13.8% | 3.6% |
| 30–34 years old | 82.4% | 14.1% | 3.5% | 79.6% | 18.2% | 2.2% |
| 35–39 years old | 79.0% | 20.2% | 0.8% | 70.1% | 25.0% | 4.9% |
| 40 years old or over | 61.0% | 34.0% | 5.0% | 69.3% | 25.3% | 5.3% |
| Educational attainment | | | | | | |
| BA+ | 78.9% | 19.2% | 1.9% | 82.2% | 14.8% | 3.0% |
| non-BA | 71.4% | 24.8% | 3.8% | 70.1% | 25.0% | 4.9% |
| Employment status | | | | | | |
| Employed | 75.0% | 22.2% | 2.8% | 76.3% | 19.2% | 4.5% |
| Non-employed | 89.5% | 10.5% | 0.0% | 71.1% | 27.7% | 1.2% |
| Individual annual income | | | | | | |
| Income ≤ 5M | 75.5% | 22.5% | 2.0% | 75.2% | 20.6% | 4.2% |
| Income > 5M | 78.0% | 18.4% | 3.5% | 79.2% | 18.9% | 1.9% |
| Marital status | | | | | | |
| Married | 71.1% | 25.1% | 3.7% | 72.4% | 22.8% | 4.8% |
| Non-married | 81.2% | 17.4% | 1.4% | 78.6% | 18.2% | 3.2% |

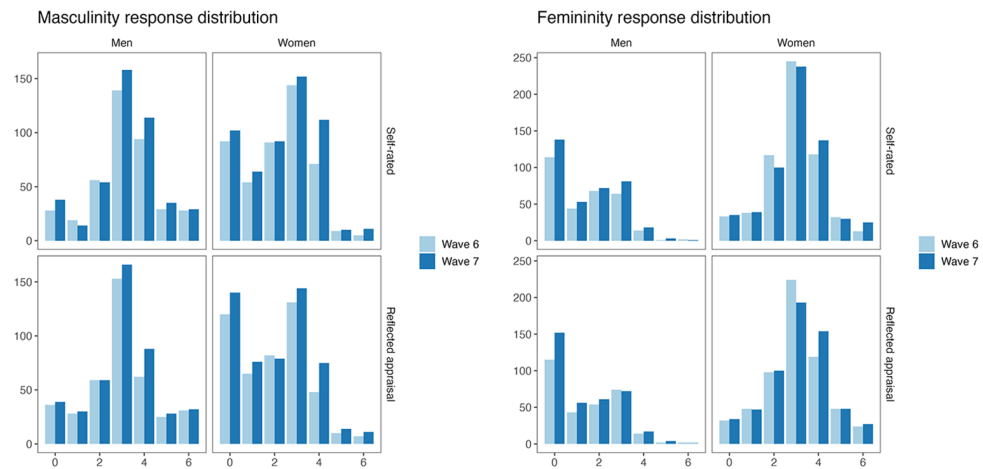
Among both women and men, younger respondents are more likely to answer all scales, while older respondents are less likely to answer gender “atypical” scales. Similarly, college-educated men and women are more likely than non-graduates to answer all scales. Among men, non-employed respondents are more likely to answer all scales than their employed counterparts, whereas the opposite pattern is observed among women. This may reflect age differences, as non-employed men are more likely to be younger and still in school, while non-employed women may have left the workforce due to family-related events such as marriage or childbearing. Higher income is also associated with greater likelihood of responding to all scales, although this association is less pronounced than that for educational attainment. Finally, non-married respondents are more likely to answer all scales than married respondents. This pattern may be confounded by age.

Among responders, the distribution of femininity and masculinity is displayed in [Figure 2](#). In the SSJDA sample the modal responses of gender-typical self-rated and

reflected-appraisal scales are 3, the scale midpoint. For the gender-*atypical* scales, the distribution of responses for men and women are more dispersed—particularly for women’s description of their masculinity. Notably, this response distribution is very different from existing patterns seen in other country contexts. Appendix Table 1 displays the median masculinity and femininity response for men and women in the 2024 General Social Survey (GSS) in the U.S. and the 2023 European Social Survey (ESS) for countries which included the gradational gender scales. This table shows that across the U.S. and European contexts for adults in the same age range as the SSJDA (i.e., 20–39), the median response for self-rated gender-typical scales was generally a 5 or 6. Though the 2023 ESS did not include reflected appraisal measures, the 2024 GSS shows that U.S. respondents’ median gender-typical reflected appraisal was also a 6.

Figure 2

Response Distribution From Gradational Scales by Gender (SSJDA)



Note. Responses were scaled from 0 (not at all) to 6 (very much).

There are several methodological and theoretical reasons we may see this middle-of-the-scale response distribution in our Japanese sample. From a methodological standpoint, differences between the GSS and the SSJDA sample may be partially explained by the lack of follow-up question that alerted respondents of a missing response. The 2024 GSS used this follow-up prompt, which reduced gender-atypical nonresponse, whereas the SSJDA did not.⁵ In Appendix Figure 1, we replicated our findings to replace all item nonresponse of sex-atypical scales with a zero to conservatively show what the hypothetical sex-atypical distribution could have been with prompting. In general, we see changes to the distributions of masculinity for women and femininity for men:

when these nonresponse cases are added to the group coded as zero, this value (0) becomes the median masculinity score for women, as opposed to the midpoint value (3). This may indicate that the middle-of-the-scale responses may be an artifact of different nonresponse patterns in this Japanese sample. Nonetheless, we may also see unimodal (instead of bimodal) SSJDA response distribution for other known differences in survey response patterns among Japanese respondents, such as higher midpoint responses on Likert scales—particularly among careless respondents (Chen et al., 1995; Masuda et al., 2017) or inappropriate translation and differences in cultural context (Behr & Shishido, 2016). The differences in response pattern may also arise due to the nature of the younger adult sampling frame in the SSJDA, which may be more likely to be gender nonconforming than in a country-wide samples (Ui & Matsui, 2008); nonetheless, this is unlikely to be the only driver, since younger birth cohorts have also been shown to have gender role attitudes similar to many of the older birth cohorts in Japan (Piotrowski et al., 2019). Though our data cannot speak to the mechanisms driving these differences in gender expression distributions, we believe this is an important area for future research.

We then evaluated response stability *within* respondents who replied to scales in both waves and differences in nonresponse between waves. Figure 3 shows the distribution of the difference in Wave 6 and Wave 7 responses among respondents who replied to the same scale in both waves. Figure 3 shows that the distributions are relatively normal, centered around 0—indicating no change in identification between waves. Despite the modal response being 0 for all combinations of masculinity and femininity changes for men and women, Figure 3 nonetheless still shows a degree of fluidity for how the same respondents may approach these gradational gender scales across time.

To show this fluidity more clearly, Table 4 shows nonresponse rates for men and women separately within each wave. We find that there is substantial variation *within* respondents over time. The majority of men who responded with a 0 for their self-rated and reflected appraisal femininity in Wave 6 would give the same response in Wave 7. However, this is not true of women who initially gave a 0 on their gender-atypical scales: while the plurality of women gave a 0 again in Wave 7, we see more variation across response categories, which may reflect changes either substantively in identification over time or a methodological artefact. For instance, response instability may indicate that women may report (and experience) higher degrees of fluidity, instability, or ambivalence in gender expression. This may also (or instead) indicate ambiguity in the measure itself, due to (for instance) the vagueness of the response scale where only the end points at 0 and 6 are labelled. Nonetheless, many men and women who provided middle-of-the-scale responses in one wave would also report a middle-of-the-scale response in another wave

5) GSS methodological statements say respondents who attempted to skip the question were provided a “gentle ‘nudge’” to encourage responses. For more information, see the codebook here: <https://gss.norc.umd.edu/content/dam/gss/get-documentation/pdf/codebook/GSS%202024%20Codebook%20R2.pdf>

—indicating that the scales may be capturing variability over time, but may have some degree of central tendency. These are novel findings, and we hope future studies may be able to replicate these findings in other Japanese samples and add more knowledge to what may drive changes over time.

Figure 3

Differences in Gradational Gender Responses Across Waves (SSJDA)

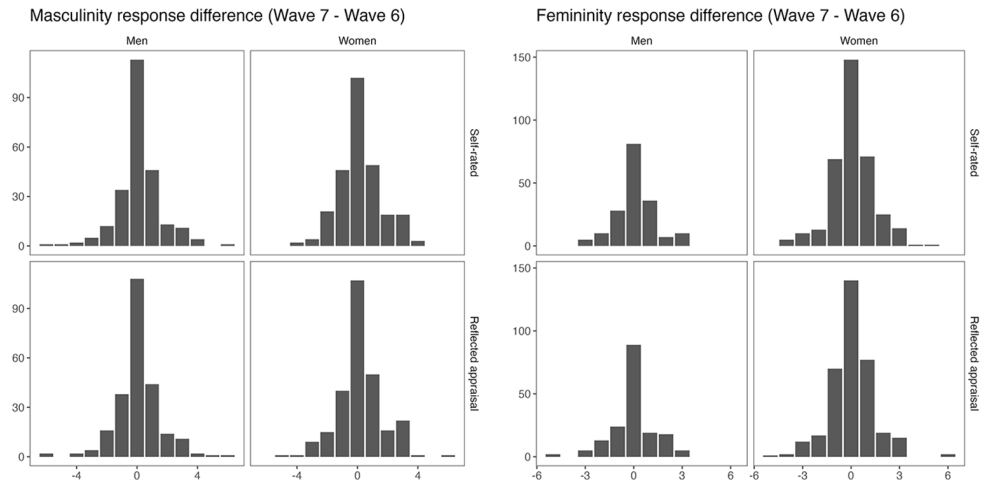


Table 4

Changes in the Distribution of Gender-Atypical Scale Scores Within Individuals Over Time

| Self-rated masculinity for women | | | | | | | | | | |
|----------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA | Total | n |
| 0 | 43.6% | 9.1% | 10.9% | 27.3% | 3.6% | 0.0% | 0.0% | 5.5% | 100.0% | 55 |
| 1 | 36.7% | 26.7% | 16.7% | 6.7% | 10.0% | 0.0% | 0.0% | 3.3% | 100.0% | 30 |
| 2 | 11.8% | 13.7% | 25.5% | 33.3% | 11.8% | 2.0% | 2.0% | 0.0% | 100.0% | 51 |
| 3 | 4.5% | 8.0% | 18.2% | 38.6% | 21.6% | 1.1% | 0.0% | 8.0% | 100.0% | 88 |
| 4 | 2.1% | 0.0% | 14.6% | 20.8% | 43.8% | 6.3% | 8.3% | 4.2% | 100.0% | 48 |
| 5 | 0.0% | 0.0% | 0.0% | 0.0% | 66.7% | 33.3% | 0.0% | 0.0% | 100.0% | 3 |
| 6 | 0.0% | 0.0% | 20.0% | 20.0% | 20.0% | 0.0% | 40.0% | 0.0% | 100.0% | 5 |
| NA | 20.5% | 7.2% | 10.8% | 20.5% | 12.0% | 2.4% | 1.2% | 25.3% | 100.0% | 83 |
| Total | 17.4% | 9.1% | 15.7% | 26.4% | 17.6% | 2.2% | 2.2% | 9.4% | 100.0% | 363 |

| Self-rated femininity for men | | | | | | | | | | |
|-------------------------------|-------|-------|-------|--------|-------|--------|-------|-------|--------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA | Total | n |
| 0 | 56.9% | 11.1% | 2.8% | 12.5% | 0.0% | 0.0% | 0.0% | 16.7% | 100.0% | 72 |
| 1 | 23.3% | 23.3% | 43.3% | 6.7% | 3.3% | 0.0% | 0.0% | 0.0% | 100.0% | 30 |
| 2 | 12.2% | 12.2% | 36.6% | 29.3% | 4.9% | 0.0% | 0.0% | 4.9% | 100.0% | 41 |
| 3 | 7.7% | 10.3% | 30.8% | 38.5% | 7.7% | 0.0% | 0.0% | 5.1% | 100.0% | 39 |
| 4 | 0.0% | 11.1% | 11.1% | 44.4% | 22.2% | 0.0% | 11.1% | 0.0% | 100.0% | 9 |
| 5 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 100.0% | 0.0% | 0.0% | 100.0% | 1 |
| 6 | 0.0% | 0.0% | 0.0% | 100.0% | 0.0% | 0.0% | 0.0% | 0.0% | 100.0% | 1 |
| NA | 30.0% | 14.0% | 4.0% | 10.0% | 0.0% | 0.0% | 0.0% | 42.0% | 100.0% | 50 |
| Total | 29.2% | 13.2% | 18.5% | 19.8% | 3.3% | 0.4% | 0.4% | 15.2% | 100.0% | 243 |

| Reflected appraisal masculinity for women | | | | | | | | | | |
|---|-------|-------|-------|-------|-------|-------|-------|-------|--------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA | Total | n |
| 0 | 49.3% | 14.1% | 8.5% | 21.1% | 0.0% | 0.0% | 1.4% | 5.6% | 100.0% | 71 |
| 1 | 23.7% | 28.9% | 13.2% | 15.8% | 10.5% | 0.0% | 0.0% | 7.9% | 100.0% | 38 |
| 2 | 15.7% | 19.6% | 21.6% | 31.4% | 2.0% | 0.0% | 2.0% | 7.8% | 100.0% | 51 |
| 3 | 9.9% | 3.7% | 17.3% | 43.2% | 17.3% | 2.5% | 3.7% | 2.5% | 100.0% | 81 |
| 4 | 0.0% | 0.0% | 10.3% | 20.7% | 41.4% | 17.2% | 3.4% | 6.9% | 100.0% | 29 |
| 5 | 33.3% | 0.0% | 0.0% | 33.3% | 33.3% | 0.0% | 0.0% | 0.0% | 100.0% | 3 |
| 6 | 0.0% | 0.0% | 16.7% | 16.7% | 0.0% | 16.7% | 50.0% | 0.0% | 100.0% | 6 |
| NA | 26.2% | 10.7% | 6.0% | 16.7% | 8.3% | 3.6% | 1.2% | 27.4% | 100.0% | 84 |
| Total | 22.9% | 11.8% | 12.4% | 25.9% | 10.7% | 3.0% | 2.8% | 10.5% | 100.0% | 363 |

| Reflected appraisal femininity for men | | | | | | | | | | |
|--|--------|-------|-------|-------|-------|------|------|-------|--------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA | Total | n |
| 0 | 63.2% | 6.6% | 9.2% | 6.6% | 0.0% | 0.0% | 0.0% | 14.5% | 100.0% | 76 |
| 1 | 13.8% | 31.0% | 17.2% | 31.0% | 0.0% | 0.0% | 0.0% | 6.9% | 100.0% | 29 |
| 2 | 18.8% | 15.6% | 43.8% | 15.6% | 3.1% | 0.0% | 0.0% | 3.1% | 100.0% | 32 |
| 3 | 11.4% | 13.6% | 25.0% | 31.8% | 9.1% | 2.3% | 0.0% | 6.8% | 100.0% | 44 |
| 4 | 0.0% | 0.0% | 11.1% | 44.4% | 44.4% | 0.0% | 0.0% | 0.0% | 100.0% | 9 |
| 5 | 100.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 100.0% | 2 |
| NA | 33.3% | 7.8% | 3.9% | 9.8% | 0.0% | 0.0% | 0.0% | 45.1% | 100.0% | 51 |
| Total | 33.7% | 11.9% | 16.5% | 17.3% | 3.7% | 0.4% | 0.0% | 16.5% | 100.0% | 243 |

Note. Rows are Wave 6 responses, and columns are Wave 7 responses within-respondent.

Conclusion

In this study, we perform a preliminary evaluation of response distributions and non-response patterns for gradational gender scales using a young adult Japanese sample. Overall, our study shows how gradational gender measures may appear differently in a different, non-Western context. Using a sample of adults in a longitudinal study (SSJDA),

we show that there is a relatively high degree of nonresponse to gradational gender scales, and this may be particularly elevated when respondents are first exposed to the scales. Among those that do respond to the gradational scales, the majority respond to both gender-typical and gender-atypical masculinity and femininity scales, followed by respondents who just respond to their gender-typical (i.e., masculinity for men, femininity for women) scales.

Responders and nonresponders varied on a variety of sociodemographic characteristics—indicating differential nonresponse for both men and women. Younger respondents, college-educated respondents, and higher income respondents are more likely to answer all scales. The response distribution for responders indicates a modal response of 3 (the gradational scale midpoint) for gender-typical scales. The responses for gender-atypical scales tended to be more uniformly dispersed among responses at the lower-end of the scale among responders. This is different from other samples, like those in the U.S., which tend to have more extreme responses (e.g., 0s and 6s, the minimum and maximum of the scales).

Finally, due to the panel nature of the data, we could evaluate response stability over time (i.e., between waves). Though most respondents tended to respond identically on each scale (i.e., a change of 0 from wave to wave), the distributions appear quasi-normal, indicating a degree of fluctuation for the same respondents at different points in time. This is a novel intervention, since many surveys using gradational gender scales only collect cross-sectional data. Our data allows us to evaluate within-respondent variation, pointing to the need to study the cause of variability: asking gradational gender scales at multiple points in time for the same respondent may, on the one hand, reveal substantive changes in gender expression, or, on the other hand, indicate response instability due to methodological challenges in the measure itself. In sum, we hope that this translation of gender scales to a different population, which experiences a different gendered context, shows the need for further research on the use of gradational gender scales in Japan and other non-Western contexts to evaluate their comparability internationally.

Funding: This study was fielded and funded by the Social Science Japan Data Archive (SSJDA) in a call for proposals to add additional questionnaire items in August 2023. SSJDA does not bear any responsibility for the analyses or interpretations presented here.

Acknowledgments: We would like to thank those at the SSJDA for their help in acquiring these data. We used the AI tool, Claude (Opus 4.5) to assist with formatting our revised manuscript to PsychOpen GOLD standards.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: The secondary data analyzed here were collected under a protocol with ethical approval at the institution of the original data collector (i.e., the Institute of Social Science at the University of Tokyo). Because we conducted analysis of an anonymized secondary data set, further ethical approval from the authors' institutions was not required.

Data Availability: Since we are not able to share the SSJDA data directly, we have been asked to show which call our data were submitted to (see [SSJDA Panel Survey, n.d.](https://ssjda.iss.u-tokyo.ac.jp/Direct/)). Individuals who wish to access the data will have to apply to the data repository through SSJDA (<https://ssjda.iss.u-tokyo.ac.jp/Direct/>). All data are available from the SSJDA directly after application approval.

Supplementary Materials

Data and questionnaires are available for application from the SSJDA website (<https://ssjda.iss.u-tokyo.ac.jp/Direct/>; for the call for data, see [SSJDA Panel Survey, n.d.](#)). Replication files and code are available from OSF (see [Pao & Uchikoshi, 2026](#)).

Index of Supplementary Materials

SSJDA Panel Survey. (n.d.). *Call for questionnaire items for the SSJDA panel survey* [Call for data].

Institute of Social Science, the University of Tokyo.

<https://csrda.iss.u-tokyo.ac.jp/english/ssjdap/call-for-proposals.html>

Pao, C., & Uchikoshi, F. (2026). *An adaptation of gradational gender scales in a Japanese sample:*

Nonresponse rates and response distributions for self-rated masculinity and femininity

[Replication files, code]. OSF. <https://osf.io/t7s98>

References

- Alexander, A. C., Bolzendahl, C., & Wängnerud, L. (2021). Beyond the binary: New approaches to measuring gender in political science research. *European Journal of Politics and Gender*, 4(1), 7–9. <https://doi.org/10.1332/251510820X16067519822351>
- Bates, N., Chin, M., & Becker, T. (Eds.). (2022). *Measuring sex, gender identity, and sexual orientation*. National Academies Press. <https://doi.org/10.17226/26424>

- Behr, D., & Shishido, K. (2016). The translation of measurement instruments for cross-cultural surveys. In C. Wolf, D. Joye, T. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 269–287). SAGE Publications. <https://doi.org/10.4135/9781473957893.n19>
- Brinton, M. C., & Oh, E. (2019). Babies, work, or both? Highly educated women's employment and fertility in East Asia. *American Journal of Sociology*, *125*(1), 105–140. <https://doi.org/10.1086/704369>
- Budnick, J., Pao, C., & Velasco, K. (2025). Queer data for sociologists of sexualities: Introducing SOGIESC measurement and methods during political suppression. *Sex & Sexualities*, *1*(1), 147–136. <https://doi.org/10.1177/3033371251329931>
- Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*, *40*(4), 519–531. <https://doi.org/10.2307/3207893>
- Cassino, D. (2020). Moving beyond sex: Measuring gender identity in telephone surveys. *Survey Practice*, *13*(1). <https://doi.org/10.29115/SP-2020-0009>
- Chen, C., Lee, S., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, *6*(3), 170–175. <https://doi.org/10.1111/j.1467-9280.1995.tb00327.x>
- Choi, J., & Merlo, A. V. (2021). Gender identification and the fear of crime: Do masculinity and femininity matter in reporting fear of crime? *Victims & Offenders*, *16*(1), 126–147. <https://doi.org/10.1080/15564886.2020.1787282>
- Connell, R. (2020). The social organization of masculinity. In C. R. McCann & S.-K. Kim (Eds.), *Feminist theory reader: Local and global perspectives* (5th ed., pp. 192–205). Routledge.
- Edwards, T. (2005). *Cultures of masculinity*. Routledge. <https://doi.org/10.4324/9780203005224>
- Evans, J. (1997). Men in nursing: Issues of gender segregation and hidden advantage. *Journal of Advanced Nursing*, *26*(2), 226–231. <https://doi.org/10.1046/j.1365-2648.1997.1997026226.x>
- Garbarski, D. (2023). The measurement of gender expression in survey research. *Social Science Research*, *110*, Article 102845. <https://doi.org/10.1016/j.ssresearch.2022.102845>
- Goffman, E. (1977). The arrangement between the sexes. *Theory and Society*, *4*(3), 301–331. <https://doi.org/10.1007/BF00206983>
- Hamilton, L. T., Armstrong, E. A., Seeley, J. L., & Armstrong, E. M. (2019). Hegemonic femininities and intersectional domination. *Sociological Theory*, *37*(4), 315–341. <https://doi.org/10.1177/0735275119888248>
- Harkness, J. (2003). Questionnaire translation. In J. Harkness, F. Van de Vijver, & P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 35–56). Wiley.
- Hart, C. G., Saperstein, A., Magliozzi, D., & Westbrook, L. (2019). Gender and health: Beyond binary categorical measurement. *Journal of Health and Social Behavior*, *60*(1), 101–118. <https://doi.org/10.1177/0022146519825749>
- Hiramori, D., & Kamano, S. (2020). Asking about sexual orientation and gender identity in social surveys in Japan: Findings from the Osaka City Residents' Survey and related preparatory studies. *Jinkō Mondai Kenkyū (Journal of Population Problems)*, *76*(4), 443–466. <https://www.ipss.go.jp/syoushika/bunken/data/pdf/20760402.pdf>

- Holbrow, H. J. (2025). *The future is foreign: Women and immigrants in corporate Japan*. Cornell University Press.
- Lagos, D. (2022). Has there been a transgender tipping point? Gender identification differences in U.S. cohorts born between 1935 and 2001. *AJS; American Journal of Sociology*, 128(1), 94–143. <https://doi.org/10.1086/719714>
- Magliozzi, D., Saperstein, A., & Westbrook, L. (2016). Scaling up: Representing gender diversity in survey research. *Socius: Sociological Research for a Dynamic World*, 2, 1–11. <https://doi.org/10.1177/2378023116664352>
- Masuda, S., Sakagami, T., Kawabata, H., Kijima, N., & Hoshino, T. (2017). Respondents with low motivation tend to choose middle category: Survey questions on happiness in Japan. *Behaviormetrika*, 44(2), 593–605. <https://doi.org/10.1007/s41237-017-0026-8>
- Nemoto, K. (2016). *Too few women at the top: The persistence of inequality in Japan*. Cornell University Press. <https://doi.org/10.7591/j.ctt1d2dn1q>
- Pao, C. (2023). Masculinity and femininity by racial identification: Racialized differences in responses to self-rated gender scales for cisgender men and women. *Socius: Sociological Research for a Dynamic World*, 9, 1–14. <https://doi.org/10.1177/23780231231186073>
- Pao, C., Donnelly Moran, K., Compton, D., Kaufman, G., & Dowling, J. A. (2025). The case for “other”: Measuring gender and sexual identity in survey research. *Sociology Compass*, 19(1), Article e70031. <https://doi.org/10.1111/soc4.70031>
- Pao, C., & Roundy, N. (2025). An introduction to masculinity and political attitudes. In M. L. McDermott & D. Cassino (Eds.), *Masculinity in American politics* (pp. 99–107). New York University Press. <https://doi.org/10.18574/nyu/9781479830725.003.0008>
- Piotrowski, M., Yoshida, A., Johnson, L., & Wolford, R. (2019). Gender role attitudes: An examination of cohort effects in Japan. *Journal of Marriage and the Family*, 81(4), 863–884. <https://doi.org/10.1111/jomf.12577>
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & Society*, 18(4), 510–531. <https://doi.org/10.1177/0891243204265269>
- Ridgeway, C. L., & Saperstein, A. (2024). Diversifying gender categories and the sex/gender system. *Annual Review of Sociology*, 50, 385–405. <https://doi.org/10.1146/annurev-soc-030222-035327>
- Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *The Journal of Applied Psychology*, 95(2), 221–235. <https://doi.org/10.1037/a0018204>
- Uchikoshi, F., Toyonaga, K., & Teramoto, E. (2025). Consequences of expanded vocationally oriented programs for gender segregation and inequality: The case of Japanese higher education. *Research in Social Stratification and Mobility*, 97, Article 101024. <https://doi.org/10.1016/j.rssm.2025.101024>
- Ui, M., & Matsui, Y. (2008). Japanese adults’ sex role attitudes and judgment criteria concerning gender equality: The diversity of gender egalitarianism. *Sex Roles*, 58(5), 412–422. <https://doi.org/10.1007/s11199-007-9346-6>

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

<https://doi.org/10.1177/0891243287001002002>

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>

Appendix

| Country | Female | | | | Male | | | |
|-------------|--------------------------|---------------------------|---------------------------------------|--|--------------------------|---------------------------|---------------------------------------|--|
| | Self-rated Femininity | Self-rated Masculinity | Refl. App. Femininity ^a | Refl. App. Masculinity ^a | Self-rated Femininity | Self-rated Masculinity | Refl. App. Femininity ^a | Refl. App. Masculinity ^a |
| Austria | 6 | 0 | — | — | 0 | 6 | — | — |
| Belgium | 5 | 1 | — | — | 0 | 5 | — | — |
| Bulgaria | 6 | 0 | — | — | 0 | 6 | — | — |
| Croatia | 6 | 0 | — | — | 0 | 6 | — | — |
| Cyprus | 6 | 0 | — | — | 0 | 6 | — | — |
| Finland | 4 | 2 | — | — | 1 | 4 | — | — |
| France | 5 | 1 | — | — | 0 | 6 | — | — |
| Germany | 6 | 0 | — | — | 0 | 5 | — | — |
| Greece | 6 | 0 | — | — | 0 | 6 | — | — |
| Hungary | 6 | 0 | — | — | 0 | 6 | — | — |
| Iceland | 5 | 1 | — | — | 1 | 5 | — | — |
| Ireland | 5 | 0 | — | — | 0 | 5 | — | — |
| Israel | 6 | 0 | — | — | 0 | 6 | — | — |
| Italy | 6 | 0 | — | — | 0 | 6 | — | — |
| Latvia | 6 | 1 | — | — | 0 | 6 | — | — |
| Lithuania | 6 | 0 | — | — | 0 | 6 | — | — |
| Montenegro | 6 | 0 | — | — | 0 | 6 | — | — |
| Netherlands | 5 | 1 | — | — | 0 | 5.5 | — | — |
| Norway | 5 | 2 | — | — | 1 | 5 | — | — |
| Poland | 6 | 0 | — | — | 0 | 6 | — | — |
| Portugal | 6 | 0 | — | — | 0 | 6 | — | — |
| Serbia | 5 | 0 | — | — | 0 | 6 | — | — |
| Slovakia | 6 | 0 | — | — | 0 | 6 | — | — |

Table A1

Median Gender Expression Response in the 2023 European Social Survey and 2024 General Social Survey

Measurement Instruments for the Social Sciences
 2026, Vol. 8, Article e18677
<https://doi.org/10.5964/miss.18677>

| Country | Female | | | | | | Male | | | | | |
|---------------------|------------|-------------|-------------------------|--------------------------|------------|-------------|------------|-------------|------------|-------------|-------------------------|--------------------------|
| | Self-rated | | Refl. App. | | Self-rated | | Refl. App. | | Self-rated | | Refl. App. | |
| | Femininity | Masculinity | Femininity ^a | Masculinity ^a | Femininity | Masculinity | Femininity | Masculinity | Femininity | Masculinity | Femininity ^a | Masculinity ^a |
| Slovenia | 6 | 1 | — | — | — | — | 0 | 6 | — | — | — | — |
| Spain | 5 | 0 | — | — | — | — | 0 | 6 | — | — | — | — |
| Sweden | 4 | 2 | — | — | — | — | 1 | 5 | — | — | — | — |
| Switzerland | 5 | 1 | — | — | — | — | 1 | 5 | — | — | — | — |
| United Kingdom | 5 | 1 | — | — | — | — | 0 | 5 | — | — | — | — |
| United States (GSS) | 5 | 1 | 6 | 1 | 1 | 1 | 0 | 5 | 0 | 0 | 5 | 5 |

Note. Data from the European Social Survey (ESS) (<https://www.europeansocialsurvey.org/>) and General Social Survey (GSS) (<https://gss.norc.umd.edu/>) are publicly available for downloading. The script used to produce these descriptive analyses are available from OSF (<https://osf.io/t7s98/>). The sample for the ESS was $n = 46,162$, removing those from the sample who did not report a sex and were between the ages of 20 and 39, to make the data more comparable to the SSJDA. The sample for the GSS was $n = 3,309$, removing those who did not report a sex and who were between the ages of 20 to 39. The median response was calculated among only those who reported a response for the scale (i.e., removing missing responses for that specific scale).

^a The ESS did not include data on reflected appraisal gradational gender responses.

Figure A1

Response Distribution Replacing All Item Nonresponse of Gender-Atypical Scales With a Zero

