Original Article

AMERICAN SOCIOLOGICAL ASSOCIATION

Journal of Health and Social Behavior
2019, Vol. 60(1) 101–118

DOI: 10.1177/0022146519825749
jhsb.sagepub.com
SAGE

Gender and Health: Beyond Binary Categorical Measurement

Chloe Grace Hart[1], Aliya Saperstein[1], Devon Magliozzi[1], and Laurel Westbrook[2]

Abstract

This study leverages multiple measures of gender from a US national online survey (N = 1,508) to better assess how gender is related to self-rated health. In contrast to research linking feminine behaviors with good health and masculine behaviors with poor health, we find that masculinity is associated with better self-rated health for cisgender men, whereas femininity is associated with better self-rated health for cisgender women. The patterns are similar whether we consider self-identification or how people feel others perceive their gender, though reflected appraisals are most strongly associated with health for cisgender women. We also find that people who report they are seen as gender nonconforming report worse health, but only when this perception does not match their gender identification. Our results demonstrate that multiple measures of gender allow researchers to disentangle how health is not only shaped by gender enactments but also shapes perceptions of gender and gender difference.

Keywords

gender nonconformity, health disparities, reflected appraisals, sex, sexual orientation, survey measurement

Disparities in health between women and men have been the norm for the past several centuries—since demographers began collecting systematic vital statistics—although the direction and magnitude of such differences vary over time and across places. In the contemporary United States, women are more likely to suffer from chronic conditions and disability, whereas men are more likely to suffer from life-threatening conditions and have lower average life expectancy (Read and Gorman 2010). In seeking to understand these gendered health profiles, sociologists and health researchers have recognized the importance of exploring both biological and social determinants of health (e.g., Annandale and Hunt 1990; Bird and Rieker 1999), including the relationships between femininity or masculinity and health behaviors (e.g., Courtenay 2000; Moore 2010).

However, a lack of clarity in the empirical measurement of gender continues to hamper our understanding of how gendered disparities in health are produced and maintained (Springer, Stellman, and Jordan-Young 2012). One form this takes is the conflation of sex and gender. Major social science surveys do not measure sex and gender separately and treat gender category terms, such as *woman* and *man*, and sex category terms, such as *female* and *male*, as synonymous (Westbrook and Saperstein 2015). As a result, many studies—including most cited here—cannot empirically distinguish between gender and sex. Throughout this paper, rather than make assumptions about whether a study measured sex and/or gender, we reference the same categories that were used in the cited publication.

[1]Stanford University, Stanford, CA, USA
[2]Grand Valley State University, Allendale, MI, USA

Corresponding Author:
Chloe Grace Hart, Sociology Department, Stanford University, 450 Serra Mall, Building 120, Room 160, Stanford, CA 94305-2047, USA.
E-mail: cghart@stanford.edu

Another limitation of current research is the treatment of gender categories as internally homogeneous. Studies that rely on binary measures mask variation within categories despite evidence that the relationship between gender and health is not neatly dichotomous. For instance, although women tend to enact health-protective behaviors and men tend to enact behaviors that are detrimental to health, there is a great deal of variation in the extent to which individual women or men enact gender-stereotypical health behaviors (e.g., Galdas et al. 2010). The widespread reliance on conventional measures—that conflate sex and gender and do not allow for diversity within categories—means that current research may both overgeneralize observed associations between gender and health and overlook key causal factors that perpetuate disparities and the resulting implications for health policy. Improvements must therefore go beyond clearly distinguishing the concepts of "sex" and "gender" to capture multiple dimensions of gender.

We demonstrate the feasibility and value of improving measures of gender in health research by exploring the relationship between gender identification and self-rated health among cisgender people. Cisgender people identify with the same sex and gender categories to which they were assigned at birth (Schilt and Westbrook 2009). We use both categorical and gradational measures of gender and a national sample of US adults to reveal nuance in the relationship between gender and health: Consistent with past research, we find that femininity and masculinity are related to self-rated health but the observed relationships for self-perceived femininity and masculinity are opposite for cisgender women and men. We also demonstrate the utility of considering external perceptions by showing that how people feel others perceive their femininity and masculinity is associated with health, in some cases in interaction with self-perceptions. Moreover, we affirm the importance of distinguishing between gender identification and sexual orientation by showing that each has a distinct relationship to self-rated health. The relationship between gender and health is complex; rather than mask this complexity, we offer a survey measurement strategy that better equips researchers to expose and explain it.

BACKGROUND

Scholars have called for nuanced measures of gender for the past several decades to provide a clearer understanding of how gender, as opposed to biological sex, is linked to health (e.g., Annandale and Hunt 1990; Bauer et al. 2017; Bird and Rieker 1999; Courtenay 2000; Galdas et al. 2010; Krieger 2003; Lagos 2018; Moore 2010). Although distinct measures of sex and gender remain unavailable in national surveys in the United States, researchers have theorized structural and individual determinants linking femininity and masculinity to health, including widespread cultural beliefs and gender-stereotypical behaviors.

Disentangling Gender and Health

Historically, beliefs about the gendered nature of health have aligned femininity with poor health and masculinity with good health. Women have long been seen as more susceptible to illness, hysteria, and hormonal instability, therefore requiring the supervision and management of doctors (Moore 2010). Masculinity, meanwhile, has historically been equated with control over the body such that men are expected to embody fortitude and invulnerability (Courtenay 2000). These historical notions of gender and health produced durable stereotypes that align femininity with vulnerability and weakness and masculinity with stoicism and strength.

Yet, even as cultural beliefs associate femininity with fragility and masculinity with robustness, researchers have found the opposite relationship between gender-stereotypical behaviors and health outcomes. In general, stereotypically feminine behaviors are protective of health, whereas stereotypically masculine behaviors are detrimental. For example, women are more likely to engage in health maintenance behaviors such as eating a healthy diet and seeking preventive medical care, whereas men are more likely to forgo medical care and engage in behaviors that damage health but affirm masculinity, such as smoking and not using sunscreen (Bird and Rieker 1999; Read and Gorman 2010; Schrock and Schwalbe 2009; Springer and Mouzon 2011). Although these gendered behaviors seem to contradict historical notions equating femininity with bad health and masculinity with good health, they could stem from acceptance of these stereotypes: Women may proactively monitor their health because they believe their bodies are vulnerable and men who perceive themselves as invincible might discount protective behaviors (Courtenay 2000; Moore 2010).

Even if, on average, women are more likely to engage in protective, feminine-stereotyped health behaviors and men are more likely to engage in detrimental, masculine-stereotyped health behaviors, women and men do not enact these behaviors

Table 1. Multiple Dimensions of Gender.

Concept	Measurement	Advantages
Categorical sex and gender	Two-step approach Question 1: Sex at birth Answer options: Female, male, intersex Question 2: Current gender Answer options: Woman, man, transgender, nonbinary, genderqueer	Allows for identification of cisgender and transgender people Offers nonbinary answer options
Gradational gender	Self-rated femininity scale Self-rated masculinity scale Reflected appraisal femininity scale Reflected appraisal masculinity scale	Allows for measuring diversity within gender categories Allows for differences across observers

uniformly. Qualitative interviews reveal that when deciding whether to seek medical care, some women and some men express stoic ideals, some describe feelings of vulnerability, and some are guided by both perspectives (Galdas et al. 2010). Furthermore, identification with stereotypical feminine or masculine traits does not necessarily predict health behaviors. Although in some contexts survey respondents rated more masculine on the Bem Sex Role Inventory (BSRI) engaged in more risky behaviors such as drinking and smoking, in other age cohorts and countries, respondents rated as feminine engaged in such behaviors or both femininity and masculinity predicted risk-taking behavior simultaneously (Emslie, Hunt, and Macintyre 2002; Hunt, Hannah, and West 2004).

The relationship between gender and health is further complicated when looking beyond behaviors to self-rated overall health or particular health diagnoses. For example, a study of coronary heart disease found that higher femininity scores on the BSRI were associated with lower mortality among men but found no association among women (Hunt et al. 2007). Meanwhile, a study of self-rated health found that among both women and men, higher femininity scores on the BSRI were associated with worse self-rated health and higher masculinity scores were linked to better self-rated health (Annandale and Hunt 1990). This work shows that femininity and masculinity are related to health in ways that are masked by categorical measures and suggests that femininity and masculinity are not related to health behaviors, health outcomes, and self-rated health in a single, straightforward way. Structural inequality—such as disparities in pay and child care provision—also plays a role in gendered health outcomes (Krieger 2003; Read and Gorman 2010). The relationship between gender and health is thus more complex than is typically accounted for in empirical research and calls for multidimensional measures of both gender and health.

Multiple Dimensions of Gender

Disentangling the relationship between various dimensions of gender and health requires precision about the underlying concepts being measured. This includes distinguishing between sex and gender, recognizing nonbinary categories, and using gradational measures to allow for variation within gender categories (see Table 1). Given that gender is interactionally negotiated, it is also important to distinguish whose determination of gender is being measured. All of these improvements would bring the empirical operationalization of sex and gender more in line with contemporary social science theory.

Separating sex and gender. Gender scholars have long recognized sex and gender as distinct constructs: *Sex* refers to the social structure that categorizes bodies using biological criteria such as genitals and chromosomes, whereas *gender* refers to behaviors that are culturally associated with those sex categories (West and Zimmerman 1987). Although this distinction has achieved cross-disciplinary consensus, it has not been implemented in most large-scale surveys (Westbrook and Saperstein 2015). The slippage between sex and gender in standard survey instruments is problematic for two reasons. First, the failure to capture information about sex and gender separately makes it impossible to disentangle the effects of sex versus gender on health disparities and other outcomes (e.g., Annandale and Hunt 1990; Geist, Reynolds, and Gaytán 2017; Sumerau et al. 2017). Second, assuming that gender is evident based on sex disregards transgender people who do not identify with

the gender category corresponding to the sex they were assigned at birth (Flores et al. 2016).

Efforts in survey research to distinguish sex from gender and recognize transgender people have centered on asking separate items about sex at birth and current gender and including a more comprehensive range of gender identity responses, such as nonbinary and genderqueer (Bauer et al. 2017; Federal Interagency Working Group 2016). These improvements allow researchers to better monitor health disparities and discrimination experienced by transgender people and better identify their health needs (GenIUSS Group 2014; Jans et al. 2016).

Beyond categorical measures. In addition to recognizing nonbinary sex and gender categories, it is also important to acknowledge diversity within gender categories. Gender is not a fixed characteristic but is "done" through social practices culturally marked as feminine or masculine, making it a complex and unstable accomplishment (West and Zimmerman 1987). Within any gender category, people enact varying levels of femininity and masculinity. Gradational measures of gender can capture this within-category variation.

To better understand the relationship between gender and health in particular, we argue that explicit, self-rated measures of femininity and masculinity are needed. Health researchers have previously measured a person's adherence to stereotypically feminine and masculine traits using the BSRI (Bem 1974), but these traits may or may not contribute to a person's gender identification (Spence and Buckner 2000). Thus, rather than present respondents with a list of traits, we ask them directly to self-identify on a pair of femininity and masculinity scales. This measure does not represent a fixed sense of gender—responses may shift over time and across contexts—but rather how a person perceives their position within broader gender constructs. Like current categorical measures of gender, these scales cannot distinguish among the various structural, interactional, and individual levels through which gender is enacted (Risman 2004); however, they offer a more gradational approach to measuring people's gendered experience. Survey research uses similar gradational measures of self-identification in other domains: for example, to capture political ideology on a spectrum from liberal to conservative alongside categorical political party affiliation (i.e., whether one votes as a Democrat, Republican, or Independent).

The analytic potential of measuring gradational gender alongside categorical gender identification has remained largely unrealized. Recent studies suggest that gradational gender identification measures can reveal information about social attitudes related to political behavior, such as believing that women face job discrimination (Bittner and Goodyear-Grant 2017; see also Wängnerud, Solevid, and Djerf-Pierre, forthcoming), that are masked by categorical gender measures alone. We demonstrate that similar gains can be made by exploring gradational gender in conjunction with health.

Self- and other perceptions. As an interactional process, gender is perceived from multiple perspectives. A person does their gender by enacting patterns of behavior that are socially understood to be feminine or masculine, but their gender is simultaneously "determined" by others who perceive that enactment (Westbrook and Schilt 2014). It is often assumed that these perspectives converge: for example, that if others consider a woman very masculine, she will also perceive herself this way. However, there can be discrepancies between how people understand their gender and how others perceive their gender (Miller and Grollman 2015). Thus, when others view a woman as very masculine, the woman might perceive herself as somewhat or not at all masculine.

Given that a person's gender self-concept does not necessarily overlap with others' interpretations, the relationship between each perspective and health may also differ. Research showing that a person's gendered sense of self informs their health behaviors (Emslie et al. 2006; Galdas et al. 2010) suggests that direct measures of how a person rates their own femininity and masculinity are needed. A case can also be made for considering how one perceives one's gender to be viewed by others, or what social psychologists call a *reflected appraisal* (Cooley 1902; Mead 1934). Research focusing on transgender people finds reflected appraisals of gender to be linked to health outcomes. One study finds that transgender people who feel others perceive them to be transgender or gender nonconforming have more often engaged in self-harm and high-risk health behaviors and that this is partially explained by the greater discrimination they report (Miller and Grollman 2015). This finding is consistent with research showing that people who experience discrimination are less likely to engage in health-promoting behaviors and more likely to have negative health outcomes (Boardman 2004; Pascoe and Richman 2009).

Although research on reflected appraisals of gender has been restricted to transgender people (see also Rider et al. 2018), there is no reason to believe

reflected appraisals of gender are not relevant to cisgender people as well. As we have noted, gender is not only experienced as a wholesale, categorical identity: People identify with and are perceived to have varying degrees of femininity and masculinity regardless of their gender category. By measuring both self-rated femininity and masculinity and reflected appraisals, along with categorical gender identification, we can empirically disentangle whether and how each dimension of gender relates to health among cisgender women and men.

DATA AND METHODS

Our data came from a 2014 online survey of 1,522 US residents ages 18 or older. The study tested several gradational measures of femininity and masculinity alongside categorical measures of sex and gender to capture multiple dimensions of people's gendered experiences (see Magliozzi, Saperstein, and Westbrook 2016). Along with these expanded measures of gender, the survey asked several dozen questions drawn from the General Social Survey (GSS), including items about self-rated mental and overall health (Smith et al. 2017).

Respondents were recruited through Amazon Mechanical Turk (MTurk) and were paid $1.50 after completing the 10-minute survey. MTurk workers have been shown to provide higher quality data than online population-based samples (Weinberg, Freese, and McElhattan 2014). However, as with other MTurk samples, our sample overrepresented younger and more highly educated Americans, and respondents were more likely to identify as white or Asian (Weinberg et al. 2014). Although the sample was not nationally representative, it was more diverse than undergraduate student samples often utilized for experimental or exploratory research, and it aligned well with national estimates for other key variables. For example, our regional distribution was within one to two percentage points of the distribution of respondents in the 2014 GSS (e.g., 38% vs. 36% living in the South); the GSS reported its sample was 55% female and 45% male, and ours was 53% cisgender women and 47% cisgender men (see Magliozzi et al. 2016 for more details). Also, as in the 2014 GSS, we found women and men report similar average family income and educational attainment. Descriptive statistics for our sample are reported in Table 2.

Sex and Gender Measures

Ours was the first national survey of US adults that incorporated not only separate measures of sex and gender but also multiple measures of gender, including both categorical and gradational items. To assess whether presentation order influenced responses, we randomized placement of the sex and gender module within the questionnaire (in the middle of the survey or at the end) and question order within the module (scales were either first or last, and the order of self-rated and reflected appraisal scales also varied). Responses to the sex and gender items did not differ significantly across display conditions. Nevertheless, all models presented here included controls for the conditions to account for any differences generated by the random assignment of respondents.

Categorical sex and gender. We queried respondents' sex and gender separately using a two-step approach shown to inclusively capture gender identity without confusing cisgender respondents (Bauer et al. 2017; GenIUSS Group 2014). Respondents were asked, "What sex were you assigned at birth? (For example, on your birth certificate)." The possible responses were female, male, and intersex; no respondents in this sample chose intersex. Respondents were then asked, "What is your current gender?," with the options woman, man, transgender, and a gender not listed here.

Among the 1,522 respondents, all but 8 were classified as cisgender. Of the other respondents, 3 were labeled male at birth and identified as women, 2 were labeled female at birth and identified as transgender, and 3 were labeled female at birth and identified as a gender not listed (e.g., genderqueer). The prevalence of transgender respondents was in line with previous research (Flores et al. 2016) but was too small for statistical analysis. Thus, we focused on cisgender adults. We dropped respondents with missing values on the variables used in our models, for a total analytic sample of 1,508 respondents.

Self-rated gender scales. We also asked respondents to rate their femininity and masculinity on two separate 7-point scales in response to the question, "In general, how do you see yourself?" (Figure 1). Although some studies measure masculinity and femininity on a single, bipolar scale (e.g., Rider et al. 2018; Wylie et al. 2010), following long-standing recommendations in psychology (see Constantinople 1973), we used two unipolar scales to avoid treating masculinity and femininity as mutually exclusive and opposite. As might be expected, we found that cisgender women rated themselves as more feminine than cisgender men on average (4.7 vs. 1.1) and cisgender men rated themselves as more masculine than cisgender women (4.7 vs. 1.2). However, we

Table 2. Descriptive Statistics.

	Cisgender Women (N = 795)		Cisgender Men (N = 713)	
	Mean	SD	Mean	SD
Self-rated femininity	4.7	1.2	1.1	1.2
Self-rated masculinity	1.2	1.2	4.7	1.2
Reflected appraisal femininity	4.8	1.2	1.0	1.2
Reflected appraisal masculinity	1.1	1.3	4.6	1.2
Self-rated gender nonconforming (%)	4.0		3.4	
Reflected appraisal gender nonconforming (%)	4.9		4.2	
Age (years)	36.1	11.8	32.7	10.5
Education (years completed)	14.6	2.7	14.6	2.7
Household income ($)	54,000	36,000	54,000	38,000
	%	N	%	N
Self-rated health				
Excellent	17.1	136	17.1	122
Good	61.3	487	58.5	417
Fair	19.0	151	22.2	158
Poor	2.6	21	2.2	16
Sexual orientation				
Gay/lesbian/homosexual	3.1	25	3.7	26
Bisexual	8.4	67	3.1	22
Straight/heterosexual	88.4	703	93.3	665
Married	45.4	361	29.3	209
Self-identified race				
White	80.6	641	82.2	586
Black	8.3	66	4.4	31
All other responses	11.1	88	13.5	96
Hispanic origin	5.5	44	9.1	65
Born outside the United States	5.3	42	5.6	40
Region				
Northeast	14.8	118	19.8	141
South	40.3	320	34.6	247
West	23.9	190	25.4	181
Midwest	21.0	167	20.2	144

Source: Authors' survey, November 2014.
Note: SD = standard deviation.

also found that less than one-third of participants rated themselves at the maximum of their sex-typical gender identification scale, a result that challenges the all-or-nothing relationship implied by binary measurements. Operationalizing gender as gradational therefore yields information not captured when gender is measured solely categorically.

Reflected appraisal gender scales. Our reflected appraisal measures asked participants to rate how others perceive their femininity and masculinity in response to the question, "In general, how do most people see you?" (see Figure 1).[1] About half of respondents reported either masculinity or femininity self-ratings that were not identical to how they felt others perceived them. Among respondents who did not report their femininity or masculinity as identical to how they thought others perceived them, the majority (83% and 84%, respectively) differed by one point (see Figure 2). Relatively small differences between reflected appraisals and self-identification are to be expected (see Burke 1991),

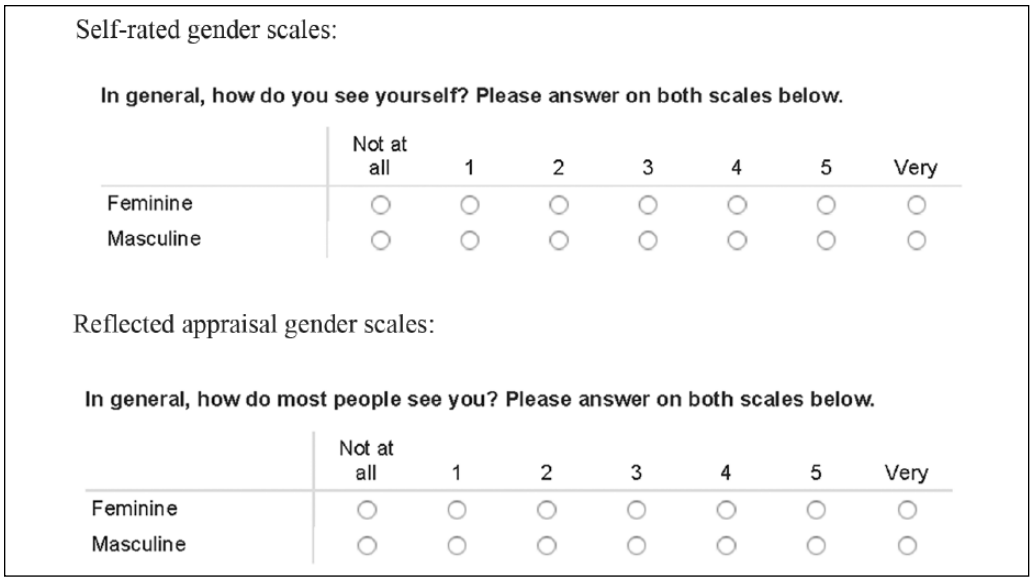

Figure 1. Gradational Gender Measures.

and the magnitude alone does not rule out important associations with health.

Using the gender identification scales, we derived a measure that we termed *self-rated gender nonconformity*, in reference to the societal expectation that one's gender identification corresponds to the sex one was assigned at birth. We defined gender nonconformity as women who consider themselves more masculine than feminine and men who consider themselves more feminine than masculine. Four percent of cisgender women and 3.4% of cisgender men in our sample rated themselves as gender nonconforming (N = 32 and 24, respectively). *Reflected appraisal gender nonconformity* refers to analogous items derived from the reflected appraisal scales (e.g., for women, those who believe that others see them as more masculine than feminine). In our sample, 4.9% of cisgender women and 4.2% of cisgender men believed that others see them as gender nonconforming (N = 39 and 30, respectively). Among gender-nonconforming cisgender women and men, just one-quarter were nonconforming on both sets of scales, underscoring the need for distinct measures.

Gender nonconformity could be conceptualized more expansively to also include those who regard themselves as equally masculine and feminine. Using this expanded definition, 11.7% of cisgender women and 10.2% of cisgender men rated themselves as gender nonconforming (N = 93 and 73, respectively), and 11.2% of cisgender women and 8.7% of cisgender men believed that others see them as gender nonconforming (N = 89 and 62, respectively). We used the more conservative definition of gender nonconformity in our primary analyses, but we discuss analyses using the more expansive definition where results differ. The question of how to conceptualize gender nonconformity is an important topic for future research that will be facilitated by expanding the available measures of gender.

Health Measures

Our dependent variable was self-rated health, a measure that significantly predicts mortality (Jylhä 2009) and has become more predictive over time as people have gained access to increasingly accurate medical information (Schnittker and Bacak 2014). Nevertheless, we caution that it should not be overinterpreted as applying to all dimensions of health: Self-rated health is one of a broad set of indicators alongside health behaviors and outcomes that only together offer a full picture of health.

To measure self-rated health, we asked the question used by the GSS: "Would you say your own health, in general, is excellent, good, fair or poor?" Very few respondents (2.5%) rated their health as poor in our relatively young sample. We collapsed responses, producing a measure of health with the categories fair/poor, good, and excellent; results did not substantively differ when all four categories were used. The self-rated health of cisgender

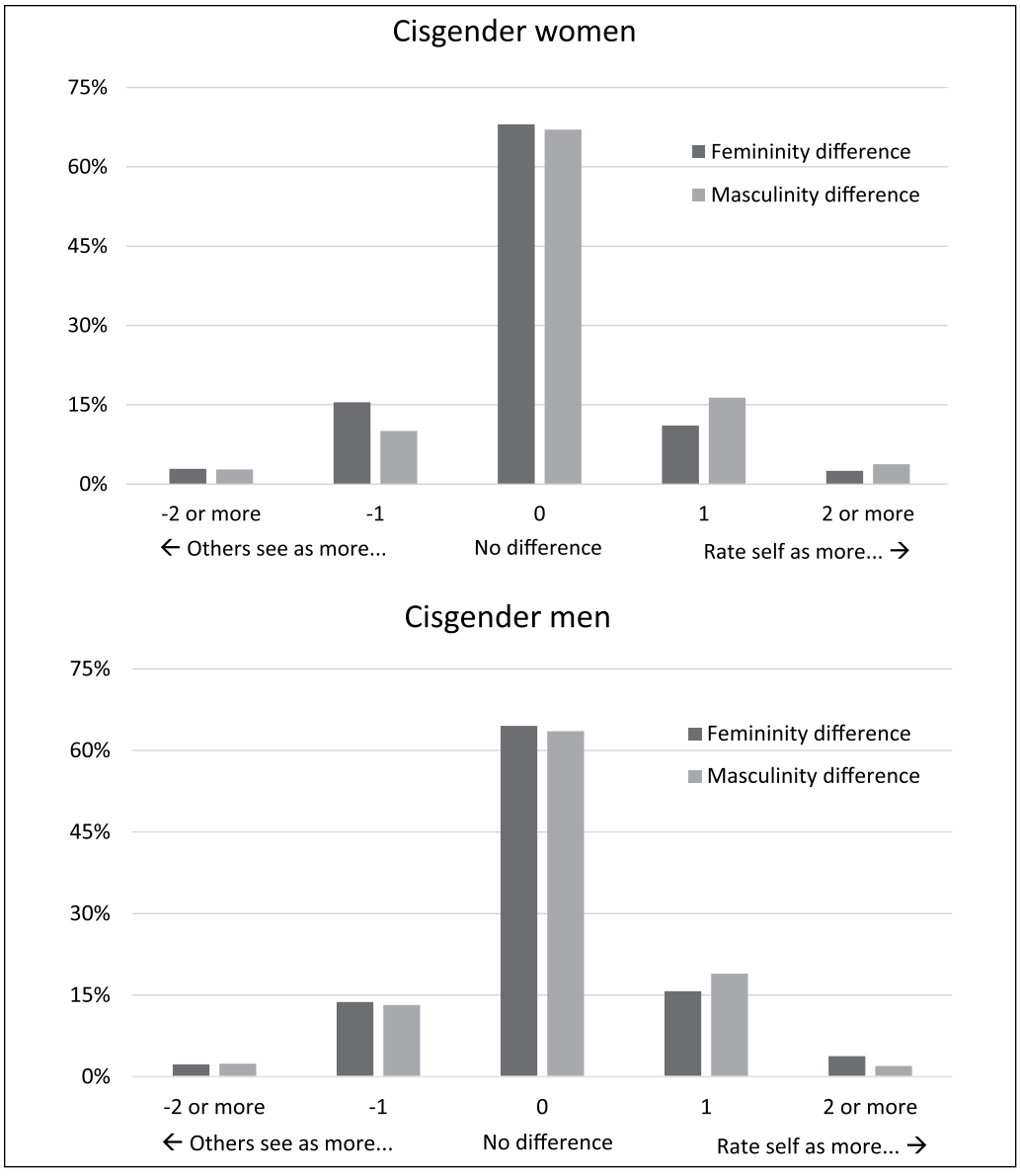

Figure 2. Differences between Self-rated and Reflected Appraisal Gender Scales.
Source: Authors' survey, November 2014.

women and men was similar: Just over 20% rated their health as fair or poor, about 60% chose good, and just under 20% chose excellent. Differences in self-rated health between cisgender women and men were not statistically significant in our data (t statistic = .85, p = .40). This is consistent with previous research given similar education and family income in our sample; in studies that find women report slightly worse self-rated health than men, the difference disappears when social determinants are included as controls (e.g., Prus 2011).

The findings reported here apply to more than one dimension of self-rated health. To supplement results presented for self-rated overall health, we ran parallel models for self-rated mental health, measured as how many of the past 30 days the respondent's mental health was not good, considering symptoms such as stress, depression, and

problems with emotions. In comparing results using self-rated mental health to self-rated overall health, we found substantively similar relationships between gender identification and both health measures (see the Supplemental Material in the online version of the article).

Analysis

To examine the relationship between gender and self-rated health, we estimated ordinary least squares (OLS) regression models predicting our three-category measure of self-rated health (positive estimates indicate better health). We also estimated both multinomial and ordered logistic regressions; all three sets of models yielded consistent estimates in both direction and statistical significance (see Supplemental Material in the online version of the article). We present the OLS estimates here for ease of interpretation.

To limit confounding with other demographic characteristics that are related to self-rated health and adjust for ways in which our sample varied in makeup from a nationally representative sample, we controlled for several additional variables. These included age, self-identified race and Hispanic origin, whether the respondent was married, years of education, income, region of residence, and whether the respondent was born in the United States. We also accounted for sexual orientation, with the options gay, lesbian, or homosexual; bisexual; and heterosexual or straight. For the most part, the relationships between our controls and self-rated health were in line with previous research utilizing data gathered through probabilistic sampling. For example, we found Hispanic origin, education, and income were positively associated with health, whereas age and bisexuality were negatively associated (see Gorman et al. 2015; Markides and Eschbach 2005). However, in our sample, there were not significant differences in self-rated health between people who identify as black and people who identify as white (cf. Boardman 2004); we also did not find a significant health penalty for people born outside of the United States (cf. Prus 2011). This may have been due to a different distribution of countries of origin in our data or the fact that our sample was comprised of younger and more highly educated individuals across subpopulations.

RESULTS

Our analysis shows that femininity and masculinity are significantly linked to health. We find that for both cisgender women and men, rating oneself high on the gender scale normatively associated with one's gender category (e.g., femininity for women) is associated with better reported health. However, the way cisgender women think others view their gender is more strongly tied to their reported health than their self-identification, whereas for cisgender men, self-ratings are most strongly linked to reported health. Finally, we find that among cisgender women and men, those who think others view them as gender nonconforming report slightly worse health on average, but those who also see themselves as gender nonconforming do not. These findings demonstrate that the relationships between femininity, masculinity, gender conformity, and health cannot be assumed to be identical for cisgender women and men. Furthermore, they show the value of multiple measures of gender in understanding the complex links between gender and health.

Self-rated Gender and Health

Previous research focusing on health behaviors suggests that the relationship between femininity, masculinity, and health works similarly for women and men, with feminine behaviors associated with better health and masculine behaviors associated with worse health. We assess whether this relationship holds between gradational gender and self-rated health by examining whether self-rated femininity and masculinity have similar relationships with self-rated health for cisgender women and men.[2] Table 3 indicates that they do not; rather, these relationships run opposite for cisgender women and men in our sample. For cisgender women, each point of self-rated femininity comes with a statistically significant health boost of .05 points so that cisgender women who rate their femininity at the maximum of 6 are expected to report health that is .30 points higher than cisgender women who rate their femininity at the minimum of 0 (Model 1a). This is a moderate difference that covers about one-third of the difference between one health designation and the next (e.g., between good and excellent). Conversely, the relationship between self-rated femininity and self-rated health for cisgender men is negative and marginally significant. Each point of femininity that cisgender men assign themselves is associated with a health decrease of .04 points ($p = .05$) so that the average difference in self-rated health between cisgender men with the minimum versus maximum femininity is about one-quarter of a point (Model 1b).

We find self-rated masculinity is not meaningfully associated with self-rated health for cisgender

Table 3. Self-rated Gender Scales as Predictors of Self-rated Health.

	Cisgender Women			Cisgender Men		
	Model 1a	Model 2a	Model 3a	Model 1b	Model 2b	Model 3b
Self-rated femininity	.05*		.06*	–.04		.01
	(.02)		(.02)	(.02)		(.02)
Self-rated masculinity		–.01	.02		.11***	.11***
		(.02)	(.02)		(.02)	(.02)
Age (years)	–.00*	–.00	–.00	–.00	–.00	–.00
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Education (years)	.02*	.02*	.02*	.04***	.04***	.04***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Income (in $10,000s)	.02**	.02**	.02**	.01	.01	.01
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Sexual orientation (reference = straight)						
Gay/lesbian/homosexual	.17	.14	.16	–.05	–.01	–.01
	(.13)	(.13)	(.13)	(.13)	(.12)	(.12)
Bisexual	–.23**	–.25**	–.24**	–.24	–.19	–.20
	(.08)	(.08)	(.08)	(.14)	(.14)	(.14)
Married	.05	.06	.05	–.01	–.03	–.03
	(.05)	(.05)	(.05)	(.06)	(.06)	(.06)
Self-identified race (reference = all else)						
White	.07	.07	.07	.11	.13	.13
	(.07)	(.07)	(.07)	(.08)	(.07)	(.07)
Black	–.04	–.02	–.03	.19	.13	.13
	(.10)	(.10)	(.10)	(.13)	(.13)	(.13)
Hispanic origin	.18	.17	.18	.12	.09	.09
	(.10)	(.10)	(.10)	(.08)	(.08)	(.08)
Born outside the United States	.06	.06	.06	–.04	–.02	–.02
	(.10)	(.10)	(.10)	(.11)	(.11)	(.11)
Region (reference = Midwest)						
Northeast	.06	.06	.06	–.02	–.02	–.02
	(.07)	(.07)	(.07)	(.07)	(.07)	(.07)
South	–.01	.00	–.01	–.03	–.02	–.02
	(.06)	(.06)	(.06)	(.07)	(.07)	(.07)
West	.06	.07	.06	–.09	–.09	–.09
	(.06)	(.06)	(.06)	(.07)	(.07)	(.07)
Constant	1.40***	1.62***	1.33***	1.39***	.84***	.80***
	(.20)	(.19)	(.22)	(.190)	(.21)	(.22)
Observations (N)	795	795	795	713	713	713
BIC	1552	1557	1558	1450	1426	1432
AIC	1467	1473	1469	1368	1343	1345
R^2	.08	.07	.08	.07	.11	.11

Source: Authors' survey, November 2014.
Note: Ordinary least squares regressions. Standard errors in parentheses. Models include controls for survey condition (not shown). The estimate for age is between –.003 and –.004 in all models but rounds to zero at two significant digits. BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion.
*$p < .05$, ** $p < .01$, *** $p < .001$.

women, whereas self-rated masculinity is strongly and positively associated with self-rated health for cisgender men. Each additional masculinity point cisgender men assign themselves is associated with .11 points better average health, an estimate that is double that of self-rated femininity for cisgender women (Models 2a and 2b). This means a cisgender man who rates his masculinity at the maximum of 6

Table 4. Comparing Self-rated and Reflected Appraisal Gender Scales as Predictors of Self-rated Health.

	Cisgender Women		Cisgender Men	
	Model 1a	Model 2a	Model 1b	Model 2b
Reflected appraisal femininity	.06**	.05		
	(.02)	(.03)		
Self-rated femininity		.00		
		(.03)		
Reflected appraisal masculinity			.10***	.03
			(.02)	(.03)
Self-rated masculinity				.09**
				(.03)
Constant	1.32***	1.32***	.91***	.82***
	(.21)	(.21)	(.21)	(.21)
Observations (N)	795	795	713	713
BIC	1549	1555	1432	1431
AIC	1464	1466	1350	1345
R^2	.08	.08	.10	.11

Source: Authors' survey, November 2014.
Note: Standard errors in parentheses. Models include all controls reported in Table 3. BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion.
**$p < .01$, *** $p < .001$.

is expected to have self-rated health that is .66 points higher than that of a cisgender man who rates his masculinity at the minimum of 0, a substantively large difference that covers most of the difference between one health designation and the next.

When we include both self-rated masculinity and femininity in the same model, only the gender scale traditionally associated with the given gender category is strongly associated with self-rated health. This suggests that for cisgender women, self-rated femininity is more closely linked to reported health than self-rated masculinity, whereas for cisgender men, self-rated masculinity is more closely related to reported health than self-rated femininity. Although we find the scale that matters most varies between cisgender women and men, Table 3 clearly demonstrates that self-rated gender is significantly associated with self-rated health. Furthermore, we find a statistically significant association between sexual orientation and self-rated health in addition to the association between self-rated gender and health, a point to which we return in the discussion.

Reflected Appraisal Gender and Health

Having established a link between self-rated gender and health in our sample, we now turn to the relationship between health and how respondents believe others perceive their gender, that is, their reflected appraisal gender. We find reflected appraisal femininity, like self-rated femininity, is positively associated with reported health for cisgender women (Table 4, Model 1a). The estimate for reflected appraisal femininity is in fact slightly larger than that of self-rated femininity: Each point of femininity that cisgender women in our sample report others attribute to them is associated with .06 points better average self-rated health. When both the self-rated and reflected appraisal femininity scales are included (Model 2a), the reflected appraisal measure remains larger and marginally statistically significant ($p = .08$), and the self-rated measure drops to zero. Thus, how a cisgender woman believes others determine her gender is more strongly associated with her self-rated health.

For cisgender men in our sample, reflected appraisal masculinity is likewise associated with self-rated health. However, the magnitude of the reflected appraisal estimate (Table 4, Model 1b) is slightly smaller than that for self-rated masculinity: Each point of masculinity that cisgender men believe others attribute to them is associated with .10 points better average health. Further, when we compare the relative strength of the relationship between self-rated masculinity and reflected appraisal masculinity for cisgender men (Model 2b),

only self-rated masculinity remains significantly associated with their self-rated health. Thus, whereas reflected appraisal gender is most strongly associated with reported health for cisgender women, self-rated gender is most strongly associated with reported health for cisgender men.

Model fit statistics provide further evidence of this gender distinction. For cisgender men in our sample, Model 2b in Table 3, which includes only self-rated masculinity, is the best fitting model according to both the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC), with lower values indicating better fit. For cisgender women, Model 1a in Table 4, which includes only reflected appraisal femininity, fits best according to both model fit statistics. For all models in Table 4, the variance inflation factor scores for all predictor variables were less than 3 (well below the conventional cutoff of 10), indicating that multicollinearity is not a concern. Thus, self-rated masculinity is most strongly associated with reported health for cisgender men in terms of effect size, statistical significance, and model fit, whereas reflected appraisal femininity is most strongly associated with reported health for cisgender women on these measures. This finding is consistent with previous research demonstrating that women consider how they are perceived by others more than men (Slevin 2010).

Gender Nonconformity and Health

The aforementioned analyses demonstrate a significant link between gradational gender measures and self-rated health. Masculinity, and particularly self-rated masculinity, is positively associated with reported health for cisgender men, whereas femininity, and particularly one's femininity as it is believed to be viewed by others, is positively associated with reported health for cisgender women. Put another way, there could be a health penalty for cisgender people whose gender identification does not align with their gender category in the normatively expected way. Although our sample of nonconforming respondents is not large, our results, combined with research highlighting the role of gender nonconformity in health for transgender people, raise further questions about the nature of this relationship. Is rating oneself as gender nonconforming (for cisgender women, those who view themselves as more masculine than feminine and for cisgender men, those who view themselves as more feminine than masculine) related to one's self-rated health?

We find that gender-nonconforming people report worse health, on average, but this difference is reduced or eliminated when self-rated gender nonconformity aligns with reflected appraisal gender nonconformity. Cisgender women who view themselves as more masculine than feminine do not have significantly different average self-rated health from other cisgender women (Table 5, Model 1a), whereas cisgender women who feel that others see them as more masculine than feminine report worse self-rated health by about one-quarter of a point on average (Model 2a). However, we also find the interaction term between self-rated and reflected appraisal gender nonconformity is positive and significant (Model 3a). The interaction indicates that cisgender women who believe that others see them as more masculine than feminine report significantly lower health on average, but cisgender women who also rate themselves more masculine than feminine report their health as a half-point *higher* on average, which more than makes up for the difference.

For cisgender men, we find that men who regard themselves as more feminine than masculine have self-rated health that is about one-quarter of a point lower than men who regard themselves as equally or more masculine than feminine, but the estimate is not statistically significant at conventional levels (Model 1b, $p = .08$). Cisgender men in our sample who believe that others view them as more feminine than masculine report significantly worse health by about one-third of a point on average relative to other cisgender men (Model 2b). Finally, as with cisgender women, we find that the negative association with gender nonconformity is mitigated if a cisgender man both rates himself as gender nonconforming and believes others view him this way (Model 3b). Although the interaction term for cisgender men is positive (.14), the estimate is smaller than that of cisgender women and not statistically significant. Thus, overall, we find that worse self-rated health is more strongly associated with gender nonconformity when people feel that others are determining their gender in a way that does not match their own self-concept, and this relationship between perceived gender nonconformity and self-rated health is similar for cisgender women and men.[3]

When we expand the concept of nonconformity to include those who report equal femininity and masculinity, the positive interaction between self-rated gender nonconformity and reflected appraisal gender nonconformity holds for both cisgender women and men (see Supplemental Material in the online version of the article). However, the weight of evidence shifts: The interaction term is now substantively large and statistically significant for

Table 5. Comparing Self-rated and Reflected Appraisal Gender Nonconformity as Predictors of Self-rated Health.

	Cisgender Women			Cisgender Men		
	Model 1a	Model 2a	Model 3a	Model 1b	Model 2b	Model 3b
Self-rated nonconformity	.01	.10	–.08	–.23	–.07	–.12
	(.12)	(.12)	(.15)	(.13)	(.14)	(.18)
Reflected appraisal nonconformity		–.23*	–.36**		–.36**	–.40**
		(.11)	(.13)		(.13)	(.15)
Self-rated nonconformity × reflected appraisal nonconformity			.51*			.14
			(.25)			(.30)
Constant	1.60***	1.61***	1.64***	1.34***	1.35***	1.35***
	(.18)	(.18)	(.18)	(.19)	(.19)	(.19)
Observations (N)	795	795	795	713	713	713
BIC	1557	1559	1562	1451	1449	1456
AIC	1473	1471	1469	1369	1363	1364
R^2	.07	.08	.08	.07	.08	.08

Source: Authors' survey, November 2014.
Note: Standard errors in parentheses. Models include all controls as previously reported. BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion.
*$p < .05$, ** $p < .01$, *** $p < .001$.

cisgender men, whereas for cisgender women, it is small and nonsignificant. We speculate that differing standards of nonconformity for women and men may help explain why the results vary in strength depending on how nonconformity is defined. Cisgender women who are equally feminine and masculine may be perceived as less nonconforming than cisgender men who are equally feminine and masculine given narrower cultural definitions of masculinity relative to femininity (McGuffey and Rich 1999). Nonetheless, using either conceptualization, we find people who see themselves as gender nonconforming and feel that others perceive them as nonconforming tend to have better health than those who feel others see them as gender nonconforming but do not agree with this assessment.

DISCUSSION

These results raise several points that merit further consideration. First, although previous research has linked femininity with positive health behaviors and masculinity with negative health behaviors, we find a more complex relationship between self-rated gender and self-rated health. This may be explained by differences between a person's gendered self-concept and the extent to which they enact gender-stereotypic behaviors relevant to health. Second, the causal pathways linking gender identification and health are complex and require further exploration; although our models imply a person's gender influences their health, aspects of a person's health could also affect the extent to which they feel feminine or masculine. Finally, our data suggest that sexual orientation and gender identification have distinct associations with health. This highlights the importance of employing distinct and specific measures of sex, gender, and sexual orientation to disentangle their relationships with health.

Distinguishing between Gender Identification and Gendered Behaviors

In our sample, cisgender adults whose gradational gender is normatively aligned with their gender category tend to have better self-rated health: Femininity is positively associated with health for women, and masculinity is positively associated with health for men. This finding seemingly runs counter to previous work showing that feminine-stereotyped behaviors are generally health promoting and masculine-stereotyped behaviors are generally damaging to health. However, the gender measures in our study capture identification rather than behavior. Although masculine-stereotyped behaviors such as smoking have negative effects on health and feminine-stereotyped behaviors such as monitoring one's diet have positive effects on health (Courtenay 2000; Moore 2010), partaking in these

gendered behaviors does not necessarily inform a person's gender identity. For example, a man who monitors his diet may still consider himself very masculine, whereas a woman who smokes may still consider herself very feminine. Our results point to the importance in distinguishing between behaviors and identification when investigating connections between gender and health.

The link between femininity and self-rated health for cisgender women and masculinity and self-rated health for cisgender men remains open for interpretation. We see several possible explanations. First, a person who does not conform to normative expectations of their gender category might be sanctioned by others for failing to adhere to gender norms, and such sanctions could be distressing and therefore detrimental to health, much like explicit discrimination; thus, conforming more strongly to those norms could be health protective. Second, some people who feel they do not conform to gender expectations overcompensate with gender-stereotyped behaviors (cf. Willer et al. 2013), and this may include behaviors that are detrimental to health. Alternately, as we discuss in the following, the relationship between self-rated gender and self-rated health may be recursive such that the two measures are inextricably bound together.

Addressing Causality

The models presented here imply a causal path in which self-rated gender informs self-rated health, in line with research that suggests a person's gender identification would play a role in their health behaviors. However, there are several alternate possibilities that could explain the observed relationship between self-rated gender and self-rated health.

First, a person's gendered sense of self could inform how he or she *report* their health, independent of their health behaviors or outcomes. How one reports and interprets one's health can be seen as a way of doing gender (Courtenay 2000; Moore 2010). For example, as masculinity is associated with invulnerability, people who regard themselves as masculine might refuse to acknowledge illness or physical limitations; those who regard themselves as feminine, meanwhile, might enact femininity by acknowledging bodily vulnerability (Galdas et al. 2010). Psychosomatic symptoms also can be real in their consequences. Over- or underestimating one's health is likely related to health behaviors and willingness to seek medical care, which affects future health outcomes. Although previous research comparing external health measures to self-rated symptoms has not found differential accuracy in women and men's self-reports (Bird and Rieker 1999), other research finds differences in health optimism between women and men (Grol-Prokopczyk, Freese, and Hauser 2011); in either case, gradational gender measures could reveal similarities and differences that categorical gender measures mask.

The causal pathway also could be reversed: A person's health might influence their gender identification. Previous research finds that men who develop health problems experience a diminished sense of masculinity, suggesting that poor health could hamper performance of gendered activities and thus shape a person's gendered sense of self (Emslie et al. 2006). Indeed, in an earlier pilot study testing our gender measures, respondents often referred to health-related behaviors and traits when explaining their gender identification. This occurred spontaneously in response to an open-ended request for feedback at the end of the survey. For example, one cisgender man who selected the maximum of the self-rated masculinity scale said, "I embrace being a male and enjoy my masculinity. I am 6'0, 245 lbs, 15% body fat, shaved head, beard. I strongly dislike men who act feminine. This includes men who listen to pop/rap, wear skinny jeans, and/or don't eat meat." He drew on both physical attributes (height, weight) and a health-related behavior (eating meat) in his gender assessment. A cisgender woman who rated herself as moderately feminine (4 on the 7-point scale) explained she chose this score "because I am morbidly obese, and as such have a more boxy figure than the typical slender hourglass that is considered a beautiful feminine shape." This respondent likewise drew on medical terminology (obesity) to explain her gender identification.

Finally, it is possible that factors not captured in these data might inform both one's gender identification and one's health. For example, regular exercise might cause cisgender women and men to evaluate their health better and also consider themselves either more feminine (due to the slimming effect) or masculine (due to increased strength). The data presented here are cross-sectional and thus cannot answer causal questions; the temporal ordering in the relationship between gender and health remains an important area for future research.

Disentangling Sexual Orientation from Gender Identification

In addition to highlighting the gradational gender heterogeneity among cisgender adults and revealing a relationship between self-rated gender and self-rated health, these data allow us to separately assess

the associations between sexual orientation, gradational gender, and health. Although sexual orientation and gender conformity are often conflated (e.g., gay men are stereotyped as more feminine than straight men), gender and sexual orientation are distinct concepts that operate independently of one another (Rubin 1984). With multiple measures of gender and sexual orientation in surveys, researchers can begin to disentangle whether the relationship between sexual orientation and health noted in previous research (e.g., Gorman et al. 2015) can be explained by conformity or nonconformity to gender expectations.

We find some evidence that gender identification and sexual orientation are independently and additively related to self-rated health. When we consider sexual orientation alone, we find—as in previous research—that bisexual people report worse health than either their straight or lesbian/gay counterparts (see Supplemental Material in the online version of the article). Including controls for self-rated gender reduces the size of the bisexual health estimate, but the relationship between bisexuality and self-rated health remains positive and statistically significant for cisgender women. For cisgender men, the coefficient for bisexuality remains positive but is no longer a statistically significant when gradational gender measures are included. However, we interpret this result cautiously as fewer cisgender men identify as bisexual (22) relative to cisgender women (67) in our sample. We also tested interactions between sexual orientation and gender conformity. They were not statistically significant for either cisgender women or men in our sample, but descriptively, our data indicate that gender conformity is positively associated with health for straight and bisexual cisgender women and negatively associated with health for lesbian cisgender women. We interpret this in line with previous research that finds gender nonconformity is more highly valued among lesbian women (Blair and Hoskin 2014). These speculative findings underscore the need to have separate measures of sex, sexual orientation, and both categorical and gradational gender in studies of health disparities to ensure these various factors are neither overlooked nor treated interchangeably.

CONCLUSION

Previous research offers conflicting accounts of how gender is related to health, leaving the link between how people experience their gender and their health unresolved. This study takes three steps toward better understanding the relationship between gender and health: (1) using distinct measures of sex, gender, and sexual orientation; (2) measuring gender identification as both gradational and categorical; and (3) measuring gender both as something one does and as something that is determined by others.

With these data, we find evidence of better self-rated health among cisgender adults whose gender identification is normatively aligned with their gender category: Femininity is positively associated with self-rated health for women, and masculinity is positively associated with self-rated health for men. We also explore the relationship between health and how people believe that others perceive their gender, a topic that previously has been studied only among the transgender population. We find that gender-reflected appraisals are pertinent to self-rated health for both cisgender women and men, but in comparing the relative strength between self-ratings and reflected appraisals, we discover a gender distinction. How cisgender men view their gender is most strongly associated with their self-rated health, whereas how cisgender women believe other people perceive their gender is most strongly linked to their self-rated health.

Finally, we investigate the experiences of cisgender people whose masculinity and femininity do not conform to normative expectations of their gender category. We find that cisgender people perceived by others as gender nonconforming tend to report worse health. However, this health disparity is alleviated for those who also assess themselves as gender nonconforming. Thus, it is not gender nonconformity per se that is negatively tied to self-rated health but rather the experience of being perceived as nonconforming when this does not align with one's self-concept.

Future health research can be improved by using the multiple measures of gender that we utilize here. Gradational measures reveal that gender is linked to self-rated health in distinct ways for cisgender women and men, demonstrating the importance of recognizing diversity within gender categories. Of course, our study also raises new questions: To what extent is one's gendered self-concept a cause or a consequence of one's self-rated health? Does the relationship between gender identification and health hold when health is operationalized using specific health behaviors or medical diagnoses? Do these patterns apply to both transgender and cisgender people and children as well as adults? How would studies of structural determinants of health also benefit from a gradational

approach to gender measurement? Health disparities research can only begin to address these questions with multiple, theoretically informed, categorical and continuous measures of gender.

SUPPLEMENTAL MATERIAL

Supplemental material is available in the online version of the article.

NOTES

1. Measures of what you think other people think are also called *third-order beliefs* (see e.g., Ridgeway and Correll 2006).
2. We present split sample models for ease of interpretation and to highlight the different results for cisgender women and men on the separate femininity and masculinity scales. However, there are no statistically significant categorical gender differences for any of the controls, so results would be similar if we pooled the sample and estimated interactions with each of the gender scales.
3. This result differs from studies of transgender people that find being perceived as a gender other than the sex one was assigned at birth is associated with negative health behaviors despite this perception aligning with those respondents' view of themselves. The discrepancy could result from different stakes in being perceived as gender nonconforming for transgender people compared to cisgender people, or it could relate to the different dimensions of health (behaviors vs. self-rated health) being assessed.

ACKNOWLEDGMENTS

We are grateful for financial support from the American Sociological Association Fund for the Advancement of the Discipline and the Clayman Institute for Gender Research, and for comments from Bethany G. Everett. A previous version of this paper was presented at the 2018 Population Association of America Annual Meeting.

REFERENCES

Annandale, Ellen, and Kate Hunt. 1990. "Masculinity, Femininity and Sex: An Exploration of Their Relative Contribution to Explaining Gender Differences in Health." *Sociology of Health & Illness* 12(1):24–46.

Bauer, Greta R., Jessica Braimoh, Ayden I. Scheim, and Christoffer Dharma. 2017. "Transgender-inclusive Measures of Sex/Gender for Population Surveys: Mixed-methods Evaluation and Recommendations." *PLoS ONE* 12(5):e0178043.

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

Bird, Chloe E., and Patricia P. Rieker. 1999. "Gender Matters: An Integrated Model for Understanding Men's and Women's Health." *Social Science & Medicine* 48(6):745–55.

Bittner, Amanda, and Elizabeth Goodyear-Grant. 2017. "Sex Isn't Gender: Reforming Concepts and Measurements in the Study of Public Opinion." *Political Behavior* 39(4):1–23.

Blair, Karen L., and Rhea A. Hoskin. 2014. "Experiences of Femme Identity: Coming out, Invisibility and Femmephobia." *Psychology and Sexuality* 6(3):229–44.

Boardman, Jason D. 2004. "Health Pessimism among Black and White Adults: The Role of Interpersonal and Institutional Maltreatment." *Social Science & Medicine* 59(12):2523–33.

Burke, Peter J. 1991. "Identity Processes and Social Stress." *American Sociological Review* 56(6):836–49.

Constantinople, Anne. 1973. "Masculinity-femininity: An Exception to the Famous Dictum?" *Psychological Bulletin* 80(5):389–407.

Cooley, Charles H. 1902. *Human Nature and the Social Order*. New York: Scribners.

Courtenay, Will H. 2000. "Constructions of Masculinity and Their Influence on Men's Well-being: A Theory of Gender and Health." *Social Science & Medicine* 50(10):1385–401.

Emslie, Carol, Kate Hunt, and Sally Macintyre. 2002. "How Similar Are the Smoking and Drinking Habits." *European Journal of Public Health* 12(1):22–28.

Emslie, Carol, Damien Ridge, Sue Ziebland, and Kate Hunt. 2006. "Men's Accounts of Depression: Reconstructing or Resisting Hegemonic Masculinity?" *Social Science & Medicine* 62(9):2246–57.

Federal Interagency Working Group on Improving Measurement of Sexual Orientation and Gender Identity in Federal Surveys. 2016. "Evaluations of Sexual Orientation and Gender Identity Survey Measures: What Have We Learned?" Retrieved December 30, 2018 (https://nces.ed.gov/FCSM/pdf/Evaluations_of_SOGI_Questions_20160923.pdf).

Flores, Andrew R., Jody L. Herman, Gary J. Gates, and Taylor N. T. Brown. 2016. *How Many Adults Identify as Transgender in the United States?* Los Angeles: Williams Institute.

Galdas, Paul M., Joy L. Johnson, Myra E. Percy, and Pamela A. Ratner. 2010. "Help Seeking for Cardiac Symptoms: Beyond the Masculine-feminine Binary." *Social Science & Medicine* 71(1):18–24.

Geist, Claudia, Megan M. Reynolds, and Marie S. Gaytán. 2017. "Unfinished Business: Disentangling Sex, Gender, and Sexuality in Sociological Research on Gender Stratification." *Sociology Compass* 11(4):1–16.

GenIUSS Group. 2014. *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-based Surveys*. Los Angeles: Williams Institute.

Gorman, Bridget K., Justin T. Denney, Hilary Dowdy, and Rose Anne Medeiros. 2015. "A New Piece of the Puzzle: Sexual Orientation, Gender, and Physical Health Status." *Demography* 52(4):1357–82.

Grol-Prokopczyk, Hanna, Jeremy Freese, and Robert M. Hauser. 2011. "Using Anchoring Vignettes to Assess Group Differences in General Self-rated Health." *Journal of Health and Social Behavior* 52(2):246–61.

Hunt, Kate, Mary Kate Hannah, and Patrick West. 2004. "Contextualizing Smoking: Masculinity, Femininity and Class Differences in Smoking in Men and Women from Three Generations in the West of Scotland." *Health Education Research* 19(3):239–49.

Hunt, Kate, Heather Lewars, Carol Emslie, and G. David Batty. 2007. "Decreased Risk of Death from Coronary Heart Disease amongst Men with Higher 'Femininity' Scores: A General Population Cohort Study." *International Journal of Epidemiology* 36(2):269–73.

Jans, Matt, David Grant, Royce Park, Bianca D. M. Wilson, Jody Herman, Gary Gates, and Ninez Ponce. 2016. "Putting the 'T' in LBGT: Testing and Fielding Questions to Identify Transgender People in the California Health Interview Survey." Presented at the 2016 Population Association of America Conference, April 1, Washington, DC.

Jylhä, Marja. 2009. "What Is Self-reported Health and Why Does It Predict Mortality? Towards a Unified Conceptual Model." *Social Science & Medicine* 69(3):307–16.

Krieger, Nancy. 2003. "Genders, Sexes, and Health: What Are the Connections—and Why Does It Matter?" *International Journal of Epidemiology* 32(4):652–57.

Lagos, Danya. 2018. "Looking at Population Health beyond 'Male' and 'Female': Implications of Transgender Identity and Gender Nonconformity for Population Health." *Demography* 55(6):2097–117.

Magliozzi, Devon, Aliya Saperstein, and Laurel Westbrook. 2016. "Scaling up: Representing Gender Diversity in Survey Research." *Socius* 2:1–11.

Markides, Kyriakos S., and Karl Eschbach. 2005. "Aging, Migration, and Mortality: Current Status of Research on the Hispanic Paradox." *Journals of Gerontology Series B: Psychological and Social Sciences* 60(2):68–72.

McGuffey, C. Shawn, and B. Lindsay Rich. 1999. "Playing in the Gender Transgression Zone: Race, Class, and Hegemonic Masculinity in Middle Childhood." *Gender & Society* 13(5):608–27.

Mead, George H. 1934. *Mind, Self, and Society*. Chicago: University of Chicago Press.

Miller, Lisa R., and Eric Anthony Grollman. 2015. "The Social Costs of Gender Nonconformity for Transgender Adults: Implications for Discrimination and Health." *Sociological Forum* 30(3):809–31.

Moore, Sarah E. H. 2010. "Is the Healthy Body Gendered? Toward a Feminist Critique of the New Paradigm of Health." *Body & Society* 16(2):95–118.

Pascoe, Elizabeth A., and Laura Smart Richman. 2009. "Perceived Discrimination and Health: A Meta-analytic Review." *Psychological Bulletin* 135(4):531–54.

Prus, Steven G. 2011. "Comparing Social Determinants of Self-rated Health across the United States and Canada." *Social Science & Medicine* 73(1):50–59.

Read, Jen'nan Ghazal, and Bridget K. Gorman. 2010. "Gender and Health Inequality." *Annual Review of Sociology* 36(1):371–86.

Rider, G. Nicole, Barbara J. McMorris, Amy L. Gower, Eli Coleman, and Marla E. Eisenberg. 2018. "Health and Care Utilization of Transgender and Gender Nonconforming Youth: A Population-based Study." *Pediatrics* 141(3):e20171683.

Ridgeway, Cecilia, and Shelley Correll. 2006. "Consensus and the Creation of Status Beliefs." *Social Forces* 85(1):431–53.

Risman, Barbara J. 2004. "Gender as a Social Structure: Theory Wrestling with Activism." *Gender & Society* 18(4):429–50.

Rubin, Gayle. 1984. "Thinking Sex: Notes for a Radical Theory of the Politics of Sexuality." Pp. 267–319 in *Pleasure and Danger: Exploring Female Sexuality*, edited by C. S. Vance. Boston: Routledge and Kegan Paul.

Schilt, Kristen, and Laurel Westbrook. 2009. "Doing Gender, Doing Heteronormativity." *Gender & Society* 23(4):440–64.

Schnittker, Jason, and Valerio Bacak. 2014. "The Increasing Predictive Validity of Self-reported Health." *PLoS ONE* 9(1):e84933.

Schrock, Douglas, and Michael Schwalbe. 2009. "Men, Masculinity, and Manhood Acts." *Annual Review of Sociology* 35(1):277–95.

Slevin, Kathleen F. 2010. "'If I Had Lots of Money . . . I'd Have a Body Makeover:' Managing the Aging Body." *Social Forces* 88(3):1003–20.

Smith, Tom W., Michael Davern, Jeremy Freese, and Michael Hout. 2017. *General Social Surveys, 1972–2016: Cumulative Codebook* (National Data Program for the Social Sciences Series, no. 24). Chicago: NORC.

Spence, Janet T., and Camille E. Buckner. 2000. "Instrumental and Expressive Traits, Trait Stereotypes, and Sexist Attitudes: What Do They Signify?" *Psychology of Women Quarterly* 24(1):44–53.

Springer, Kristen W., and Dawne M. Mouzon. 2011. "'Macho Men' and Preventive Health Care: Implications for Older Men in Different Social Classes." *Journal of Health and Social Behavior* 52(2):212–27.

Springer, Kristen W., Jeanne Mager Stellman, and Rebecca M. Jordan-Young. 2012. "Beyond a Catalogue of Differences: A Theoretical Frame and Good Practice Guidelines for Researching Sex/Gender in Human Health." *Social Science & Medicine* 74(11):1817–24.

Sumerau, J. E., Lain A. B. Mathers, Alexandra C. H. Nowakowski, and Ryan T. Cragun. 2017. "Helping Quantitative Sociology Come out of the Closet." *Sexualities* 20(5–6):644–56.

Wängnerud, Lena, Maria Solevid, and Monika Djerf-Pierre. Forthcoming. "Moving beyond Categorical Gender in Studies of Risk Aversion and Anxiety." *Politics & Gender*.

Weinberg, Jill, Jeremy Freese, and David McElhattan. 2014. "Comparing Data Characteristics and Results

of an Online Factorial Survey between a Population-based and a Crowdsource-recruited Sample." *Sociological Science* 1:292–310.

West, Candace, and Don H. Zimmerman. 1987. "Doing Gender." *Gender & Society* 1(2):125–51.

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.

Westbrook, Laurel, and Kristen Schilt. 2014. "Doing Gender, Determining Gender: Transgender People, Gender Panics, and the Maintenance of the Sex/Gender/Sexuality System." *Gender & Society* 28(1):32–57.

Willer, Robb, Christabel L. Rogalin, Bridget Conlon, and Michael T. Wojnowicz. 2013. "Overdoing Gender: A Test of the Masculine Overcompensation Thesis." *American Journal of Sociology* 118(4): 980–1022.

Wylie, Sarah A., Heather L. Corliss, Vanessa Boulanger, Lisa A. Prokop, and Bryn S. Austin. 2010. "Socially Assigned Gender Nonconformity: A Brief Measure for Use in Surveillance and Investigation of Health Disparities." *Sex Roles* 63(3–4):264–76.

AUTHOR BIOGRAPHIES

Chloe Grace Hart is a PhD candidate at Stanford University whose main areas of research include gender inequality, health disparities, and sexual violence.

Aliya Saperstein is an associate professor of sociology at Stanford University whose research examines how categories of difference, such as race/ethnicity and sex/gender, are operationalized in survey research and the consequences of those methodological decisions for studies of stratification.

Devon Magliozzi is a PhD candidate at Stanford University whose research considers how gender, race, and class hierarchies are constructed, contested and concealed in settings from workplaces to neighborhoods.

Laurel Westbrook is an associate professor of sociology at Grand Valley State University whose research focuses on the inner workings of the sex/gender/sexuality system, including how knowledge production, gendered violence, and social movements both maintain and alter the system.