Continuous Gender Identity and Economics

By ANNE ARDILA BRENØE, LEA HEURSEN, EVA RANEHILL, AND ROBERTO A. WEBER*

* Brenøe: University of Zurich, Schönberggasse 1, 8001 Zurich, Switzerland (anne.brenoe@econ.uzh.ch). Heursen: Humboldt University of Berlin, (lea.heursen@hu-berlin.de)., Germany (lea.heursen@hu-berlin.de). Ranehill: University of Gothenburg, Vasagatan 1, 41124 Gothenburg, Sweden (eva.ranehill@economics.gu.se). Weber: University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland (roberto.weber@econ.uzh.ch). This work was supported by the University Research Priority Program "URPP Equality of Opportunity" of the University of Zurich.

Over the past few decades, economists have become increasingly interested in the topic of gender. Part of this work documents gender gaps in basic economic preferences, such as risk tolerance (e.g., Niederle 2016), and how these gaps correlate with important outcomes, such as educational choice (e.g., Buser, Niederle and Oosterbeek 2014).

To date, economic research exploring gender preference gaps has, with a few prominent exceptions (e.g., Akerlof and Kranton 2000), focused on differences by *biological sex*—a binary classification as either a "man" or "woman"—rather than *gender identity*. This approach is sensible, as this binary classification is ubiquitous in datasets. At the same time, substantial research in gender studies argues for a conceptualization of *gender* as a manifestation of individual traits and behaviors, social and personal perceptions of identity, and agreement with or divergence from societal norms (e.g., Knaak, 2004; Westbrook & Saperstein, 2015). In this literature, gender is often measured through subjective assessments of masculinity, femininity, or possibly both (Bem, 1974; Magliozzi et al., 2016).

Whether this richer, non-binary and subjective conceptualization of gender is useful for economics is not immediately obvious. On one hand, using a subjectively constructed self-assessment as an explanatory variable complicates analysis due to potential noise and bias. On the other hand, substantial variation within biologically "male" and "female" samples in both behavioral tendencies like risk preferences (Nelson, 2015) and in economic outcomes (Goldin, 2016) suggests that a richer classification may capture important dimensions of individual heterogeneity useful for positive economics and for the design of more carefully targeted policy interventions. Moreover, increasing identification by individuals as non-binary and the growing inclusion of richer measures of gender identity in administrative datasets creates opportunities for using these measures in economic research.

Our research studies the value of incorporating richer, subjective, self-assessed notions of gender identity into economics. To be useful for economists, it is necessary, first, to have simple measures that can be widely employed in administrative datasets and, second, to document that such measures add value to understanding economic questions beyond the binary measure of (mostly) biological sex.

In a first step, we conducted a survey to validate a concise measure of continuous gender identity. We then provide preliminary evidence on the extent to which this measure adds explanatory power beyond that of a binary indicator for biological sex in understanding the propensity to exhibit important economic traits. This complements the richer study of this second question in Brenøe, et al. (2021).

I. Validating a Single-Item Measure of Continuous Gender Identity

We collected answers to various scales measuring gender identity used in gender research, along with a novel single-item question. We then investigate the degree to which variation in the former can be accounted for by the latter.

A. Methods

The survey sample consisted of 54 women and 46 men recruited from the University Registration Center for Study Participants at the University of Zurich. [1] Participants completing the online survey received a participation fee of CHF 10. The survey included five measures for non-binary gender, presented in fixed order as follows:

BSRI: In the 60-item Bem (1974) Sex Role Inventory, respondents rate themselves, for each item, on a 7-point Likert scale ranging from "never or almost never true" to "always or almost always true." Each item of the BSRI is a characteristic, coded as either feminine ("love children"), masculine ("defend my own beliefs") or neutral ("conscientious"). Hence, this scale measures masculinity and femininity as independent dimensions. A drawback of the BSRI is that the masculinity-femininity classifications are derived from somewhat dated gender stereotypes and may therefore measure the extent to which one conforms to these stereotypes and expectations rather than one's own sense of gender.

Two-Dimensional Scale: The two-dimensional scale by Magliozzi et al. (2016) measures both first-order ("how do you see yourself?") and third-order ("how do most

[1] We exclude one male participant who indicated having responded unreliably.

people see you?") masculinity and femininity on separate dimensions. For each dimension, participants indicate their response on 7-point Likert scales, ranging from "not at all" to "very" masculine or feminine, respectively.

Single Item Continuous Gender Identity (CGI): Our own scale measures first-order perceptions of gender identity ("Where would you put yourself on this scale?") by eliciting self-placement on a 7-point scale, ranging from "very masculine" to "very feminine." The main difference with the Magliozzi scale is measurement of masculinity and femininity in a single dimension. Following Magliozzi et al., we also elicited a measure of third-order beliefs ("Where would other people put you on this scale?").

OSRI: We searched for a modernized version of the BSRI, settling on the 20-item open-source Open Sex-Role Inventory (2019). Masculine items ("I like guns") or feminine items ("I have kept a personal journal") are rated on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree." While not a published instrument, De Roover and Vermunt (2019) compared masculinity and femininity across sexual orientations using a large dataset.

TMF: In the unidimensional 6-item Traditional Masculinity Femininity scale developed by Kachel et al. (2016) respondents rate themselves from "very masculine" to "very feminine" on a 7-point scale. The six individual questions address gender-role adoption ("I consider myself as..."), preference ("Ideally, I would like to be..."), and identity ("Traditionally, my interests/attitudes and beliefs/behavior/outer appearance would be considered as...").

We also asked a set of demographic questions. All participants reported a current gender that matched their (reported) biological sex at birth; henceforth, we use biological sex to distinguish between men and women.

B. Analysis and Choice of Gender Measure

To compare our measures, we standardize all scores. We score the BSRI following the test manual and get results similar to Bem (1981, p. 71). 37.0% (13.3%) of the women (men) in our data classify as feminine, 14.8% (37.9%) as masculine and 27.8% (24.4%) as androgynous. In the two-dimensional scale by Magliozzi et al. (2016), and in our unidimensional scale, we find high correlations of almost 0.9 between responses to first- and third-order questions. We also use principal component analysis to extract a measure of the underlying latent continuous gender identity from the seven existing scales (online Appendix Table A1).

Table 1 presents the correlations between the continuous gender measures, while Figure 1 illustrates these relationships. The corre-

TABLE 1 – CORRELATION MATRIX OF FEMININITY-MASCULINITY SCALES FROM ONLINE SURVEY

	Single Item CGI	BSRI Fem	BSRI Masc	Mag Fem	Mag Masc	OSRI Fem	OSRI Masc	TMF
BSRI Fem	**0.292**							
	(0.003)							
BSRI Masc	**-0.220**	-0.199						
	(0.029)	(0.049)						
Mag Fem	**0.903**	0.305	-0.165					
	(0.000)	(0.002)	(0.104)					
Mag Masc	**-0.913**	-0.223	0.270	-0.813				
	(0.000)	(0.027)	(0.007)	(0.000)				
OSRI Fem	**0.527**	0.390	-0.016	0.600	-0.467			
	(0.000)	(0.000)	(0.878)	(0.000)	(0.000)			
OSRI Masc	**-0.561**	-0.207	0.236	-0.531	0.533	-0.234		
	(0.000)	(0.040)	(0.019)	(0.000)	(0.000)	(0.020)		
TMF	**0.914**	0.329	-0.279	0.882	-0.845	0.568	-0.623	
	(0.000)	(0.001)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	
First Component	**0.918**	0.457	-0.333	0.915	-0.879	0.671	-0.683	0.948
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: p-values are reported in parentheses. Fem (Masc) refers to the femininity (masculinity) scores of the two-dimensional BSRI, OSRI and Magliozzi (Mag) scales. Single Item CGI is the score from our single question's unidimensional scale, ranging from "very masculine" to "very feminine"; TMF is an alternative unidimensional measure. First Component is the first component from a principal component analysis of the seven continuous scales (excluding our single item measure). *Source:* Brenøe, et al., 2021.

lations all have the appropriate sign, indicating reliability of the different measures of gender identity. The coefficients are smaller in magnitude for both dimensions of the BSRI than for other scales, but the correlations are nevertheless statistically significant. The correlations for the separate dimensions of the two-dimensional scales are negative, indicating that substantial variation on these scales might be unidimensional in nature. Our scale also correlates highly with the first component extracted from principal component analysis.

Our unidimensional scale demonstrates both variation and overlap in reported gender identity between men and women (Figure A1 in the online Appendix). Participants used the full 7-point scale to self-identify, with men (women) using the five most masculine (feminine) categorical responses. Thus, not all women (men) consider themselves equally feminine (masculine), but instead perceive their gender identity heterogeneously.

We conclude that (i) our single-item measure captures a substantial share of the variation in gender identity measured by other scales and (ii) there is substantial variation in continuous gender identity. We next provide preliminary evidence on the relationship between gender identity and economic preferences.

II. Preliminary Evidence on Continuous Gender and Economic Decision Making

We implemented a computerized experiment with 120 participants. In this (pre-registered; doi.org/10.17605/OSF.IO/PHYT6) pilot experiment, we elicited our measure of continuous gender identity and biological sex

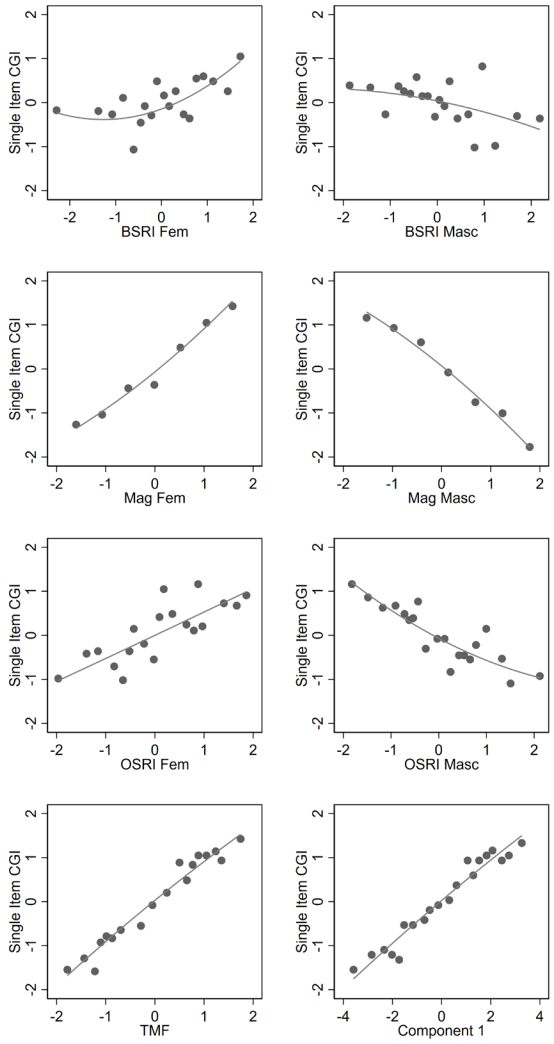


FIGURE 1. BINNED SCATTER PLOTS DEPICTING THE RELATIONSHIP BETWEEN THE DIFFERENT CG SCALES AND OUR SINGLE ITEM

Note: Fem (Masc) refers to the femininity (masculinity) scores of each of the BSRI, OSRI, and Mag (Magliozzi) scales. Single Item represents our single question, ranging from "very masculine" to "very feminine"; TMF has a similar scale. First Component is the first component from a principal component analysis of the seven continuous gender scales (excluding our single item measure).

along with four measures of economic preferences for which previous research has documented robust gender gaps: risk taking, competitiveness, preferences for equality over efficiency, and overconfidence (Fisman, Kariv, & Markovits, 2007; Gillen, Snowberg, & Yariv, 2015; Gneezy & Potters, 1997; Niederle & Vesterlund, 2007). We followed the experimental procedures used in earlier research as closely as possible. To control for possible measurement error in non-binary gender identity, we followed the procedure proposed by Gillen et al. (2019) and collected a second set of responses to the gender identity scale in an online follow-up survey three weeks after the laboratory experiment.

Table 2 presents results from regression analysis of the incentivized behavioral measures on biological sex and continuous gender identity. Panel A presents OLS regressions of each preference measure on biological sex, finding that women are less risk seeking, competitive, overconfident, and prioritize efficiency less than equality compared to men, consistent with previous research. Panel B repeats this exercise using self-reported gender identity instead of biological sex, finding similar results—gender identity predicts all four economic preference measures. However, as shown in Panel C, including both biological sex and gender identity as explanatory variables, the latter displays substantial explanatory power beyond that of biological sex only for risk attitudes—in this case, the continuous measure of gender identity is statistically significant while biological sex is not. While these are only

TABLE 2 – REGRESSION OF INCENTIVIZED BEHAVIORAL MEASURES ON BIOLOGICAL SEX AND CONTINUOUS GENDER IDENTITY

	Risk	Competitiveness	Overconfidence	Equality vs. Efficiency (ρ)
Panel A. Biological sex				
Biological Female	-0.674	-0.453	-0.514	-1.280
	(0.179)	(0.084)	(0.176)	(0.543)
Panel B. Gender identity				
Single Item CGI (ORIV)	-0.435	-0.208	-0.246	-0.558
	(0.092)	(0.053)	(0.110)	(0.332)
Panel C. Gender identity and biological sex				
Biological Female	-0.125	-0.418	-0.437	-1.304
	(0.314)	(0.155)	(0.262)	(0.858)
Single Item CGI (ORIV)	-0.380	-0.024	-0.053	0.017
	(0.175)	(0.094)	(0.169)	(0.533)
Observations	120	120	120	114
Mean of Dependent Variable	-0.000	0.533	0.000	5.325

Notes: Robust standard errors in parentheses. The estimates in each column and panel come from a separate regression. Panel A regresses the four incentivized behavioral measures [risk (standardized, mean zero, standard deviation one), competitiveness (binary), overconfidence (standardized) and ρ (deciles, sample with few GARP violations in allocation choices following Fisman et al. 2007)] on biological sex and uses HC3 standard errors. Panels B and C instrument our standardized single item CGI question (11-point scale) elicited in the lab with a similar question in the follow up survey (12-point scale). All regressions control for session fixed effects and a constant. *Source:* Brenøe, et al., 2021.

preliminary findings in the early stage of a larger data collection, they suggest added explanatory power from incorporating self-reported measures of continuous gender identity, mainly in the domain of risk.

III. Conclusion

While much work remains to determine whether continuous measures of self-reported gender identity can be useful for understanding economic behavior and valuable for policymaking, this paper reports two important steps in this direction. First, we identify a short single-item question that correlates with richer measures used in gender research, thereby providing a simple measurement instrument. Second, we provide preliminary evidence on the usefulness of this measure for better accounting for the relationship between gender and a set of important economic preferences and beliefs.

REFERENCES

Akerlof, George A. and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics* 115(3): 715–753.

Brenøe, Anne. A., Lea Heursen, Eva Ranehill and Roberto Weber. 2021. **"**Gender Identity and Economic Decision Making." Unpublished.

Bem, Sandra L. 1974. "The Measurement of Psychological Androgyny." *Journal of Consulting and Clinical Psychology* 42(2): 155–62.

Bem, Sandra L. 1981. *Bem Sex-Role Inventory: Professional Manual.* Palo Alto, CA: Consulting Psychologists Press.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics*, 129(3): 1409–1447.

De Roover, Kim, and Jeroen K. Vermunt. 2019. "On the Exploratory Road to Unraveling Factor Loading Non-Invariance: A New Multigroup Rotation Approach." *Structural Equation Modeling: A Multidisciplinary Journal* 26(6): 905–23.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review* 97(5): 1858–76.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy* 127(4): 1826-1863.

Gneezy, Uri, and Jan Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *The Quarterly Journal of Economics* 112(2): 631–45.

Goldin, Claudia, Lawrence F. Katz. 2016. "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family Friendly Occupation." *Journal of Labor Economics* 34(3): 705-745.

Kachel, Sven, Melanie C. Steffens, and Claudia Niedlich. 2016. "Traditional Masculinity and Femininity: Validation of a New Scale Assessing Gender Roles." *Frontiers in Psychology* 7: 956.

Knaak, Stephanie. 2004. "On the Reconceptualizing of Gender: Implications for Research Design." *Sociological Inquiry* 74(3): 302–17.

Magliozzi, Devon, Aliya Saperstein, and Laurel Westbrook. 2016. "Scaling Up: Representing Gender Diversity in Survey Research." *Socius: Sociological Research for a Dynamic World* 2.

Nelson, Julie A. 2015. "Are Women Really More Risk-Averse than Men? A Re-Analysis of the Literature Using Expanded Methods." *Journal of Economic Surveys* 29(3): 566–85.

Niederle, Muriel, "Gender" Handbook of Experimental Economics, second edition, Eds. John Kagel and Alvin E. Roth, Princeton University Press, 2016, 481-553.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122(3): 1067–1101.

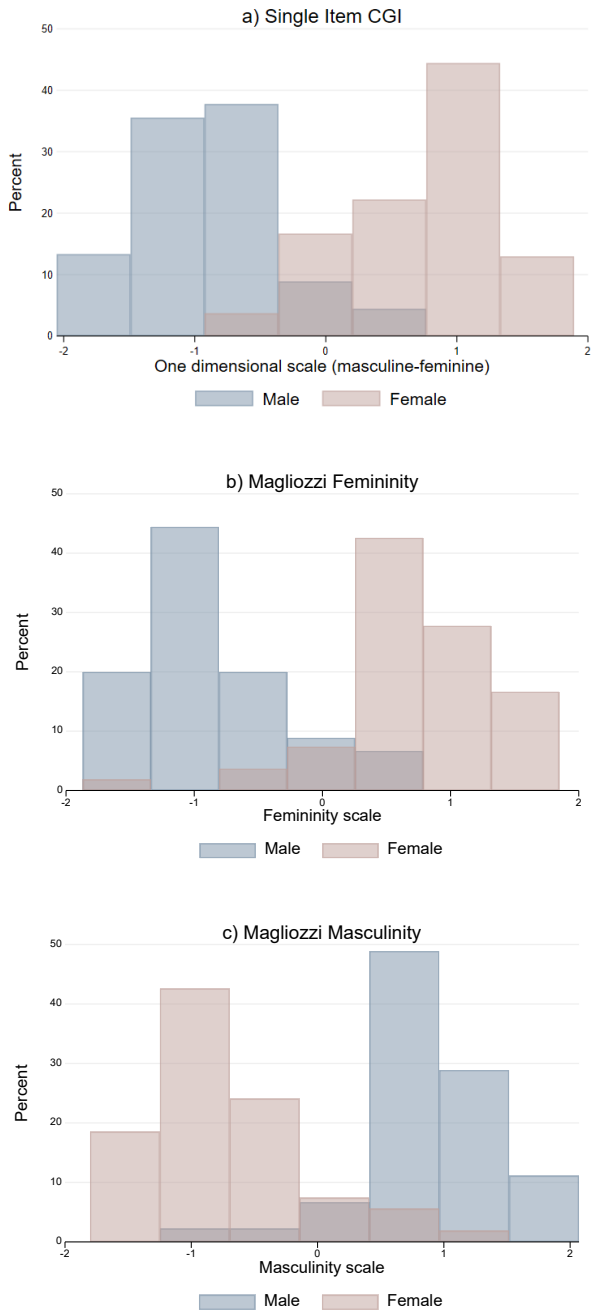
Westbrook, Laurel, and Aliya Saperstein. 2015. "New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys." *Gender & Society* 29(4): 534–60.

Continuous Gender and Economics

By ANNE ARDILA BRENØE, LEA HEURSEN, EVA RANEHILL, AND ROBERTO WEBER

ONLINE APPENDIX

FIGURE A1. DISTRIBUTIONS OF MAGLIOZZI FEMININITY, MASCULINITY AND OUR UNIDIMENSIONAL SCALE



Note: Scores are from the first-order gender identity questions and are standardized to have a mean of zero and a standard deviation of one. For our single item CGI question in Panel (a), the unidimensional scale ranges from "very masculine" to "very feminine" on a 7-point scale. For the Magliozzi scales in Panels (b) and (c), the scales range from "not at all" to "very" masculine and feminine, respectively.

TABLE A1 – PRINCIPAL COMPONENT ANALYSIS OF THE SEVEN CONTINUOUS GENDER SCALES

Panel A. Component loadings				
	Comp1	Comp2	Comp3	Unexplained
BSRI Fem	0.236	0.121	0.824	0.129
BSRI Masc	-0.172	0.797	-0.341	0.119
Mag Fem	0.473	0.134	-0.159	0.120
Mag Masc	-0.454	0.055	0.234	0.171
OSRI Fem	0.347	0.501	0.225	0.242
OSRI Masc	-0.353	0.279	0.239	0.398
TMF	0.490	-0.010	-0.133	0.085
Panel B. Eigenvalues and variance explained				
	Eigenvalue	Difference	Proportion	Cumulative
Component 1	3.745	2.708	0.535	0.535
Component 2	1.037	0.084	0.148	0.683
Component 3	0.953	0.346	0.136	0.819
Component 4	0.607	0.219	0.087	0.906
Component 5	0.388	0.219	0.055	0.962
Component 6	0.169	0.068	0.024	0.986
Component 7	0.100	.	0.014	1.000

Note: This table shows the results from a principal component analysis of the seven continuous gender scales (excluding our single item measure). Panel A presents the component loadings for the first three components; the final column, 'unexplained' refers to the proportion of the variance which cannot be explained when only these first three components are considered. Taken together, the seven components explain 100 percent of the variance. Panel B lists the eigenvalues corresponding to each component (column 1), and the difference between these eigenvalues. The final two columns report the proportion and cumulative proportion of the variance which can be explained by the relevant components.