Sex Roles (2020) 83:685–705
https://doi.org/10.1007/s11199-020-01141-1

ORIGINAL ARTICLE

Seeing the World in Pink and Blue: Developing and Exploring a New Measure of Essentialistic Thinking about Gender

Karisa Y. Lee[1,2] · Harry T. Reis[1] · Ronald D. Rogge[1]

Published online: 15 April 2020

Abstract

Essentialism, or the belief that certain categories have fundamental, intrinsic, and stable essences, pervasively influences social judgments. Among the many groupings that describe people, gender is the most essentialized category yet relatively little is known about individual differences in gender essentialism. To explore this construct in Study 1 with 2996 U.S. participants, we developed a new measure, named the Gender Essentialism Measure (GEM), that offers two advantages over prior measures: (a) we used Item Response Theory to optimize the measure's psychometrics and (b) we adopted a multidimensional conceptualization, incorporating four core facets of gender essentialism (Biological Determinism, Social Determinism, Immutability, and Inductive Potential). Study 2 used a large U.S. sample ($n = 2803$) to characterize individual differences in essentialistic thinking about gender. Gender essentialism was associated with endorsement of sexism, system-justifying ideologies, relatively inflexible and dispositional thinking about others, and lesser empathic concern and perspective-taking. Studies 3 and 4, using samples of 133 and 118 U.S. participants, respectively, demonstrated that gender essentialism predicts greater acceptance of existing gender disparities. Our research indicates that understanding and addressing the societal and personal impact of gender stereotyping would benefit from going beyond sexism to also considering the role of essentialistic thinking about gender. The GEM offers researchers, practitioners, and policymakers a psychometrically optimized, multi-faceted tool to assess the extent and prevalence of gender-essentialistic beliefs.

Keywords Essentialism · Gender role attitudes · Gender differences · Gender roles · Sex and gender measures

First coined by Medin and Ortony (1989), the term *psychological essentialism* denotes the belief that certain categories have essences, representing one way in which people may think about and categorize members of groups. Beliefs about essentialism have been shown to affect people's thinking about social categories such as race (Williams and Eberhardt 2008), sexual orientation (Haslam and Levy 2006), mental illness (Haslam and Ernst 2002), and social class (Kraus and Keltner 2013). In these domains, holding essentialist beliefs has been shown to influence, for example, automatic categorization processes (Bastian et al. 2011), hierarchy justification (Brescoll et al. 2013), stereotype threat (Dar-Nimrod and Heine 2006), perception of within-group similarities and between-group differences (Martin and Parker 1995; Yzerbyt et al. 2001), prejudice toward essentialized groups (Haslam and Levy 2006; Haslam et al. 2000, 2002; Jayaratne et al. 2006), and the desire to cross group boundaries (Shelton and Richeson 2005; Williams and Eberhardt 2008).

Existing research establishes that among the many social categories that might be used to describe people, gender is the most essentialized category (Gelman et al. 1986; Gelman and Taylor 2000; Prentice and Miller 2006; Taylor 1996). When essentialist thinking is applied to gender, specific male or female "essences" are perceived as strongly determining gender-related characteristics, suggesting that there are stable, internal qualities that characterize the sexes regardless of contextual factors (Mahalingam and Rodriguez 2003). Lay essentialist theories about gender differences are common, ranging from pop culture (e.g., "men are from Mars, women are from

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11199-020-01141-1) contains supplementary material, which is available to authorized users.

✉ Harry T. Reis
reis@psych.rochester.edu

[1] Department of Psychology, The University of Rochester, River Campus Box 270266, Rochester, NY 14627, USA

[2] Facebook, Menlo Park, CA, USA

Venus") to more scientifically grounded studies that attribute myriad behavioral differences between men and women to differences in hormones, neural organization, and other physiological attributes (see Richardson 2013, for a review).

Given their pervasiveness, it is important to better understand the nature and ramifications of holding essentialistic, as opposed to more fluid, views of gender. Research on this subject is limited, even as more general interest in essentialism has grown. The work reported in the present article was designed to improve on existing measures of gender essentialism and then to use this measure to identify key correlates and consequences of holding essentialistic beliefs about gender.

Although existing measures have provided useful preliminary insights, because those measures were largely adapted from measures of personality or race essentialism, they have two potential shortcomings: (a) Because essentialism may have positive, negative, or ambivalent consequences depending on its domain and context (Haslam and Levy 2006; Haslam et al. 2000, a scale intended for one social category may not be optimal for a different social category and (b) existing measures focus almost exclusively on biological determinism, and thus they do not account for other components of essentialism (e.g., discreteness, uniformity, informativeness, immutability, stability, inherence, and exclusivity; Haslam et al. 2000). To advance understanding of this construct, therefore, a more comprehensive measure is needed.

Psychological Essentialism: Basic Concepts

Psychological essentialism is the belief that members of a category share deep, underlying, stable similarities that give rise to observable, meaningful differences that differentiate them from members of other groups (Medin and Ortony 1989). Phrased another way, psychological essentialism refers to the belief that things that look alike share a nontrivial, immutable, and fundamental essence that determines what they are (Keller 2005).

Rothbart and Taylor (1992) posited that people often perceive a shared essence in social groups, thus treating social categories as "natural kinds," despite the fact that these categories may be better understood in terms of context-dependent sociocultural attributes and conventions. Research on social categories such as race and gender has shown that essentialist beliefs are associated with increased stereotype endorsement (Bastian and Haslam 2006; Brescoll and LaFrance 2004; Martin and Parker 1995; Williams and Eberhardt 2008) as well as lesser motivation to change members of essentialized groups, including oneself (Prentice and Miller 2006; Yzerbyt et al. 2001), to eliminate racial disparities, and to bridge group boundaries (Shelton and Richeson 2005; Williams and Eberhardt 2008).

Research on the impact of essentialism on social evaluation has yielded more complex findings. Although essentialism is sometimes associated with greater prejudice (Haslam et al. 2000; Keller 2005; Williams and Eberhardt 2008), other research indicates that these attitudes depend on the particulars of the group being essentialized. For example, when membership in groups, particularly stigmatized groups, has behavioral implications, perceptions of control are reduced, which decreases culpability and often increases tolerance and sympathy. Thus, perceiving a causal genetic role in sexual orientation predicts lesser prejudice toward lesbians and gay men (Haslam and Levy 2006; Horvath and Ryan 2003; Jayaratne et al. 2006). Also, attributing criminal or deviant behavior to uncontrollable causes such as genes is linked to greater tolerance for the perpetrator (Dar-Nimrod et al. 2011). Further, Haslam et al. (2004) found that certain highly essentialized characteristics are rated among the most desirable personality traits (e.g., intelligence, creativity, imaginativeness). Clearly, then, holding essentialist beliefs need not imply negative consequences.

Although biological sex is genetically determined and, with rare exceptions, dichotomous, gender is considered a multifaceted construct that comprises both biological (e.g., genitalia, gonadal hormones) and sociocultural (e.g., gender roles, identity) elements. An essentialist view assumes that human essences underlie a coherent and meaningful set of gender differences in behavior (Prentice and Miller 2006) and that these differences are fixed at birth, stable across the lifespan, immutable, and would not be reduced under different sociocultural conditions. Importantly, laypeople may attribute these essences to nature—to biological, genetic, and/or hormonal factors—or to nurture—durable social influences like cultural socialization or social class. As Wood and Eagly (2012, p. 72) explain, laypeople "might be thinking of nurture or nature (or perhaps both) when they ascribe differing essences to the two sexes."

There is little doubt that gender is a key social category that influences social perception throughout life**.** This tendency starts early—for example, children as young as 3.5 years-old show evidence of gender stereotyping (Reis and Wright 1982). Young children believe that a newborn will develop gender-typical personality traits and behaviors, regardless of the environment in which he or she is raised (Taylor 1996). As for adults, people made stronger inductive inferences about a novel attribute and viewed their own performance as relatively immune to change, when they learned that it distinguished them from a member of the other gender (Prentice and Miller 2006). In another study, participants who read a genetic explanation for a gender difference in "plant identification ability" were more likely to believe that people cannot change and more strongly endorsed gender stereotypes, compared with those who read a sociocultural explanation for this fictive sex difference (Brescoll and LaFrance 2004). Believing that

gender differences in math ability have a genetic basis has been shown to impair women's math performance in much the same way that stereotype threat manipulations do (Dar-Nimrod and Heine 2006).

Conceptualization and Measurement of Gender Essentialism

Existing research typically reduces essentialism to biological or genetic determinism. However, Haslam et al. (2000) showed that essentialist beliefs about social categories are more complex, consisting of two distinct factors. The *natural kind dimension* encompasses beliefs about the immutability, naturalness, discreteness, necessary features, and historical invariance of categories. The second dimension, the *reification factor*, comprises elements of uniformity, informativeness, inherence (i.e., the extent to which the category has an underlying reality), and exclusivity. Haslam et al. (2000, 2002) found that the extent to which social categories are essentialized varies greatly, and that whereas certain groups are highly naturalized, others are highly reified, and still others are both naturalized and reified. Thus we felt that it was important to incorporate both of these elements to more completely understand gender essentialism. In particular, we theorized that belief in the immutability of gender categories and in the inductive potential of biological sex—or, in other words, how much information a person's biological sex is believed to reveal about him or her—would be especially important. Immutability is also important to essentialism because it would be especially pertinent to transgender prejudice.

It is also possible to conceive of an essence as being shaped by nature *or* by nurture (or both). Rangel and Keller (2011) proposed a second and complementary type of psychological essentialism: *social determinism*—the belief that personal characteristics and behavioral tendencies are determined by enduring social factors, such as socialization, social class, or social origin. Belief in social and genetic determinism represent independent but potentially non-opposing aspects of a more general essentialist-belief system because both help explain concepts such as stereotyping, prejudice, perceived group homogeneity, and the perceived inductive potential of "essence-related person information" (Rangel and Keller 2011, p. 12).

Given that people essentialize different social categories to different degrees (Haslam et al. 2000) and given theoretical claims and empirical evidence that the organization of essentialist beliefs is both domain- and context-specific (Haslam et al. 2004; Haslam and Levy 2006), an update to existing measures of gender essentialism is needed. The notion that psychological essentialism is not necessarily limited to genetic determinism has been expressed only a handful of times (Gelman and Hirschfeld 1999; Yzerbyt et al. 1997), despite Haslam et al.' (2000, p. 123) warning against "treat[ing] essentialism as equivalent to the understanding of social categories as natural kinds." Nevertheless, nearly all past research on gender essentialism has used a one-dimensional measure of belief in biological determinism.

The present research addressed three limitations of past work. First, unlike prior studies that reworded measures of general essentialism to apply to gender, we broadened the item pool to more thoroughly capture domain-specific elements of gender essentialism (as we described previously). We then applied item response theory (IRT; Hambleton et al. 1991) so that the new measure would be psychometrically optimized (as we describe in the following). Second, we included items reflecting social determinism (Rangel and Keller 2011). Third, because essentialism should not be reduced to biological determinism, as we discussed, we constructed our measure to include other elements, such as beliefs in immutability and inductive potential. The measure we developed in Study 1 addressed these limitations by including items that tap multiple facets of gender-essentialist beliefs rather than just a single (i.e., biological) component.

Correlates of Essentialist Beliefs about Gender

The findings we reviewed here notwithstanding, research has not yet provided a comprehensive characterization of the attitudinal, cognitive, and personality traits that are associated with essentialist beliefs about gender. Given that gender categorization begins early and pervades many areas of social life, we sought to identify a nomological network of beliefs, traits, and judgments associated with gender-essentialist thinking (Study 2). Studies 3 and 4 examined consequences of gender essentialism for accepting existing gender disparities.

We selected the beliefs and traits investigated in Study 2 either because prior research has linked them to essentializing social categories other than gender or because we believed that they might be particularly relevant to gender essentialism. Space limitations preclude a full review for each set of variables, but we briefly describe the rationale for including these six attributes in our research.

First, to encompass the various distinct contemporary forms of sexism, we included measures of internal and external motivation to respond without sexism (Klonis et al. 2005), hostile and benevolent sexism (Glick and Fiske 2001), and old-fashioned and modern sexism (Swim et al. 1995). Each of these conceptualizations has its own documented etiology and behavioral consequences—for example, benevolent sexism has more positive content than hostile sexism—but because all are based on viewing men and women as distinct categories, we expected all of them to be correlated with essentialism.

Next, because stereotypes also help justify existing social hierarchies (Yzerbyt et al. 1997), we examined whether essentialistic beliefs would similarly satisfy system-justifying motives. This idea was plausible because system-justifying motives imply that differences between social groups are "natural" and unlikely to change (Brescoll et al. 2013). In support of this idea, Morton et al. (2009) found that essentialism is often invoked by members of high-status groups to protect their current advantages when threatened by social change (Keller 2005; Rangel and Keller 2011). We therefore expected that gender essentialism would be positively associated with various types of system-justifying beliefs.

Our third attribute was based on the idea that a central component of essentialism is the belief that group membership is stable and immutable (Yzerbyt et al. 1997). Lay theories about the stability of personality traits differ, spanning a continuum from a belief in stability (entity theorists) to a belief in malleability (incremental theorists; Dweck et al. 1995). Past research has found correlations between essentialist thinking outside the realm of gender and entity theorizing (e.g., Keller 2005) and we therefore expected a similar result for gender essentialism.

A fourth area of our investigation focused on cognitive flexibility. Psychological essentialism can be considered a cognitive heuristic that facilitates categorization and simplifies certain social judgments. Essentialist beliefs imply order and uniformity, and they provide unambiguous answers to questions about how people differ. Research has shown that people who hold essentialist beliefs tend to be higher on various epistemic needs (e.g., desiring definitive and straight-forward answers; Keller 2005; Rangel and Keller 2011). Thus, we expected gender essentialism to be positively related to a variety of constructs (e.g., need for cognitive closure, personal need for structure, intolerance for ambiguity) reflecting cognitive inflexibility and fulfillment of epistemic needs.

A fifth focus was based on the idea that the ability to empathize with a person's circumstances and to take their perspective relies, to some extent, on consideration of this person as a distinct individual rather than an exemplar of a category (Batson et al. 1997). Because holding essentialist beliefs predicts increased perceived group homogeneity (Dar-Nimrod and Heine 2011; Rothbart and Taylor 1992), we expected gender essentialism to be negatively related to perspective-taking.

Finally, Studies 3 and 4 focused on the effect of gender essentialism on acceptance of gender disparities. Because essentialistic thinking may provide ideological support for prevailing status hierarchies, as we described previously, we hypothesized that gender essentialism would be associated with increased acceptance of disparities between men and women, such as those that currently exist in employment, income, and social status. Prior studies have shown comparable findings for race and social-class essentialism and acceptance of existing racial disparities (Williams and Eberhardt 2008) and higher and lower class individuals (Kraus and Keltner 2013).

The Present Research

The present research contributes to the literature by adopting a comprehensive approach to the study of gender essentialism. In Study 1, we develop and validate a new measure of gender essentialism, using item response theory to enhance the measure's psychometric properties (Hambleton et al. 1991). Study 2 is intended to identify personal, relational, and epistemic correlates of gender essentialism in order to establish its role in a person's larger theory of mind. Study 3 explores the implications of essentialistic views of gender for tolerance of gender disparities. Study 4 examines the causal effect of essentialist beliefs on emotional reactions to existing gender disparities through an experimental manipulation. All hypotheses described in the present article were pre-registered in the first author's Master's Thesis proposal as well as filed and approved at the authors' university before any data were collected.

Study 1

In Study 1, we developed the Gender Essentialism Measure (GEM), using a combination of exploratory factor analyses (EFA) and item response theory. We began by generating an initial pool of 55 items, following the idea, explained previously, that the structure of essentialism is context-dependent and multidimensional, encompassing elements such as high inductive potential, sharp category boundaries, historical invariance, immutability, and homogeneity. We then collected pilot data from 395 respondents ($n = 221$, 56% female; $M_{age} = 30.1$, $SD = 13.3$, range 18–68; $n = 332$, 84% Amazon Mechanical Turk (MTurk), $n = 63$, 16% psychology subject pool), and we ran EFAs to identify the dominant dimensions within this preliminary item pool. These results suggested a four-factor solution, identifying four distinguishable facets: biological/genetic determinism (e.g., "In most cases, gender differences can be traced back to biological causes"), inductive potential (e.g., "It is possible to know about many aspects of a person once you know his or her biological sex"), immutability (e.g., "A person's gender is fixed at birth"), and social determinism (e.g., "In most cases, gender differences can be traced back to socio-cultural causes"). Based on the EFA results, items with poor loadings were dropped and additional items were written to create a revised pool of 61 items which were then given to a new sample of 2996 respondents.

Study 1 augmented classical test theory approaches with IRT to create a psychometrically optimized multidimensional

measure of essentialism. When used with large and diverse samples, IRT identifies items with the highest precision (i.e., lowest noise) across the broadest range of the construct being assessed, and it is likely to yield measures that function well in a wide range of samples (Hambleton et al. 1991). A known disadvantage of classical test theory is the assumption that precision is constant for people at all levels of a given trait; however, it is clear that many scales violate this assumption. IRT recognizes that measurement precision varies across the entire range of a trait, and thus it estimates latent trait scores for each respondent and, based on those scores, calculates response curves for each item, which reveal how much information or precision each item offers across the entire range of the latent trait being measured. These item information curves indicate exactly how much information each item provides for assessing the construct of interest and can be summed to create a test information curve, which displays how informative the scale is as a whole (see Hambleton et al. 1991, for more detailed information about IRT and a discussion of its advantages over classical test theory). Thus, our use of IRT in our research offers two specific psychometric benefits. First, the resulting scale has maximum variance (which increases the scale's ability to produce substantive findings). Second, scales optimized in this way maximize the ability to distinguish among individuals across the full range of the measure—that is, distinguishing among individuals who score in the low range or in the high range on this measure.

Method

Participants

A total of 3090 individuals responded to an online survey, resulting in a final sample of 2996 respondents after data cleaning. Participants were predominantly female ($n = 2064$, 68.9%), with a mean age of 39.5 years ($SD = 11.6$, range = 18–71). Most of the participants had attended college ($n = 773$, 25.8% some college; $n = 1042$ 34.8% bachelor's degree; $n = 1049$, 35.0% graduate degree), and $n = 132$ (4.4%) had completed high school or less. A plurality of the respondents (41.8%) identified as Democratic, with 13.9% Republican, 33.6% Independent, and 10.7% other. Religiosity was distributed fairly evenly, with 33.5% stating that religion was "not at all" important to them, 22.7% "slightly," 19.4% "moderately," and 24.4% "very."

Procedure

Prior to data collection, all procedures involved in the present research were reviewed and approved by the University of Rochester Research Subjects Review Board. Respondents had to be at least 18 years-old and were recruited via the ResearchMatch.org online community ($n = 2834$, 91.7%) and MTurk ($n = 256$, 8.3%). The survey took 15–20 min to complete and contained the following questions: A pool of 61 gender essentialism items, the infrequency subscale of the Attentive Responding Scale (ARS-18) (Maniaci and Rogge 2014) to assess attentiveness, and seven demographic questions. As a recruitment incentive, feedback on the extent to which respondents essentialize gender was given at the end of the survey.

We began with an initial pool of 61 items, selected from several existing scales, items adapted from existing scales, and items written by the authors to reflect dimensions of essentialism other than genetic/biological determinism. We used or adapted 32 items from the Biological Basis, Discreteness, and Informativeness Scales (Bastian and Haslam 2006), the Belief in Genetic Determinism Scale (Keller 2005), the Belief in Social Determinism Scale (Rangel and Keller 2011), the Causes of Group Differences Questionnaire (Martin and Parker 1995), the Gender Theory Questionnaire (Coleman and Hong 2008), the Race Conceptions Scale (Williams and Eberhardt 2008), and the Lay Theory of Race Scale (No et al. 2008).

We wrote 29 additional items to reflect aspects of essentialist beliefs about gender not covered by the aforementioned scales, based on Haslam et al.'s (2000) nine elements of essentialistic beliefs, specifically (a) informativeness (sample items: "Ultimately, you are born a man or a woman, and this affects most of the things you do in life" and "Sex has little to do with a person's abilities and personal qualities" [reverse-scored]), (b) necessity (sample item: "The presence of a Y chromosome in men gives rise to many 'male' characteristics and abilities"), (c) immutability (sample item: "Even if people undergo sex-change surgery, at heart they are still the sex they were born"), (d) uniformity (sample item: "Members of the same sex are usually more similar to each other than to members of the opposite sex"), and (e) stability (sample item: "It is impossible to imagine a culture that treats men and women exactly the same"). Participants responded to all items on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The relevance of all items to the construct of essentialism was determined by the authors.

Data Cleaning

Prior to analysis, the dataset was subjected to two rounds of cleaning, using procedures recommended by Funk and Rogge (2007) and determined before examining the data. First, we used the six-item infrequency subscale of the ARS-18 to identify responses that were invalid due to inattention. Respondents rated each item on a 5-point Likert-type scale and received higher scores for each increasingly implausible response. These scores were summed and a cut-score of 6 was used to identify 71 (2.3%) of the initial 3090 responses as

invalid. Next, 23 (.8%) responses were identified as multivariate outliers, based on Mahalanobis distances (Tabachnick and Fidell 2013) and were deleted. These two steps left a final sample of 2996 participants.

Results

Identifying Dimensions for IRT

To identify the dimensional structure within the pool of 61 items, we ran a preliminary EFA with principal axis factoring and an oblimin rotation (allowing the factors to correlate). This step ensures that items being subjected to IRT analyses together were indeed measuring common constructs. A scree plot suggested a five-factor solution, accounting for 42% of the variance. However, when five factors were extracted, only four factors emerged that contained multiple items with reasonably strong factor loadings (i.e., loading at least .40 or more). The four factors extracted corresponded with the dimensions suggested by the pilot data: Biological/Genetic Determinism (gender differences are fundamental biologically determined), Social Determinism (gender differences are shaped by socio-cultural factors), Immutability (gender is a discrete and immutable category with sharply defining features), and Inductive Potential (i.e., the belief that knowledge of a person's gender serves as a rich source of information and inferences). (Results from this initial EFA can be found in the online supplement: Section 1, Table 1s.)

IRT Analyses and Item Selection

We subjected each set of items (loading at least .40 on their respective factors) to a separate IRT analysis using Samejima's (1997) graded response model in Multilog 7.0 (Thissen et al. 2002) with marginal maximum-likelihood estimation. The six most effective items were selected for each subscale of the GEM based on the information they provided as assessed by IRT (i.e., their relative ability to discriminate among participants who score at different levels on the subscale being assessed). This yielded a 24-item version of the GEM, a scale that is optimized in two ways: (a) it creates a composite with the highest possible level of variance and (b) it maximizes the ability to discriminate between individuals who score low and high on the trait in question. (Further details about this analysis, including item information curves for all four subscales, are presented in the online supplement: Section 1, Figs. 1 s and 2 s.)

Factor Analyses

We performed factor analyses to verify the underlying structure of the GEM. Following best practices (Fabrigar et al. 1999), we randomly split the sample into two halves to run Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) on separate sets of respondents in order to cross-validate the dimensional results. Table 1 presents results of an EFA using principal axis factoring with an oblimin rotation in the first random half-sample ($n = 1511$; 1044, 69.1% women). Both the scree plot and the Kaiser–Guttman criterion suggested a four-factor solution that accounted for 58.9% of the variance. Factor pattern coefficients, listed in the four columns under the heading EFA in Table 1, show that the items all loaded strongly on their intended factors, and virtually not at all with the other three factors (thereby confirming the discrete four-factor structure of the GEM).

To verify the dimensional structure of the GEM, we conducted a CFA on the 24 items of the GEM in the second sample half of Study 1 ($n = 1485$; 1038, 69.9% women) using Mplus 7.11 (Muthén and Muthén 2012). We utilized four widely used fit indices to determine the acceptability of model fit: (a) the model Chi-squared statistic (the primary index of absolute model fit), (b) the standardized root-mean-square residual (SRMR; values less than .08 suggest acceptable fit), (c) Bentler's comparative fit index (CFI; values above .90 suggesting acceptable fit), and (d) the root-mean-square error of approximation (RMSEA; values less than .07 indicate acceptable fit). The CFA model with four subscales demonstrated excellent fit: $\chi^2(246) = 62.3$, RMSEA = .051, 95% CI [.049, .053], CFI = .940, SRMR = .041. In contrast, a model placing all 24 items on a single global subscale demonstrated unacceptable fit: $\chi^2(252) = 12{,}168$, RMSEA = .126, 95% CI [.124, .128], CFI = .625, SRMR = .108. As the two rightmost columns of Table 1 show, all the items yielded strong path coefficients to their respective latent factors in the four-factor CFA model. Furthermore, the four latent dimensions of gender essentialism were correlated very similarly in both the EFA and CFA subsamples, as Table 2 shows. Thus, the CFA sample results cross-validated results from the EFA sample, establishing that the GEM has a stable factor structure.

The CFA results showed that the GEM subscales of Inductive Potential, Immutability, and Biological Determinism were reasonably correlated (rs between .51 and .80), indicating that they shared between 26 to 64% of their variance. This suggests that in addition to using each subscale as a distinct construct, researchers may wish to collapse these three scales into a composite reflecting essentialist beliefs (excluding socially based essentialism) about gender. In the interest of parsimony, these scales will be combined and referred to as gender essentialism throughout the rest of this paper. Given the smaller correlations between Social Determinism and the other three scales, that subscale will be retained as a separate construct in the analyses that follow. Combining data from the EFA and CFA subsamples, Cronbach's alpha for the four subscales were: Inductive Potential, $\alpha = .83$; Immutability, $\alpha = .89$; Biological Determinism, $\alpha = .85$;

Table 1 *GEM correlational structure with EFA and CFA in separate random sample halves, study 1*

GEM Items	EFA Pattern Coefficients $n = 1511$				CFA Path Coefficients $n = 1485$	
	F1	F2	F3	F4	β	S.E.
(a) FACTOR 1 (F1): Inductive Potential						
The innate properties of a person's gender determine what the person is like.	.71	−.05	.05	−.02	.72	.011
When getting to know a person, it is possible to get a good picture of the kind of person they are based on their biological sex.	.68	−.05	.02	−.12	.59	.014
It is easy to accurately predict a person's gender from his or her abilities and interests.	.63	.04	.02	.12	.69	.011
Knowing about someone's personality gives a good indication of the person's gender.	.61	.05	−.08	.09	.61	.013
What a person is like (such as his or her abilities and traits) is deeply ingrained in his or her gender, and cannot be changed much.	.55	−.02	.04	.13	.65	.012
It is possible to know about many aspects of a person once you know his or her biological sex.	.51	.00	.13	.17	.72	.011
(b) FACTOR 2 (F2): Social Determinism						
The social background a person comes from is strongly reflected in the development of the person's gender-related attributes.	−.01	.70	.03	.05	.63	.013
A person's gender-related behavior is largely the product of his or her social origin.	−.07	.70	.03	−.04	.73	.011
In most cases, gender differences can be traced back to socio-cultural causes.	.00	.66	−.05	−.16	.75	.010
The type of social environment a person grows up in is evident in the development of the person's gender-related behavior.	−.05	.64	.02	.08	.60	.014
When the sexes differ in some behavior, that difference is mainly due to how they have been treated by their parents and society.	.02	.62	.02	−.12	.70	.011
The way a man or woman turns out depends largely on what society expects of him or her.	.09	.62	−.08	.02	.60	.014
(c) FACTOR 3 (F3): Immutability						
A person's gender is clearly defined; you are either a female or a male	.02	.02	.84	.01	.84	.007
A person's gender is fixed at birth.	−.04	.00	.80	.04	.79	.008
No one can truly change his or her sex—you are who you are.	.06	.04	.78	.01	.79	.008
Even if people undergo sex-change surgery, at heart they are still the sex they were born.	.10	.02	.75	−.06	.74	.009
Gender is not set in stone and can be changed. (Reverse coded)	−.01	−.04	.72	−.02	.72	.010
A person's gender is not easily defined. (Reverse coded)	−.10	−.06	.68	.07	.70	.011
(d) FACTOR 4 (F4): Biological Determinism						
Many forms of men's and women's behavior are biologically determined.	−.07	−.02	−.02	.80	.68	.011
When trying to understand differences in men's and women's behavior, one should always look first to biology.	.09	−.07	.07	.61	.75	.010
Behavioral differences between men and women are largely determined by their genetic background.	.16	−.08	−.01	.60	.72	.010
In most cases, gender differences can be traced back to biological causes.	.05	−.03	.10	.59	.70	.011
To a large extent, a person's sex biologically determines his or her abilities and traits.	.33	.03	.09	.44	.73	.010
Men and women are fundamentally good at different things.	.10	−.03	.19	.42	.64	.012

Note. Table entries are factor loadings, with factor-definitions printed in black. As discussed in the text, these subscales can be used as four subscales or two. CFA model demonstrated excellent fit: $\chi^2(246) = 62.3$, RMSEA = .051, 95% CI [.049, .053], CFI = .940, and SRMR = .041. Respondents rated all items on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*)

Table 2 *Correlations Among the EFA Factors (below Diagonal) and CFA Latent Variables.(above Diagonal), Study 1.*

	F1 *r*	F2 *r*	F3 *r*	F4 *r*
FACTOR 1 (F1): Inductive Potential	–	−.25	.51	.80
FACTOR 2 (F2): Social Determinism	−.15	–	−.32	−.45
FACTOR 3 (F3): Immutability	.40	−.28	–	.63
FACTOR 4 (F4): Biological Determinism	.59	−.34	.54	–

Social Determinism, $\alpha = .83$. (A ready-to-use version of the GEM with scoring instructions can be found in the online supplement.)

Study 2

Study 2 was intended to explore the personal, relational, and epistemic correlates of gender essentialism with the newly developed GEM. In order to describe the role of essentialist beliefs about gender in a person's conception of the social world, we constructed a nomological network of attitudes and beliefs that may be associated with essentialized views of gender—that is, we examined correlations between the GEM scores and a constellation of variables that should help elucidate the place of essentialism within a person's larger cognitive and attitudinal approach to the social world.

Method

Participants

Given the large number of constructs we sought to examine, to minimize participant burden, we created five versions of our survey. Each version was pretested to take approximately 25 min to complete. Participants were randomly assigned to one of the five versions. Some of the more conceptually important scales were included in more than one version of the survey.

Prior to data collection, all procedures involved in our research were reviewed and approved by the University of Rochester Research Subjects Review Board. We initially set a sample-size target of 350 individuals per subsample in order to achieve power of .95 (two-tailed alpha = .05) to detect a correlation of .20 (GPower). In all, 3082 individuals participated. After cleaning (details follow), the final sample consisted of 2803 respondents (76.9% female; M_{age} = 43.4 years, *SD* = 13.9, range = 18–86,). There were 556 respondents for Version A, 561 for Version B, 561 for Version C, 586 for Version D, and 539 for Version E. Participants were recruited online from ResearchMatch.org (93.5%) or other online sources (2.6%) or were University of Rochester students who received course extra-credit (3.7%). A majority had attended college (27.1% some college, 34.3% bachelor's degree, 34.0% graduate degree), whereas 4.6% completed high school or less. A plurality of the respondents (44.5%) identified as Democrats, with 17.7% Republican, 28.1% Independent, and 9.7% other. Religiosity was distributed with 30.3% describing religion as "not at all" important, 23.2% "slightly," 20.1% "moderately," and 26.5% "very." There were no significant differences on any of these demographic variables as a function of which survey version was completed.

Measures

Participants in all five samples completed demographic questions, a subset of self-report scales designed to assess a variety of different constructs (discussed previously) as well as the newly developed GEM scale. We used multiple scales to address related constructs to be sure that our findings reflected the construct in question and not the idiosyncrasies of any particular measure.

Sexism

The Internal Motivation to Respond Without Sexism Scale (IMS-S) and External Motivation to Respond Without Sexism Scale (EMS-S) are validated measures that assess both sources of motivation to respond without sexism (Klonis et al. 2005). A sample item of the IMS-S is: "I attempt to act in nonsexist ways toward women because it is personally important to me." A sample item of the EMS-S is: "If I acted sexist toward women, I would be concerned that others would be angry with me." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Higher scores on both scales indicate stronger motives to respond with sexism. Scores were averaged across items and Cronbach's alphas were .82 and .78 for the internal and external motivation scales, respectively. Both scales were included on Versions C and D of the survey.

The Ambivalent Sexism Inventory taps two positively correlated components of sexism: hostile and benevolent sexism (Glick and Fiske 2001. The hostile sexism subscale assesses sexist antipathy ("Women seek to gain power by getting control over men"). The benevolent sexism subscale assesses a subjectively positive orientation toward women ("Women should be cherished and protected by men"). Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged across items; higher scores on both scales indicate higher levels of each form of sexism. Cronbach's alphas were .88 and .85 for the hostile and benevolent subscales, respectively, and both subscales were included on all five survey versions.

The Old-Fashioned and Modern Sexism Scales measure two components of sexism (Swim et al. 1995). The former is typified by endorsement of traditional gender roles and differential treatment of men and women ("When both parents are employed and their child gets sick at school, the school should call the mother rather than the father"), whereas the latter is marked by resentment about perceived special treatment of women and denial of continuing discrimination ("Over the past few years, the government and news media have been showing more concern about the treatment of women than is warranted by women's actual experiences"). Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged across items; higher scores on both scales indicate higher levels of sexism. Cronbach's alphas were .61 and .86 for the old-fashioned and modern sexism scales, respectively, and both scales were included on Versions A and B.

System-Justifying Beliefs

Altemeyer's (1996) Right-Wing Authoritarianism Scale taps three aspects of authoritarian beliefs: conservatism, authoritarian submission, and aggression and was included on Version C. Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged across items so that higher scores indicated greater right-wing authoritarianism. A sample item is "Our country desperately needs a mighty leader who will do what has to be done to destroy the radical new ways and sinfulness that are ruining us." Cronbach's alpha was .95.

The Social Dominance Orientation Scale measures preferences for inequality among social groups and was included on Version B (Pratto et al. 1994). A sample item reads: "Some groups of people are simply not the equals of others." Respondents answered items using a 1 (*very negative*) to 7 (*very positive*) scale. Scores were averaged across items so that a higher score indicated greater endorsement of social inequality. Pratto et al. (1994) provide evidence of the scale's predictive and discriminant validity. Cronbach's alpha was .89.

Patriotism and nationalism were assessed with scales created by Kosterman and Feshbach (1989). A sample patriotic item reads: "I feel a great pride in that land that is our America," and a sample nationalist item is: "Generally, the more influence America has on other nations, the better off they are." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. For each scale, scores were averaged across items so that a higher score indicated higher levels of patriotism and nationalism, respectively. Cronbach's alphas were .92 and .85 for the two scales, respectively. These measures were included in Versions A and E.

The Free-Market Ideology Scale ($\alpha = .83$) looks at the extent to which a person's worldview endorses free-market economics (Lewandowsky et al. 2013). A sample item is: "An economic system based on free markets unrestrained by government interferences automatically works best to meet human needs." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged across items so that a higher score indicated greater endorsement of free-market ideology. This scale was included on Versions A and E of the survey.

Cultural and economic conservatism were assessed with versions of the Middendorp Cultural and Economic Conservatism Scale adapted by Crowson (2009). A sample item from the 12-item cultural conservatism scale ($\alpha = .79$) reads: "Abortion should remain illegal under all circumstances," whereas a sample item from the nine-item economic conservatism scale ($\alpha = .87$) is: "Our country can only get ahead if the government gives the industry free reign to control its own affairs." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. For each scale, scores were averaged across items so that a higher score indicated higher levels of cultural and economic conservatism, respectively. Both scales were included on survey Versions C and D.

The Protestant Ethic Scale taps devotion to work, individual achievement, and discipline, and it has been strongly related to anti-Black prejudice (Katz and Hass 1988). A sample item reads: "Most people who don't succeed in life are just plain lazy." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater endorsement of the Protestant Ethic. Cronbach's alpha was .77. We included this scale on Versions B and C.

Implicit Person Theories

The Person Theory Measure, developed and validated by Dweck et al. (1995), assesses belief in the stability of human characteristics. A sample item is: "Everyone is a certain kind of person and there is not much that can be done to really change that." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater endorsement of a fixed mindset about human qualities. Cronbach's alpha was .90. This scale was included on Version C of the survey.

Poon and Koehler's (2006) measure of dispositional inference assays the tendency to predict behavior from given traits, to infer a trait from an observed behavior, and to assume temporal and cross-situational stability of traits. A sample item reads: "Person A behaved in a more polite way than Person B in a particular situation. What is the probability that Person A is more strongly characterized by the trait polite than Person B?" Respondents provided a probability from 0 to 100%. Scores were averaged so that a higher score indicated a greater tendency to make dispositional inferences regarding four

traits: intelligence (Version A), friendliness (Version B), ambition (Version C), and politeness (Version D). Cronbach's alphas for the four versions of this scale were .76, .81, .80, and .77, respectively.

Cognitive Flexibility

The Epistemic Belief Inventory ($\alpha = .79$) assesses beliefs about simplicity and directness in the nature and acquisition of knowledge (Schraw et al. 2002) and was included on Version A. A sample item is: "Too many theories just complicate things." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater endorsement of a simplistic, inflexible orientation to knowledge.

The Need for Cognitive Closure Scale measures the need for definitive and simple answers (Webster and Kruglanski 1994) and was included on Version A. We used a validated short version of the scale (Roets and Van Hiel 2011); a sample item reads: "I dislike questions which could be answered in many different ways." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater need for cognitive closure. Cronbach's alpha was .84.

The Personal Need for Structure Scale assesses the desire to structure and organize the environment (Neuberg and Newsom 1993; Thompson et al. 1993) and was assessed on Version D. A sample item is: "I enjoy having a clear and structured mode of life." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater need for structure. Cronbach's alpha was .86.

Altemeyer's (2002) DOG Scale measures dogmatism, a person's tendency toward unjustified certainty, and it was included on Version E. A sample item reads: "The things I believe in are so completely true, I could never doubt them." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater need for cognitive closure. Cronbach's alpha was .93.

The Need for Cognition Scale is a widely used scale tapping the tendency to engage in and enjoy thinking. We used a validated short version of the scale (Cacioppo et al. 1984) on Version B. One reverse-scored sample item is: "I like tasks that require little thought once I've learned them." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated greater need for cognition. Cronbach's alpha was .90.

The Intolerance for Ambiguity Scale assesses the general tendency to perceive ambiguous material or situations as threatening (MacDonald 1970) and was included on Version C. A sample item is: "There's a right way and a wrong way to do almost everything." Items were answered using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated more intolerance for ambiguity. Cronbach's alpha was .72.

The Need to Evaluate Scale measures one's tendency to engage in evaluative responding (Jarvis and Petty 1996) and was assessed on Versions D and E. A sample item reads: "I like to have strong opinions even when I am not personally involved." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that a higher score indicated higher evaluative needs. Cronbach's alpha was .85.

Empathic Concern and Perspective-Taking

The empathic concern subscale of the Interpersonal Reactivity Index (Davis 1980) assesses the tendency to feel compassion and concern for unfortunate others and was included on Versions A and B. A sample item is: "I often have tender, concerned feelings for people less fortunate than me." The perspective-taking subscale measures the tendency to spontaneously adopt others' point of view and was included on surveys C and E; a sample item states: "I sometimes try to understand my friends better by imagining how things look from their perspective." Respondents answered items using a 1 (*strongly disagree*) to 7 (*strongly agree*) scale. Scores were averaged so that higher scores indicated greater empathic concern and perspective-taking, respectively. Cronbach's alphas were .76 and 80, respectively.

Data Cleaning

Prior to analysis, the dataset was subjected to two steps of data cleaning, again following pre-established procedures. First, as in Study 1, we used six items from the infrequency subscale of the ARS-18 to identify responses that reflected inattention (Maniaci and Rogge 2014). This step classified 276 (9.0%) of the initial 3082 responses as invalid. Next, three responses (.1%) were identified as multivariate outliers, based on Mahalanobis distances (Tabachnick and Fidell 2013). These two steps eliminated 279 (9.1%) respondents, leaving a final sample of 2803.

Results

The six items comprising each of the four GEM subscales showed strong coherence (Biological Determinism: $\alpha = .81$; Discreteness/Immutability: $\alpha = .88$; Inductive Potential: $\alpha = .81$; Social Determinism: $\alpha = .81$). As discussed in Study 1, given the overlap of the first three subscales of the GEM, we combined these subscales into a single 18-item GEM composite ($\alpha = .90$), representing essentialist beliefs

about gender. (A full table of results including correlations with each of the four subscales of the GEM can be found in the online supplement, Section 2, Table 2 s.)

A *t*-test revealed a significant difference in mean GEM composite scores as a function of gender, $t(2786) = -10.36$, $p < .001$, $\eta_p^2 = .037$. Men scored higher ($M = 11.01$, $SD = 3.50$) than women ($M = 9.46$, $SD = 3.27$). There was no corresponding difference in Social Determinism scores ($M_{men} = 4.37$, $SD = 1.24$; $M_{women} = 4.37$, $SD = 1.27$), $t(2786) = .04$, $p = .970$, $\eta_p^2 = .000$. For all of the following analyses, we tested for gender differences by including the various predictors and gender in regression analyses. When we did this, none of the correlational results became nonsignificant (nor did they change much in magnitude). Additionally, for each of the 35 constructs includd in Study 2, we examined interactions between participant gender and the three-subscale GEM composite as well as between participant gender and the Social subscale. Only five of 70 (7.1%) gender interactions were significant, and none of them accounted for more than 1.4% of the variance. This level of difference seems likely to be due to chance, so we do not address them here but we report these interactions in the online supplement (Section 2, Table 3s).

Analysis of variance (ANOVA) revealed significant differences in GEM scores as a function of political affiliation. As shown in Table 3, Tukey's post-hoc test revealed that Republicans scored significantly higher than Democrats, Independents, and Others. In fact, all four groups differed significantly from one another (except the comparison of Independents to Others). An ANOVA revealed a corresponding difference in Social Determinism as a function of political affiliation. Tukey's post-hoc test showed that Republicans scored significantly lower than Democrats and Independents, but not Others. No other comparisons were significant.

An ANOVA revealed significant differences in mean GEM composite scores as a function of geographic locale. Tukey's post-hoc tests revealed that respondents living in the Southern United States scored significantly higher than those living in the West, the Northeast, and the Midwest. None of these latter regions differed significantly from one another. There was no difference in Social Determinism scores as a function of where participants lived.

Finally, religious importance was positively correlated with scores on the three-subscale GEM composite, $r(2800) = .34$, $p < .001$, and negatively correlated with scores on the Social Determinism subscale, $r(2800) = -.07$, $p < .001$. Education level was negatively correlated with scores on the GEM composite, $r(2800) = -.11$, $p < .001$, and positively correlated with Social Determinism scores, $r(2800) = .07$, $p < .001$.

Sexism

As shown in Table 4a, the three-subscale GEM composite was negatively correlated with the Internal Motivation to Respond Without Sexism Scale and positively correlated with the External Motivation to Respond Without Sexism Scale. Social Determinism was positively correlated with both. These results indicate that individuals with a more essentialistic view of gender are more likely to feel external—but not internal—pressure to avoid displaying sexism. However those endorsing Social Determinism report feeling both kinds of pressure to be nonsexist. Additionally, the GEM composite was strongly positively correlated with hostile, benevolent, old-fashioned, and modern sexism. Social Determinism, on the other hand, was negatively related to hostile and modern sexism, but unrelated to benevolent and old-fashioned sexism. Thus, essentialistic thinking about gender is consistently related to the varied forms of sexism that we studied such that stronger endorsement of essentialistic beliefs are associated with higher levels of sexism.

System-Justifying Beliefs

As expected, the three-subscale GEM composite was strongly and positively correlated with system-justifying beliefs (see

Table 3 *GEM composite and social determinism subscales as a function of political affiliation and U.S. region of habitation, study 2*

(a) Political Affiliation

	Republican	Democrat	Independent	Other			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*(3, 2798)	*p*	η_p^2
3-Subscale GEM Composite	12.12_a (3.18)	8.87_b (3.14)	9.85_c (3.11)	9.72_c (3.59)	123.74	< .001	.117
Social Determinism Subscale	4.21_a (1.17)	4.44_b (1.32)	4.40_b (1.20)	$4.37_{a,b}$ (1.31)	3.90	.009	.004

(b) Region

	South	West	Northeast	Midwest			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*(3, 2798)	*p*	η_p^2
Three-subscale GEM Composite	10.27_a (3.57)	9.39_b (3.18)	$9.59_{b,c}$ (3.34)	9.77_c (3.32)	8.56	< .001	.009
Social Determinism Subscale	4.32_a (1.31)	$4.39_{a,b}$ (1.23)	4.47_b (1.26)	$4.36_{a,b}$ (1.23)	1.79	.147	.002

Note. Means sharing the same subscript across a row are not significantly different from each other (Tukey's HSD, $p \geq .05$)

Table 4 *Bivariate correlations of the GEM composite and social determinism subscale with sexism and system-justifying beliefs, study 2*

		3-Subscale GEM Composite	Social Determinism Subscale
Construct	Sample	*r*	*r*
(a) Sexism			
Internal motivation to respond without sexism	C, D	−.36***	.12***
External motivation to respond without sexism	C, D	.24***	.07*
Hostile sexism	A, B, C, D, E	.54***	−.08***
Benevolent sexism	A, B, C, D, E	.58***	−.03
Old-Fashioned sexism	A, B	.41***	−.03
Modern sexism	A, B	.53***	−.14***
(b) System-Justifying Beliefs			
Right-Wing authoritarianism	C	.62***	−.11*
Social dominance orientation	B	.46***	−.06
Patriotism	A, E	.29***	−.17***
Nationalism	A, E	.53***	−.07*
Free-Market ideology	A, E	.36***	−.18***
Cultural conservatism	C, D	.63***	−.04
Economic conservatism	C, D	.39***	−.15***
Protestant ethic	B, C	.38***	.06*

Note. Sample A, $n = 556$; Sample B, $n = 561$; Sample C, $n = 561$; Sample D, $n = 586$; Sample E, $n = 539$

*$p < .05$. ***$p < .001$.

Table 4b). The largest correlations were found between the GEM composite and right-wing authoritarianism and cultural conservatism, followed by nationalism, social dominance orientation, economic conservatism, Protestant ethic, free-market ideology, and patriotism. Thus, as predicted, holding an essentialistic view of gender is positively associated with valuing existing social hierarchies and the systems that keep them in place. In contrast, Social Determinism was negatively correlated with most of these beliefs, indicating that social determinism went along with a less conservative ideology. This GEM subscale was not significantly related to social dominance orientation and cultural conservatism, and it was actually positively correlated with Protestant ethic (although the magnitude of this effect is small). Because Protestant work ethic is the belief that a person's hard work and effort is the sole predictor of success in life, it may serve the slightly more indirect function of rationalizing social inequality, as opposed to constructs like social dominance orientation and right-wing authoritarianism, which represent a more extreme and outright endorsement of social hierarchies.

Implicit Person Theories

As shown in Table 5, the three-subscale GEM composite was positively related to person theorizing (belief in the stability of human characteristics), whereas the Social Determinism subscale was not. However, both the GEM composite and the Social Determinism subscale were positively correlated with lay dispositionism. Thus, both types of essentialistic thinking about gender relate to the tendency to use a person's traits to explain and predict behavior and to expect cross-situational and temporal stability of traits.

Cognitive Flexibility

Table 5 also reports correlations that support our hypothesis that essentialistic beliefs may simplify categorization and social judgment, as reflected in epistemic beliefs, dogmatism, and intolerance for ambiguity. Results for the Social Determinism subscale were more ambiguous, correlating significantly only with intolerance for ambiguity. In this realm, social deterministic beliefs were unrelated to most indicators of characteristic essentialist thinking—in some cases, even serving as a way to reduce some of the ambiguity that essentialist thinkers seem to find uncomfortable and to make quick and definitive evaluations about their social world. Taken together, these results indicate that holding an essentialist view of gender is associated with a strong preference for stability, order, structure, and unambiguous and simple answers.

Table 5 *Bivariate correlations of the GEM composite and social determinism subscale with additional indicators, study 2*

		3-Subscale GEM Composite	Social Determinism Subscale
Construct	Sample	*r*	*r*
(a) Implicit Person Theories			
Person theory measure	C	.24***	.03
Lay dispositionism	A, B, C, D	.17***	.10***
(b) Cognitive Flexibility			
Epistemic beliefs	A	.60***	−.04
Need for cognitive closure	A	.12**	.00
Personal need for structure	D	.11*	.04
Dogmatism	E	.51***	.01
Need for cognition	B	−.25***	.03
Intolerance for ambiguity	C	.35***	.11*
Need to evaluate	D, E	.06†	.05
(c) Empathic Concern and Perspective-Taking			
Empathic concern	A, B	−.12***	−.08*
Perspective-taking	C, E	−.22***	−.00

Note. Sample A, $n = 556$; Sample B, $n = 561$; Sample C, $n = 561$; Sample D, $n = 586$; Sample E, $n = 539$
* $p < .05$. ** $p < .01$. *** $p < .001$.

Empathic Concern and Perspective-Taking

Table 5 shows that the three-subscale GEM composite was negatively related to both empathic concern and perspective-taking, consistent with the more impersonal and inflexible way that people who essentialize gender tend to think about others. These results are an important first step in extending research on gender essentialism into the field of close relationships because many important interpersonal processes rely on the ability to feel compassion for and take the perspective of relational partners. The Social Determinism subscale was also negatively related to empathic concern, but not significantly related to perspective-taking.

Discussion

It will be useful to summarize these many results to provide an integrated account of how an essentialistic view of gender relates to a person's larger worldview. People who essentialize gender were more prone to stereotypic biases in social judgments, specifically with regard to negative stereotypes about women's abilities and roles, consistent with their greater endorsement of system-justifying and hierarchy-enhancing ideologies (e.g., the Protestant work ethic, conservative sociopolitical attitudes, social dominance orientation). This pattern of results highlights the role of essentialist beliefs in helping people explain gender differences that they may encounter in their lives and further in justifying observed social inequities (the subject of Studies 3 and 4). This status-legitimizing function is also supported by the finding that men essentialize gender significantly more than women do, a result that is consistent with research showing that members of high-status groups tend to endorse essentialist beliefs to a greater degree than members of low-status groups (Kraus and Keltner 2013). The inverse relationship we found between gender essentialism and indicators of cognitive flexibility demonstrates a more epistemic basis to these beliefs, namely that an essentialistic view of gender also helps provide order, structure, and definitive answers in the face of potential ambiguity and uncertainty. This preference for stability and uniformity is likewise reflected in the tendency of essentialistic thinkers to adopt dispositional and entity theories about the stability of human behavior.

Given the somewhat exploratory nature of the Social Determinism subscale, our findings should be interpreted tentatively. For some variables, social determinism yielded results contrary to the GEM composite, but in most cases, social determinism was unrelated to the hypothesized characteristics, functions, and consequences of essentialist thinking. For example, regarding sexism, endorsing a social-deterministic view of gender was *negatively* related to hostile and modern sexism—constructs that represent denial of sexism and resentment toward women and feminists. However, social determinism was *unrelated* to old-fashioned and benevolent sexism—constructs that embody reverence for traditional gender roles. Similarly, social determinism was negatively related to those system-justifying beliefs that represent an overtly aggressive stance on authoritarianism (e.g., right-wing authoritarianism), again showing the opposite pattern as the GEM composite. In contrast, social

determinism was *unrelated* to less inflammatory beliefs that focus on global attitudes toward social inequality (e.g., social dominance orientation). Perhaps most intriguingly, social determinism showed the *same positive* association as the GEM composite with concepts like belief in the Protestant work ethic—a conceptual stepping stone for rationalizing social hierarchies by implying that lack of effort is to blame for low social status—intolerance for ambiguity, and lay dispositionism, as well as the *same negative* relationship with empathic concern, suggesting that social determinism may reduce uncertainty and lessen sympathy for less fortunate others.

The differentiation in our findings for the GEM composite and Social Determinism subscale underscores the importance of distinguishing these two types of essentialism. Although Rangel and Keller (2011) proposed that a general belief in social determinism represents an independent, yet influential, component of essentialist thinking, our results indicate that the predictors and consequences of these two types of essentialism warrant further exploration. In particular, our preliminary findings support viewing these two types of essentialism as conceptually orthogonal, rather than opposing, as much prior research and theory has assumed.

Study 3

In Study 3, we used the GEM to examine whether more essentialistic views of gender predict tolerance of existing disparities between men and women (e.g., employment, income, status). Understanding the system-justifying function of essentialistic beliefs might help explain how such disparities continue to exist despite policies aimed at reducing them (Brescoll et al. 2013; Yzerbyt et al. 1997). We predicted that an essentialistic view of gender would be related to increased acceptance of inequality and reduced motivation to address these disparities.

Method

Participants

We set a sample-size target of 160 in order to achieve power of .90 (two-tailed alpha = .05) to detect a correlation of .25 (GPower). A total of 170 United States members of MTurk completed a two-part survey on gender disparities in America in exchange for payment ($.30 for completing both parts). After cleaning (see details in the following), the sample consisted of 133 individuals (68.4% female; M_{age} = 38.5 years, SD = 12.9, range = 19–71).

Procedure

Prior to data collection, all procedures involved in our research were reviewed and approved by the University of Rochester Research Subjects Review Board. In Part 1, participants completed the GEM, the Ambivalent Sexism Inventory (Glick and Fiske 2001), and several demographic questions, for which they were paid $.20. All measures were administered and tallied in the same manner as described in Study 2. In order to minimize response set effects, we asked about gender disparities in a separate measure, administered approximately one week later. Participants were recontacted to complete a brief follow-up study for an additional $.10. These instructions stated: "There are disparities between men and women in this country in a number of areas (e.g., inequalities in employment, income, social status). On many different dimensions, women do not do as well as men." Participants then used a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*) to indicate their endorsement of 10 statements (plus one filler item) about gender inequities (modeled after Williams and Eberhardt's 2008, study of tolerance for racial inequities). A sample item is: "Gender disparities pose a major problem to American society" (reverse-scored). (The full list is presented in the online supplement, Section 3.) A total score for comfort with gender disparities was calculated by averaging responses across the 10 items (α = .92).

Data Cleaning

In both parts of the study, we included an attention check adapted from the Directed Questions Scale (DQS; Maniaci and Rogge 2014). This item stated: "To show that you are paying attention, please skip this question." Thirty-seven participants (21.8%) failed one or both attention checks and were excluded from analyses.

Results

Comfort with gender disparities was regressed on participants' gender, the GEM composite, hostile and benevolent sexism (included as covariates to establish unique effects of essentialism over and above sexism), and all two- and three-way interactions. None of the interactions was significant, so these effects were dropped from the model. Together, the four main effects accounted for 72.5% of the variance in comfort with gender disparities (R^2 = .73), $F(4, 128) = 84.97$, $p < .001$. Controlling for the other main effects, all three scales emerged as significant predictors of comfort with gender disparities: the GEM composite (β = .52), $F(1, 128) = 59.55$, $p < .001$; hostile sexism (β = .51), $F(1, 128) = 68.73$, $p < .001$; and benevolent sexism (β = −.20), $F(1, 128) = 11.88$, $p < .001$. Participant gender was not significant in this analysis (β = .04), $F(1,$

128) = .85, p = .640. An additional exploratory analysis was performed substituting the Social Determinism subscale for the GEM, but it did not significantly predict comfort with gender disparities (β = −.06), F(1, 128) = 1.06, p = .305.

Discussion

Study 3 found that a more essentialist view of gender was associated with greater comfort with and acceptance of gender disparities. Importantly, this relationship persists when controlling for measures of explicit sexism. Although believing that men and women have inherently separate and meaningful "essences" may lead to seeing gender differences as natural, further inferring that those differences are unproblematic, inevitable, and acceptable may represent a "naturalistic fallacy"—a tendency to believe that what is natural is also desirable. Thus, gender inequities are seen as less concerning and unlikely to change, a pattern of thinking that may reduce personal motivation to support policies that address gender disparities.

Study 4

Study 4 had two main purposes. First, because feeling concerned about inequality is an important precursor to moving for change, we examined emotional reactions to existing gender disparities. Second, Study 3 used a correlational design, obviating causal interpretation—it is possible, for example, that both gender essentialism and tolerance of gender disparities reflect the influence of other factors, such as sexism or social dominance orientation. The reverse pathway is also plausible—that greater acceptance of gender disparities may lead people to essentialize gender. Therefore, in Study 4, we experimentally manipulated views of gender to be more or less essentialistic and subsequently assessed emotional reactions to a real news article about gender inequality in America. We hypothesized that participants exposed to a more biologically essentialistic view of gender would feel less moved and upset by existing gender inequality than would those exposed to a non-essentialistic view of gender.

Method

Participants

Because the effect size in Study 3 was relatively large (β = .52), we set a medium effect size (d = .50) as a target, which indicated a sample size target of 64 per group (n = 128) in order to achieve power of .80 (two-tailed alpha = .05). University students (n = 118; 61.8% female; M_{age} = 19.1 years, SD = 1.3, range = 18–25) participated in exchange for extra credit in their Psychology courses. Participants completed the Modern and Old-Fashioned Sexism Scales (Swim et al. 1995) prior to coming to the lab. These measures were formatted identically to Study 2. By using a different sexism measure than in Study 3, we hoped to show that the effects of gender essentialism do not depend on which specific measures are controlled.

Materials and Procedure

Prior to data collection, all procedures involved in our research were reviewed and approved by the University of Rochester Research Subjects Review Board. Participants arrived for a study ostensibly concerned with awareness of media issues. To bolster the cover story, participants first answered questions about their media consumption and their beliefs about media trustworthiness, scientific research, and media reports about scientific research. Next, we manipulated essentialist beliefs by randomly assigning participants to read an article describing experimental evidence that either supported or undermined a biological basis for gender differences. This priming manipulation was disguised by embedding the headline for either an essentialist article ("Scientists Identify Difference Between Male and Female Brains") or a non-essentialist article ("Scientists Say That Gender Roles Have No Genetic Basis") within a list of headlines on scientific topics unrelated to gender (e.g., "Scientists Discover a New Test For Malaria, No Blood Required") and by telling participants they would be randomly assigned to read and answer comprehension questions for one article on the list.

After reading this article, participants were told they would be randomly assigned to evaluate two additional articles, this time about American social issues. They were then shown a second list of headlines that included a distracter article intended to reduce awareness of the study's focus on gender ("California Bill Calls for Cuts in Emissions"), the headline for the target article detailing existing gender inequalities ("Gender Inequality in the U.S. Today"), and several additional headlines on topics unrelated to gender (e.g., "Banks Say No to Marijuana Money, Legal or Not"). All participants then read and answered comprehension questions about the distracter article on environmental legislation and then the gender inequities article.

To assess emotional reactions to gender inequities, we gave participants an article that appeared on the Trust Women PAC website on March 19, 2012 (online supplement, Section 3). Headlined "Gender Inequality in the U.S. Today," the article argues that despite surface changes in earning power, gender inequality is still a very real issue in America, pointing out that although both genders are equally present in the workforce, men continue to dominate top management positions in many professions. The article was inconclusive about reasons

underlying this inequity, allowing participants to draw their own conclusions. After reading the article, participants indicated the degree to which each of 11 adjectives described their present mood using a 5-point scale (ranging from 1 = *not at all* to 5 = *extremely*), again adapted from Williams and Eberhardt (2008). Four of these items were summed in a composite reflecting emotional engagement (moved, concerned, upset, and nervous) and four were summed in a composite reflecting emotional disengagement (comfortable, indifferent, relaxed, and apathetic). Also, three filler items were included: educated, informed, and knowledgeable. Finally, participants provided demographic information and were debriefed and thanked.

We manipulated essentialist beliefs by having participants read one of two mock articles ostensibly taken from the newspaper, *The Los Angeles Times*. The essentialist article (see online supplement, Section 3) described research that had discovered a hard-wired difference between the male and female brain that explains "some of the stereotypical disparities in male and female behavior." To maximize believability, this article was adapted from an actual media report of research conducted at the University of Pennsylvania Perelman School of Medicine (Penn Medicine News 2013). The non-essentialist article advanced a less biological basis for gender differences (online supplement, Section 3), in which researchers were described as failing to discover any substantial structural differences between men's and women's brains that could explain stereotypical differences between men's and women's behaviors. This article was also based on an actual article describing research conducted at Washington University in St. Louis (Gorman 2014), adding arguments for a non-biological conception of gender paraphrased from Fine (2010). In this version, we created the following quote, ascribing it to the lead author:

> [o]ur results are lending weight to the growing evidence that behavioral differences between men and women are shaped by the social and cultural environments in which they grow up—that they can't be attributed to some sort of biological or evolutionary explanation.

Results

As in Study 3, we regressed emotional engagement on modern sexism, old-fashioned sexism, a dummy-coded variable representing participants' gender, and a dummy-coded variable representing experimental condition. The main effects of these four variables accounted for 23.6% of the variance in emotional engagement ($R^2 = .24$), $F(4, 113) = 8.75$, $p < .001$. The effect of experimental condition was significant ($\beta = -.23$), $F(1, 113) = 7.30$, $p = .008$, controlling for participants' gender, modern sexism, and old-fashioned sexism. Participants in the gender essentialist condition were significantly less emotionally engaged ($M = 2.16$, $SD = .60$) than participants in the non-essentialist condition ($M = 2.49$, $SD = .70$), ($\beta = -.23$), $F(1, 113) = 7.30$, $p = .008$. The gender difference was not significant ($\beta = -.17$), $F(1, 113) = 2.98$, $p = .087$. Not surprisingly, greater emotional engagement was linked to greater old-fashioned sexism ($\beta = .26$), $F(1, 113) = 8.19$, $p = .005$, and less modern sexism ($\beta = -.34$), $F(1, 113) = 11.11$, $p = .001$. None of the two-way or three-way interactions was significant.

Similar analyses accounted for 10.6% of the variance in emotional disengagement ($R^2 = .11$), $F(4, 113) = 3.35$, $p = .012$. Participants in the gender essentialist condition ($M = 3.18$, $SD = .64$) were similarly emotionally disengaged as participants in the non-essentialist condition ($M = 2.90$, $SD = .71$, although the difference was not significant, $\beta = .17$), $F(1, 113) = 3.62$, $p = .059$. Men ($M = 3.29$, $SD = .75$) were more emotionally disengaged than women were ($M = 2.90$, $SD = .61$; $\beta = .21$), $F(1, 113) = 4.24$, $p = .04$. Old-fashioned and modern sexism did not account for unique variance in emotional disengagement, ($\beta = -.03$), $F(1, 113) = .08$, $p = .79$, and ($\beta = .07$), $F(1, 113) = .43$, $p = .51$, respectively. None of the two-way or three-way interactions was significant.

Discussion

Taken together, participants in the gender essentialist condition were less emotionally engaged than participants in the non-essentialist condition. These results establish that gender essentialistic beliefs can be considered causal of emotional reactions to actual gender disparities that currently exist in the United States, thereby ruling out the reverse causal pathway and influence by third variables. Participants led to think of gender in essentialist terms were significantly less moved and upset, and more emotionally disengaged, than those led to view gender in non-essentialist terms. Responding to social inequality with these emotions can serve as a catalyst for igniting change and motivated action. Thus, seeing gender differences as essentialistic and natural may lead not only to less emotional distress, but may also lessen motivation to reduce the gender gap.

General Discussion

The present research was designed to advance our understanding of gender essentialism. The main contributions of the present research over prior measures fall into two general categories: psychometric improvement and substantive content. Regarding psychometrics, in Study 1, we used Item Response Theory to develop a psychometrically optimized measure that

will be amenable to diverse research contexts and that may be broadly useful for improving the power and precision of personality and social psychology measures. Psychometrically, the items selected for the GEM have maximum variance, which better affords power for identifying substantive associations with other constructs. Moreover, the measure maximizes the scale's ability to distinguish among individuals who score in the low or high range of gender essentialism, which should improve the sensitivity of future studies.

As for the second contribution, the GEM is the first measure of gender essentialism that directly considers the complex and multifaceted nature of psychological essentialism. Our findings strongly support previous calls to avoid reducing essentialism to biological determinism (Haslam et al. 2000) by including three new facets: immutability, inductive potential, and social determinism. We very much hope that the GEM will allow future researchers to explore the separate contributions of each of these factors, which should help the field develop a more nuanced understanding of how gender essentialism affects behavior. For example, we would speculate that the immutability factor would be particularly relevant to prejudice against transgender individuals. More pointedly, our work strongly supports prior work that has suggested rethinking the treatment of biological and social determinism as bipolar opposites (Rangel and Keller 2011), as has been common in the literature (and as we discuss in the following).

Study 2 used the newly developed GEM to create a nomological network describing the functions and correlates of endorsing an essentialistic view of gender. Studies 3 and 4 used both correlational and experimental methods to demonstrate that gender essentialism leads to greater acceptance of existing disparities between women and men. Our three studies demonstrate the usefulness of the multi-faceted GEM for advancing our understanding of gender essentialism. In Study 2, we used a large national U.S. sample to describe how endorsing essentialist views of gender relates to a person's larger theory of mind. These various findings show how essentialist beliefs about gender, over and above the effects of sexism, are intimately linked to relatively inflexible, deterministic, and dispositional accounts of human behavior. Moreover, epistemically, essentialistic views of gender provide important cognitive underpinnings to system-justifying ideologies, various forms of sexism, and beliefs in the legitimacy of existing gender inequalities in the current social structure. Taken in conjunction with prior findings that gender is the most essentialized of all human categories (Gelman et al. 1986; Gelman and Taylor 2000; Prentice and Miller 2006; Taylor 1996), our findings suggest that gender essentialism pervasively influences human activity. Although essentialist beliefs in general have been linked to prejudice, discrimination, system-justification, and cognitive inflexibility, our research is the first known to point more specifically to a link between essentialist beliefs about gender and many of these constructs.

Although social determinism is sometimes construed as the opposite of biological determinism, our results suggest that for certain constructs (e.g., dispositional thinking, certain system-justifying and epistemic beliefs, and empathic concern), they may function similarly. These findings indicate that social determinism is better considered as conceptually distinct, yet still representative of essentialistic thinking about gender. It would be preferable to think of the opposite of essentialist thinking as seeing individuals as individuals—that is, as persons whose attributes vary according to their individual biology, development, and environments rather than categories to which they happen to belong (Carothers and Reis 2013). More fully characterizing the function and consequences of holding a socially deterministic view of gender represents an exciting avenue for future research.

Another noteworthy contribution of the current work is in advancing our understanding of the association between essentialist thinking about gender and interpersonal constructs such as empathic concern and perspective-taking. These latter two constructs are foundational to many close relationship processes, such as intimacy, responsiveness, and constructive communication (Reis and Clark 2013). These results therefore suggest one plausible mechanism that may undermine successful romantic relationships for heterosexual persons high in gender essentialism. More generally, because much of social life is stratified on the basis of gender, just how individuals conceptualize gender may influence their thinking and behavior in relationships. For example, attributions for a relationship partner's behavior during conflicts may differ as a function of whether the behavior is attributed to "something about one's gender" as opposed to "something about the individual." Future research might profitably focus on how gender essentialism relates to actual and self-reported behavior in romantic heterosexual relationships.

Studies 3 and 4 shed light on the role of ideology in justifying existing inequalities between women and men, with implications for tolerance of existing inequalities and for motivation to enact social change. By allowing people to see gender differences as natural, inevitable, and inherent, essentialist beliefs help legitimize different and unequal treatment of men and women. Study 4 focused on a contemporary debate that discriminates against a lower-status group because this sort of situation emphasizes the system-justification function of essentialistic thinking. But essentialistic thinking may also extend to other circumstances. For example, some gender inequalities may give advantages to women, such as in child custody cases where the legal system often favors mothers over fathers. It will be important in future research to determine whether gender essentialists are similarly less distressed by such instances.

The present findings may provide insights that can help make sense of ongoing debates about sexism and discrimination, and in particular the resistance of gender inequalities to

change. To the extent that these inequalities are supported by coherent belief systems, incorporating such facets as lay dispositionism, cognitive style, and perspective-taking, they should be resistant to modification. Nevertheless, our findings suggest novel avenues for intervention. For example, research on mindsets has shown that subtle interventions can move people from being entity theorists to a more growth-oriented mindset, which in turn affects outcomes such as academic performance, persistence, and reduced aggressive reactions and stress responses to peer exclusion and victimization (see Yeager and Dweck 2012, for a review). Similar ideas might be used to devise interventions that nudge people away from essentialist thinking about gender to more individual-based beliefs.

Limitations and Future Research Directions

Several limitations of our research merit attention. First, all of the current research was conducted using Western samples, however, it seems reasonable to expect that cultures which have more or less essentialistic views about gender might provide different results. For example, it will be valuable to explore the operation of essentialism in cultures where essentialistic beliefs about gender are more formally integrated into societal and government structures, as well as into key social hierarchies. A second limitation involves the manipulation used in Study 4, which pertained specifically to biologically based brain differences. Nevertheless, because our theorizing and findings indicate that essentialism may take forms other than biological determinism, future research is also needed to determine whether other bases of gender essentialism, such as social determinism, would foster increased tolerance for inequality. A third limitation is that all of the present studies relied on self-reports. It will be valuable to extend this research to behavioral outcomes, such as in situations where actual discrimination takes place (e.g., in job hiring) or in resolving actual conflicts in romantic relationships. Finally, it will also be useful in future studies to determine whether social desirability biases might influence correlations obtained between the GEM and other measures.

It would also be fruitful to compare and contrast the influence of gender essentialism and other forms of essentialism on a variety of attitudinal and behavioral outcomes. That is, although studies have shown that gender is the most essentialized dimension (Prentice and Miller 2006), essentialized thinking with regard to other social categories also has important consequences, as shown for example, in research on essentialistic thinking about race (Williams and Eberhardt 2008), nationality (Rad and Ginges 2018), and personality traits (Haslam et al. 2004). It is plausible that these varied types of essentialism share common elements, as suggested by Dar-Nimrod and Heine (2011), but that each social category also has unique causes and consequences. (In psychometric terms, this would mean that a higher-order factor exists that reflects shared variance among various lower-order factors.) In future research it would be informative to identify these shared and unique components of essentialistic thinking across diverse social categories.

Practice Implications

Our research suggests several useful implications for practice and policy. First, as a tool, the GEM might be used to assess the extent and prevalence of gender-essentialistic beliefs in various settings. The GEM incorporates a broader set of underlying conceptual criteria than prior measures do, providing a multifaceted perspective for professionals wishing to examine the influence of gender-based thinking in various settings—for example, in schools, organizations, and policy contexts. To be more specific, because the GEM goes beyond the most obvious facet of essentialistic thinking—biological determinism—to encompass inductive potential, immutability, and social determinism, it is capable of identifying effects of these somewhat more subtle yet nonetheless important forms of essentialistic thinking.

A further practical limitation is highlighted by considering the fact that all of our demonstrated findings were obtained over and above the effects of sexism. That is, all of our analyses controlled for one or another existing measures of sexism, and they still found noteworthy implications of essentialistic thinking about gender. Given the goal of ameliorating the effects of gender stereotyping in contemporary institutions such as government, industry, and education, it will be important to do so in a way that not only continues current efforts to correct instances of sexism but also goes one step further by encouraging non-essentialistic thinking about gender—in other words, that fully treats individuals as individuals rather than as members of a gendered class (even if one's attitudes toward that class are benign).

Conclusion

Despite many years and much well-meaning attention, gender inequalities have remained stubbornly resistant to change. One fundamental underpinning for this resistance may be the continued popularity of belief systems that portray gender differences as large, pervasive, and inherent. Literally thousands of articles, in both the scientific and lay literatures, can be easily found which describe how men and women differ. Whatever the case-by-case merit of these articles may be, as a set they perpetuate the belief that there is something fundamentally and profoundly different about men and women and that these differences pervade most human activities. In turn,

these beliefs make it easier to justify the perpetuation of gender-based inequalities. Better understanding of the nature and implications of essentialistic thinking about gender will contribute to societal efforts to ameliorate the unequal treatment of men and women.

Compliance with Ethical Standards

Ethical Standards Statement All procedures performed in these studies were in accordance with the ethical standards of the institutional and/or national research committee (Research Subjects Review Board #00057154) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Altemeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press https://www.hup.harvard.edu/catalog.php?isbn=9780674053052.

Altemeyer, B. (2002). Dogmatic behavior among students: Testing a new measure of dogmatism. *The Journal of Social Psychology, 142*, 713–721. https://doi.org/10.1080/00224540209603931.

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*, 228–235. https://doi.org/10.1016/j.jesp.2005.03.003.

Bastian, B., Loughnan, S., & Koval, P. (2011). Essentialist beliefs predict automatic motor responses to social categories. *Group Processes and Intergroup Relations, 14*, 559–567. https://doi.org/10.1177/1368430210385258.

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin, 23*, 751–758. https://doi.org/10.1177/0146167297237008.

Brescoll, V. L., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science, 15*, 515–520. https://doi.org/10.1111/j.0956-7976.2004.00712.x.

Brescoll, V. L., Uhlmann, E. L., & Newman, G. E. (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology, 105*, 891–908. https://doi.org/10.1037/a0034701.

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*, 306–307. https://doi.org/10.1207/s15327752jpa4803_13.

Carothers, B. J., & Reis, H. T. (2013). Men and women are from earth: Examining the latent structure of gender. *Journal of Personality and Social Psychology, 104*, 385–407. https://doi.org/10.1037/a0030437.

Coleman, J. M., & Hong, Y. Y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity, 7*, 34–53. https://doi.org/10.1080/15298860600980185.

Crowson, H. M. (2009). Does the DOG scale measure dogmatism? Another look at construct validity. *The Journal of Social Psychology, 149*, 265–283. https://doi.org/10.3200/SOCP.149.3.365-383.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science, 314*, 435. https://doi.org/10.1126/science.1131100.

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137*, 800–818. https://doi.org/10.1037/a0021860.

Dar-Nimrod, I., Heine, S. J., Cheung, B. Y., & Schaller, M. (2011). Do scientific theories affect men's evaluations of sex crimes? *Aggressive Behavior, 37*, 440–449. https://doi.org/10.1002/ab.20401.

Davis, M. A. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85 https://www.uv.es/~friasnav/Davis_1980.pdf.

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry, 6*, 267–285. https://doi.org/10.1207/s15327965pli0604_1.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299. https://doi.org/10.1037/1082-989X.4.3.272.

Fine, C. (2010). *Delusions of gender: How our minds, society, and neurosexism create difference*. New York: W. W. Norton and Company https://wwnorton.com/books/Delusions-of-Gender.

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the couples satisfaction index. *Journal of Family Psychology, 21*(4), 572–583. https://doi.org/10.1037/0893-3200.21.4.572.

Gelman, S. A., & Hirschfeld, L. A. (1999). How biological is essentialism? In D. Medin & S. Atran (Eds.), *Folk biology* (pp. 403–446). Cambridge, MA: MIT Press.

Gelman, S. A., & Taylor, M. G. (2000). Gender essentialism in cognitive development In P. H. Miller & E. K. Scholnick (Eds.), Toward a feminist developmental psychology (pp. 169-190). New York: Routledge. https://books.google.com/books?hl=en&lr=&id=AKptqkv1XYoC&oi=fnd&pg=PA169&ots=9U87PZsmZL&sig=ged9DLnNyxpB4nzOEbsBxw4yQus–v=onepage&q&f=false

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57*, 396–404. https://doi.org/10.2307/1130595.

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*, 109–118. https://doi.org/10.1037/0003-066X.56.2.109.

Gorman, J. (2014, January 4). The brain, in exquisite detail. Retrieved from https://www.nytimes.com/2014/01/07/science/the-brain-in-exquisite-detail.html .

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage https://us.sagepub.com/enus/nam/fundamentals-of-item-response-theory/book3067.

Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social and Clinical Psychology, 21*, 628–644. https://doi.org/10.1521/jscp.21.6.628.22793.

Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin, 32*, 471–485. https://doi.org/10.1177/0146167205276516.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127. https://doi.org/10.1348/014466600164363.

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*, 87–100. https://doi.org/10.1348/014466602165072.

Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin, 30*, 1661–1673. https://doi.org/10.1177/0146167204271182.

Horvath, M., & Ryan, A. M. (2003). Antecedents and potential moderators of the relationship between attitudes and hiring discrimination

on the basis of sexual orientation. *Sex Roles, 48*, 115–130. https://doi.org/10.1023/A:1022499121222.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*, 172–194. https://doi.org/10.1037/0022-3514.70.1.172.

Jayaratne, T. E., Ybarra, O., Sheldon, J. P., Brown, T. N., Feldbaum, M., Pfeffer, C. A., … Petty, E. M. (2006). White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward blacks, and gay men and lesbians. *Group Processes and Intergroup Relations, 9*, 77–94. https://doi.org/10.1177/1368430206059863.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55*, 893–905. https://doi.org/10.1037/0022-3514.55.6.893.

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*, 686–702. https://doi.org/10.1037/0022-3514.88.4.686.

Klonis, S. C., Plant, E. A., & Devine, P. G. (2005). Internal and external motivation to respond without sexism. *Personality and Social Psychology Bulletin, 31*, 1237–1249. https://doi.org/10.1177/0146167205275304.

Kosterman, R., & Feshbach, S. (1989). Toward a measure of patriotic and nationalistic attitudes. *Political Psychology, 10*, 257–274. https://doi.org/10.2307/3791647.

Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology, 105*, 247–262. https://doi.org/10.1037/a0032895.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—Therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science, 24*, 622–633. https://doi.org/10.1177/0956797612457686.

MacDonald, A. P. (1970). Revised scale for ambiguity tolerance: Reliability and validity. *Psychological Reports, 26*, 791–798. https://doi.org/10.2466/pr0.1970.26.3.791.

Mahalingam, R., & Rodriguez, J. (2003). Essentialism, power and cultural psychology of gender. *Journal of Cognition and Culture, 3*, 157–174. https://doi.org/10.1163/156853703322148525.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61–83. https://doi.org/10.1016/j.jrp.2013.09.008.

Martin, C. L., & Parker, S. (1995). Folk theories about sex and race differences. *Personality and Social Psychology Bulletin, 21*, 45–57. https://doi.org/10.1177/0146167295211006.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosinadou & A. Ortony (Eds.), *Similarity and analogical reasoning (pp. 179–195)*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511529863.009.

Morton, T. A., Postmes, T., Haslam, S. A., & Hornsey, M. J. (2009). Theorizing gender in the face of social change: Is there anything essential about essentialism? *Journal of Personality and Social Psychology, 96*, 653–664. https://doi.org/10.1037/a0012966.

Muthén, L. K., & Muthén, B. O. (2012). Mplus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén. https://www.statmodel.com/download/usersguide/Mplususer guide Ver_7_r3_web.pdf

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology, 65*, 113–131. https://doi.org/10.1037/0022-3514.65.1.113.

No, S., Hong, Y., Liao, H., Lee, K., Wood, D., & Chao, M. (2008). Lay theory of race affects and moderates Asian Americans' responses toward American culture. *Journal of Personality and Social Psychology, 95*, 991–1004. https://doi.org/10.1037/a0012978.

Penn Medicine News (2013, December 2). Brain connectivity study reveals striking differences between men and women. Retrieved from https://www.pennmedicine.org/news/news-releases/2013/december/brain-connectivity-study-revea.

Poon, C. S., & Koehler, D. J. (2006). Lay personality knowledge and dispositionist thinking: A knowledge-activation framework. *Journal of Experimental Social Psychology, 42*, 177–191. https://doi.org/10.1016/j.jesp.2005.04.001.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*, 741–763. https://doi.org/10.1037/0022-3514.67.4.741.

Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science, 17*, 129–135. https://doi.org/10.1111/j.1467-9280.2006.01675.x.

Rad, M. S., & Ginges, J. (2018). Folk theories of nationality and anti-immigrant attitudes. *Nature Human Behaviour, 2*(5), 343–347. https://doi.org/10.1038/s41562-018-0334-3.

Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as component of psychological essentialism. *Journal of Personality and Social Psychology, 100*, 1056–1078. https://doi.org/10.1037/a0022401.

Reis, H. T., & Clark, M. S. (2013). Responsiveness. In J. A. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships (pp. 400–423)*. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195398694.013.0018.

Reis, H. T., & Wright, S. (1982). Knowledge of sex-role stereotypes in children aged 3 to 5. *Sex Roles, 8*, 1049–1056. https://doi.org/10.1007/BF00290999.

Richardson, S. S. (2013). *Sex itself: The search for male and female in the human genome*. Chicago: University of Chicago Press https://www.press.uchicago.edu/ucp/books/book/chicago/S/bo16835663.html.

Roets, A., & Van Hiel, A. (2011). Item selection and validation of a brief, 15-item version of the need for closure scale. *Personality and Individual Differences, 50*, 90–94. https://doi.org/10.1016/j.paid.2010.09.004.

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). London: Sage https://www.researchgate.net/profile/Marjorie_Taylor/publication/232599266_Category_labels_and_social_reality_Do_we_view_social_categories_as_natural_kinds/links/59cc157eaca272bb050c65c1/and-social-reality-Do-we-view-social-categories-as.

Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). New York: Springer. https://doi.org/10.1016/j.paid.2010.09.004.

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the epistemic belief inventory (EBI). In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 261–276). Mahwah, NJ: Lawrence Erlbaum Associates https://psycnet.apa.org/record/2001-18187-012.

Shelton, J. N., & Richeson, J. A. (2005). Intergroup contact and pluralistic ignorance. *Journal of Personality and Social Psychology, 88*, 91–107. https://doi.org/10.1037/0022-3514.88.1.91.

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68*, 199–214. https://doi.org/10.1037/0022-3514.68.2.199.

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). New York: Pearson Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571. https://doi.org/10.2307/1131718.

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571. https://doi.org/10.2307/1131718.

Thissen, D., Chen, W. H., & Bock, D. (2002). *Multilog user's guide: Multiple, categorical item and test scoring using item response theory (version 7.0) [computer software]*. Lincolnwood, IL: Scientific Software International.

Thompson, M., Naccarato, M., Parker, K., & Moskowitz, G. (1993). *The development and validation of the personal need for structure (PNS) measure*. Unpublished manuscript.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*, 1049–1062. https://doi.org/10.1037/0022-3514.67.6.1049.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology, 94*, 1033–1047. https://doi.org/10.1037/0022-3514.94.6.1033.

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In M. P. Zanna & J. M. Olson (Eds.), *Advances in experimental social psychology* (Vol. 46, pp. 55–123). Cambridge: Academic Press. https://doi.org/10.1016/B978-0-12-394281-4.00002-7.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302–314. https://doi.org/10.1080/00461520.2012.722805.

Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 20-50). Oxford, England: Blackwell. https://perso.uclouvain.be/vincent.yzerbyt/Yzerbytetal.1997.pdf.

Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review, 5*, 141–155. https://doi.org/10.1207/S15327957PSPR0502_5.