Original Manuscript

Social Psychological and Personality Science
1-13

DOI: 10.1177/1948550620988425
journals.sagepub.com/home/spp
SAGE

Patterns of Implicit and Explicit Stereotypes III: *Long-Term Change in Gender Stereotypes*

Tessa E. S. Charlesworth[1] and Mahzarin R. Banaji[1]

Abstract

Gender stereotypes are widely shared "collective representations" that link gender groups (e.g., male/female) with roles or attributes (e.g., career/family, science/arts). Such collective stereotypes, especially *implicit* stereotypes, are assumed to be so deeply embedded in society that they are resistant to change. Yet over the past several decades, shifts in real-world gender roles suggest the possibility that gender stereotypes may also have changed alongside such shifts. The current project tests the patterns of recent gender stereotype change using a decade (2007–2018) of continuously collected data from 1.4 million implicit and explicit tests of gender stereotypes (male-science/female-arts, male-career/female-family). Time series analyses revealed that, over just 10 years, both implicit and explicit male-science/female-arts and male-career/female-family stereotypes have shifted toward neutrality, weakening by 13%–19%. Furthermore, these trends were observed across nearly all demographic groups and in all geographic regions of the United States and several other countries, indicating worldwide shifts in collective implicit and explicit gender stereotypes.

Keywords

implicit social cognition, gender stereotypes, stereotype change, time series analyses (ARIMA)

> "There is nothing so obdurate to education or criticism as the stereotype."
>
> —Walter Lippmann (1922, p. 99)

The modern usage of the term "stereotype" was coined in 1922 by Walter Lippmann to capture the notion that certain representations are pervasive across people and time. The term was chosen intentionally to reference the "stereotypes" of printing plates, which allowed text to be easily duplicated and spread *widely* and almost *without change*. Today, these two features—that stereotypes are widespread and generally stable—continue to shape much of popular thinking and scientific theories about this unique mental construct (Haines et al., 2016). Furthermore, the stability of stereotypes that concern universally defined groups, such as men and women, are expected to be even more unshakable than other stereotypes (e.g., race stereotypes; Fiske, 2017) and especially so when they are assessed with *implicit* measures that are less controllable (e.g., Bargh, 1999; Lai et al., 2016).

And yet, it is also clear that, when it comes to gender roles and expectations, much has changed over the past several decades. Worldwide, women have made remarkable gains in the workforce including science, technology, engineering, and mathematics (STEM) fields; in the United States and internationally, women leaders are revealing their capacity leadership (e.g., in response to the COVID-19 pandemic; Garikipati & Kambhampati, 2020) and scientific discoveries (e.g., through patents; Intellectual Property Office, 2019). Although stereotypes are certainly not easy to change, these many societal transformations in women's representation may prompt shifts in the content of gender stereotypes over extended time. Indeed, as we elaborate below, *social role theory* posits that stereotypes are an approximate mirroring of the world as it is presented, thus suggesting that changes in the roles that women and men play in the world will ultimately shift the content of stereotypes about women and men (Eagly & Wood, 2012; Koenig & Eagly, 2014). For instance, as more women enter STEM and the workforce, the world signals a shift in the association of male-career/female-family or the association of male-science/female-arts; these shifts should be visible in the strength of gender stereotypes, both implicit and explicit. This is a reasonable view, but long-term, fine-grained temporal evidence of actual change in implicit and explicit gender stereotypes has remained elusive.

[1] Harvard University, Cambridge, MA, USA

Corresponding Author:
Tessa E. S. Charlesworth, Department of Psychology, Harvard University, Cambridge, MA 02138, USA.
Email: tet371@g.harvard.edu

Here, we provide the first comprehensive test of long-term change in both implicit and explicit gender stereotypes, using continuous, monthly data from 2007 to 2018. We model the rate and direction of change in implicit and explicit gender stereotypes to illuminate whether such collective representations have shifted alongside the changing gender landscape or whether they have instead "remained obdurate" in the face of such transformations. Additionally, we compare patterns across multiple demographic subgroups to identify whether stereotype change (a) is unfolding in parallel across groups (indicating societal-level change) or (b) is limited to specific groups in society (indicating group-level change). This approach can help illuminate the most likely sources behind *why* gender stereotype change happens.

Empirical and Theoretical Evidence for Gender Stereotype Stability

For centuries, the social categories "male" and "female" have been framed as essential and stable (Ellemers, 2018). Perhaps as a consequence, the stereotyped roles and attributes associated with male/female often also reveal persistence across time. For instance, in the domain of career/family stereotypes (associated with agency/communion stereotypes), the qualities most associated with women have reflected domestic, communal qualities for the last century, with women seen as "charming" (in 1910), "placid" (in 1950) and "maternal" (in 1990; Garg et al., 2018). Similarly, data show that there has been little change between 1983 and 2010 in explicit ratings of the roles played by the typical man or woman (e.g., leader vs. caregiver), *traits* (e.g., agentic vs. communal), *occupations* (e.g., chemist vs. elementary teacher), and *physical appearance* (e.g., strong vs. dainty; Haines et al., 2016). Together, the empirical evidence suggests that, although some explicit stereotypes appear to have shifted in the far past (e.g., Garg et al., 2018; Nesbitt & Penn, 2000), contemporary data often reveal stable gender representations.

Theoretically, gender stereotypes may be stable for a number of reasons. Individual-level cognitive biases, such as confirmation bias or subtyping (Higgins & Bargh, 1987), can make it difficult to notice the changing gender roles in society and thereby limit the opportunity for individuals to challenge and change their stereotypes. Societal-level processes also play a role in stereotype maintenance. For instance, the pervasive communication of gender stereotypes in language (e.g., Bolukbasi et al., 2016; Charlesworth et al., 2021) creates the perception that stereotypes are collective representations consensually held by everyone in society (Durkheim, 1924; Moscovici, 2000). As a result of being so widespread and embedded in cultural products, gender stereotypes may be (falsely) perceived as true and therefore as unnecessary to change. Furthermore, even when real-world changes have occurred in the gender distributions of roles and occupations (e.g., in some subfields of science; Charlesworth & Banaji, 2019b), it may take additional time for these observed changes to trickle into reported stereotypes due to a "cultural lag" (Diekman et al., 2010; Koenig & Eagly, 2014). In sum, many empirical and theoretical perspectives would anticipate little to no change in implicit and explicit gender stereotypes over the past decade.

Theoretical and Empirical Evidence for Gender Stereotype Change

Nevertheless, even in just the past decade, women have made continued gains into traditionally male-dominated roles (although men's entrance into female-dominated roles has lagged behind; Croft et al., 2015). In social role theory, the authors proposed that, as such boundaries of gendered roles shift over time, so too will the associated gender stereotypes (Koenig & Eagly, 2014). Changes toward more equitable representation of women in science and the workforce should precipitate a weakening of the male-science/female-arts and male-career/female-family associations. Such predictions are supported by experimental evidence of explicitly expressed stereotypes (e.g., Koenig & Eagly, 2014, 2019) as well as correlational evidence from single points in time (e.g., Miller et al., 2015). But what of evidence *across time*, tracking change in societal-level (aggregated) gender stereotypes?

Studies of explicit stereotypes occasionally support the possibility of gender stereotype change. For instance, in the domain of science/arts, children today draw more female scientists than children from 5 decades ago, indicating a weakening of the men-science/women-arts stereotype (Miller et al., 2018). And in the domain of career/family (or agency/communion), stereotypes have also shifted, although in the direction of *stronger* stereotypes: Women today are explicitly stereotyped as *more* communal than in the past, with 54% of respondents in 1946 indicating that women were more communal, whereas 97% indicated this stereotype in 2018 (Eagly et al., 2019). Furthermore, it is worth noting that many observed changes are in the past, while contemporary data indicate relative stability (e.g., England et al., 2020). Finally, it is worth considering whether stereotype change extends to *implicit* gender stereotypes, measures of stereotypes that are less controllable and thus often argued to be less changeable, especially over the long term (Lai et al., 2013). The present work brings contemporary, continuous data of explicit and implicit stereotypes to test whether such representations have changed alongside shifts in society at large.

The Scope of Change: Widespread Across Society or Isolated to Some Groups?

To understand the patterns and sources of long-term gender stereotype change, we consider not only whether change has occurred over the past decade but also the scope of such change. Is change widespread across most demographics and geographic locations indicating societal-level change? Or is change isolated to some demographics indicating group-level change shaped by group-specific experiences and motivations? As we elaborate elsewhere (Charlesworth & Banaji, in press), the answer to these questions can guide an understanding of the most likely *source* of change. If we observe societal-level change, then the sources are widespread, such as worldwide

Table 1. Sample Demographics Across Gender-Career and Gender-Science Stereotypes: 2007–2018.

Sample Demographics	Gender-Career	Gender-Science	Total
N	886,254	500,385	1,386,639
M_{age} (years)	29.41	28.26	28.99
SD_{age} (years)	12.3	12.39	12.34
% Female	68.05	66.08	67.34
% White	74.33	75.98	74.92
% Black	6.47	5.13	5.99
% Asian	9.16	8.86	9.05
% $\geq$College	93.04	81.57	88.9
% Liberal	47.13	53.85	49.56
% Neutral	29.9	26.39	28.63
% Conservative	22.97	19.76	21.81
% Christian	54.44	41.39	49.72
% Jewish	2.79	2.71	2.76
% Other religion	7.81	10.98	8.95
% Nonreligious	34.97	44.93	38.57
% U.S. citizen	75.95	78.47	76.86

changes in the representation of women in the workforce and achievements of gender equality (Dorius & Firebaugh, 2010). If, however, we observe group-level change (change driven by only some groups), then the sources are group specific. For example, conservatives, men, or older respondents may change slower because they have higher system justification and motivations to maintain the status quo (Jost, 2015). To identify this scope (and potential source) of gender stereotype change, we therefore compare the rate and direction of change across seven demographic groupings (gender, age, race, politics, religion, education, and geography).

Method

Data Source

Data were retrieved from the Project Implicit website (https://implicit.harvard.edu) for the male-career/female-family or the male-science/female-art Implicit Association Test (IAT). Clean data are available at https://osf.io/psru2. Data were subset to include only completed sessions with IAT D scores and explicit stereotypes as well as complete demographics. Eighty percent of complete sessions were retained for the male-career/female-family IAT, and 77% of complete sessions were retained for the male-science/female-art IAT (Supplemental Table S1), resulting in a final sample of nearly 1.4 million completed sessions, the largest sample ever used to investigate gender stereotype change over time (see Table 1 for sample demographics).

For implicit stereotypes, data were included from January 2007 through December 2018. As we describe in Supplemental Materials (SM), methodological changes in the recording of explicit stereotypes occurred in June 2016 (specifically, the design switched from multiple questions to a single question per page) that resulted in an unexplained drop in explicit stereotypes. To ensure accurate inference from consistent measures, explicit stereotype data were analyzed January 2007–May 2016; analyses with all data after May 2016 are reported in Supplemental Table S3 and visualized in Supplemental Figures S1 and S2 in the SM.

Measures

IAT. Implicit gender stereotypes were measured using the IAT (Greenwald et al., 1998), with category labels of "male" and "female" and attributes of either "science"/"liberal arts" for male-science/female-arts, or "career"/"family" for male-career/female-family (see Supplemental Table S2 for stimuli).

Explicit measures. Explicit gender stereotypes were measured using 7-point Likert-type scales assessing the degree to which an attribute was female/male, from -3 (*strongly female*) to $+3$ (*strongly male*), with 0 representing a neutral association. Explicit stereotypes were separately assessed in two questions, one for each attribute (e.g., the association of *science* with female/male, and, separately, the association of *arts*). To better approximate the relative nature of the IAT, relative explicit stereotype scores were created by subtracting the "incongruent" association from the "congruent" association (e.g., [male vs. female-*science*] $-$ [male vs. female-*arts*]). Thus, -6 reflects a strong explicit counterstereotype association (e.g., male-arts/female-science), and $+6$ reflects a strong stereotypic association (e.g., female-arts/male-science). Additionally, to assess whether one facet of the association (e.g., male vs. female-*science* or male vs. female-*arts)* is driving the observed change, we report additional analyses for each individual explicit scale (see Supplemental Figure S3 and Supplemental Table S4 for individual, decomposed trends).

Demographic measures. Respondents indicated their age/birth year, gender, level of education, political ideology, ethnicity/race, and country of residence. U.S. respondents also provided their current state and county of residence (for a list, see Section 1 in SM).

Analytic Strategy

Examining stereotype change with autoregressive integrated moving average (ARIMA) time series. Long-term change in cross-sectional data (i.e., collected from different groups of people across time points) is well suited to analysis by ARIMA time series models (Cryer & Chan, 2008; Jebb et al., 2015). As elaborated elsewhere (Charlesworth & Banaji, 2019b; Varnum & Grossmann, 2017a), ARIMA models offer necessary advantages over previous multiple regression approaches, including the ability to (1) account for temporal autocorrelations (i.e., measures close in time are often highly dependent), (2) model nonlinear patterns and seasonality, and (3) offer forecasts of possible future patterns' change. The present project therefore uses ARIMA models to examine the overall and demographic patterns of change.

Examining demographic differences in stereotype change over time. To identify and interpret whether all groups have changed in (1) similar, parallel ways (indicating societal-level change) or in (2) dissimilar, nonparallel ways (indicating group-level change), we adopt a two-step analytic approach, as outlined in Charlesworth and Banaji (in press).

In Step 1, we examine demographic *gaps* (e.g., subtracting the time series of women's gender stereotypes from the time series of men's gender stereotypes) to provide a succinct summary of whether groups have changed with *parallel* or *nonparallel* trends. If the gap between the two subgroups has remained stable, that indicates the subgroups are changing at parallel rates and directions. If, however, the *gap* between two subgroups has increased (diverged) or decreased (converged) over time that indicates the two subgroups are changing at nonparallel rates or directions. Formally, if the ARIMA series does *not* include a differencing parameter (i.e., $d = 0$ in the parameters $[p, d, q]$), and the forecasts are hovering around neutral, then we would conclude that the demographic gap is stable, and the two subgroups are changing in parallel. Alternatively, if the ARIMA series for the demographic gap includes a differencing parameter (i.e., $d > 0$), and forecasts that are different from neutrality, then we would conclude change is nonparallel across groups. In Step 2, we examine the individual trajectories for each demographic subgroup (e.g., men only, women only). Similar to post hoc t tests informing interpretation of an omnibus ANOVA, examining individual trajectories can reveal *why* the demographic gaps tested in Step 1 are converging, diverging, or remaining stable (e.g., whether the demographic gap is diverging because men or women are changing faster).

Controlling for sample changes over time, demographic covariates, and repeat test-takers. The current analyses of long-term stereotype change may be confounded by three features of the available cross-sectional data: (1) changes in the sample composition over time (e.g., increasingly young or female), (2) demographic intercorrelations (e.g., demographic differences observed across one demographic, such as race, could arise because of a correlated demographic, such as political orientation), and (3) repeat visitors causing regression to the mean (because repeat test-takers may create a drift toward more neutral stereotypes).

To address the first two challenges (sample change and demographic covariation), we use a weighting approach: We split the sample across the demographic comparison of interest and across years (creating a demographic-by-year subsample) and then re-weight each demographic-by-year subsample to match an overall target weight (see also, Charlesworth & Banaji, in press). For example, to compute the trajectories for men versus women, we split the data by gender and by year and then re-weight each gender's yearly sample to match the full sample target demographics on age, race, politics, education, and religion (see Supplemental Figure S7 in SM). Ultimately, the demographic composition of the male sample in 2007 will match the male sample in 2008, 2009, and so on as well the female samples in 2007, 2008, and so on.

To address the third challenge (repeat visitors), we performed additional analyses isolated to only those respondents who indicated never having taken an IAT before. The trends were consistent regardless of whether we subset to those "novice" IAT-takers or not, indicating the conclusions were not driven by regression to the mean among repeat test-takers (see Supplemental Table S5 and Supplemental Figure S4).

Examining geographic differences in stereotype change over time. Because separating data by both time and geography greatly reduces the number of tests available at each time-by-geography observation, time series trajectories were computed at the yearly (rather than monthly) level. Additionally, we attempted to follow the practice of Nosek and colleagues (2009) wherever possible and use a cutoff of 100 observations per geographic region per year to ensure that the estimates were sufficiently precise. This cutoff was employed for United Nations (UN) regions and U.S. states. However, because there are many fewer international than U.S. respondents, using this cutoff at the level of individual countries resulted in a significant loss of data: Only 11–13 countries remained. As such, we adopted a more lenient cutoff for individual countries, including data from countries with at least 20 observations per year (similar to cutoffs used at the county level; Orchard & Price, 2017). Ultimately, this approach resulted in sample sizes of $N = 12$ United Nations regions, $N = \sim 30$–50 international countries, and $N = \sim 50$ U.S. states. These many individual trajectories cannot easily be summarized into binary comparison pairs, so we summarize geographic differences as the *percentage* of regions that weakened in stereotype strength over time (i.e., changed toward neutrality). All individual geographic trajectories are reported in Supplemental Tables S9–S12 (for UN regions), Supplemental Tables S13–S16 (for countries), and Supplemental Tables S17–S20 (for U.S. states).

Of note, the primary data were all drawn from the U.S.-hosted website and may therefore not representatively capture the stereotypes of non-U.S. countries, especially those of non-English-speaking countries (e.g., China, Sweden, Spain). To address this concern, we acquired data from the three largest data sets of the Project Implicit country-specific websites: China, Sweden, and Spain. Each of these websites hosts the same male-science/female-arts test but in the specific language of the country (i.e., in Mandarin, Swedish, and Spanish). We then assessed the possible deviation between U.S. website data and country-specific website data. Both the overall means and trends of change were consistent regardless of the data source (see Supplemental Table S21 and Supplemental Figures S11 and S12 in SM).

Results

Overview of Results

At the broadest level, we find that both male-career/female-family and male-science/female-arts stereotypes were strong and significant on both implicit and explicit measures and were

Table 2. Means and Correlations of Implicit (2007–2018) and Explicit (2007–2016) Gender Stereotypes.

		Implicit Stereotypes				Explicit Stereotypes				Implicit–Explicit Correlation	
Stereotype	Demographic Group	*M*	*SD*	95% CI	*d*	*M*	*SD*	95% CI	*d*	*r*	95% CI
Male-career/female-family	Overall	.37	.37	[.37, .37]	1.00	1.67	1.81	[1.67, 1.68]	0.92	.15	[.15, .16]
	Women	.40	.36	[.40, .41]	1.12	1.66	1.82	[1.65, 1.67]	0.91	.16	[.15, .16]
	Men	.31	.39	[.31, .31]	0.78	1.70	1.80	[1.69, 1.71]	0.95	.15	[.15, .16]
Male-science/female-arts	Overall	.34	.41	[.33, .34]	0.81	1.69	1.73	[1.69, 1.70]	0.98	.22	[.21, .22]
	Women	.31	.42	[.31, .32]	0.76	1.67	1.75	[1.66, 1.68]	0.96	.22	[.22, .23]
	Men	.38	.41	[.37, .38]	0.92	1.75	1.70	[1.73, 1.76]	1.03	.20	[.20, .21]

observed in nearly every demographic group and every geographic region. Despite these strong and widespread magnitudes, both male-career/female-family and male-science/female-arts stereotypes have *weakened toward neutrality* over the past decade, on both implicit and explicit measures. Implicit stereotypes weakened by 13% (male-career/female-family) and 17% (male-science/female-arts) between 2007 and 2018, and explicit stereotypes weakened by 19% (male-career/female-family) and 14% (male-science/female-arts) between 2007 and 2016. Of course, even with this weakening, the stereotypes remain far from eradicated: The ARIMA model forecasts indicate that implicit male-career/female-family stereotype could take at least *134 years* from 2018 to touch neutrality, and even the faster-changing implicit male-science/-female-arts stereotype will take between 37 and 74 years from 2018 to pass neutrality.

Nevertheless, such change toward neutrality in implicit and explicit gender stereotypes was found to be *parallel* across most demographic subgroups: For 26 of 28 (93%) demographic comparisons (e.g., liberal vs. conservative, old vs. young), subgroups were moving at similar rates in the same direction. Additionally, nearly all geographic regions moved toward neutrality over the past decade: Every single UN region moved toward neutrality for both implicit and explicit gender stereotypes, as did 82%–91% of countries and 92%–98% of U.S. states.

Overall magnitude and correlations. Implicit male-career/female-family and male-science/female-arts stereotypes were strong in the stereotypic direction (Table 2). Small gender differences emerged on both tests: For male-science/female-arts, women ($D_{women} = .31$) had weaker implicit stereotypes than men ($D_{men} = .38$), $t(349{,}298) = 49.71$, $p < .001$, $d = .15$; for male-career/female-family, however, women ($D_{women} = .40$) had stronger implicit stereotypes than men ($D_{men} = .31$), $t(518{,}662) = -112.71$, $p < .001$, $d = .26$.[1]

Explicit gender stereotypes also revealed strong associations in the stereotypic direction (Table 2). Small gender differences emerged, but the effect sizes were near zero: Women had weaker explicit stereotypes on both male-science/female-arts ($M_{women} = 1.67$, $M_{men} = 1.75$), $t(144{,}531) = 10.07$, $p < .001$, $d = .04$, and male-career/female-family ($M_{women} = 1.66$, $M_{men} = 1.70$), $t(219{,}585) = 6.56$, $p < .001$, $d = .02$. Finally, both tests indicated small but significant implicit–explicit correlations (Table 2).

Overall patterns of change. Both implicit male-career/female-family and male-science/female-arts stereotypes have weakened toward neutrality over the past 12 years (from 2007 to 2018), by approximately 13% and 17%, respectively (Figure 1 and Table 3). Male-science/female-arts stereotypes have moved faster and more consistently, as indicated by a drift parameter in the best-fitting ARIMA model. Male-career/female-family stereotypes, in contrast, have moved relatively more slowly and inconsistently, revealing no drift parameter and forecasts that include both movement toward and away from neutrality.

Explicit male-career/female-family and male-science/female-arts stereotypes have also weakened toward neutrality (from 2007 to 2016) by approximately 19% and 14%, respectively (Figure 1 and Table 3). Both explicit stereotypes include drift parameters in their ARIMA models, implying consistent change over time. Additionally, looking at the individual explicit stereotype scales (such as the association of male over female to *career* alone) indicates that all individual explicit stereotypes are also weakening: Male-career associations weakened by 23%, female-family associations by 14%, male-science associations by 16%, and female-arts associations by 10% (see SM for discussion).

Demographic differences. Of 28 demographic comparisons (e.g., comparisons of men–women, Black–White for both implicit and explicit gender stereotypes), 26 (93%) revealed *parallel* change (Figure 2 and Supplemental Figure S8 and Supplemental Tables S6–S8 in SM), indicating widespread consistent patterns of change. This conclusion holds for both implicit (13 of 14 comparisons indicated parallel change) and explicit stereotypes (13 of 14 comparisons). Although some demographic comparisons indicated differencing parameters in their ARIMA models, the demographic gaps hovered around neutrality.

The two exceptions were (1) a gender difference on implicit male-science/female-arts stereotypes and (2) a Black–White race difference on explicit male-career/female-family stereotypes. First, on implicit male-science/female-arts stereotypes, men were found to have changed at a slower rate (weakening by 10% or .04 IAT points) than women (weakening by 21%

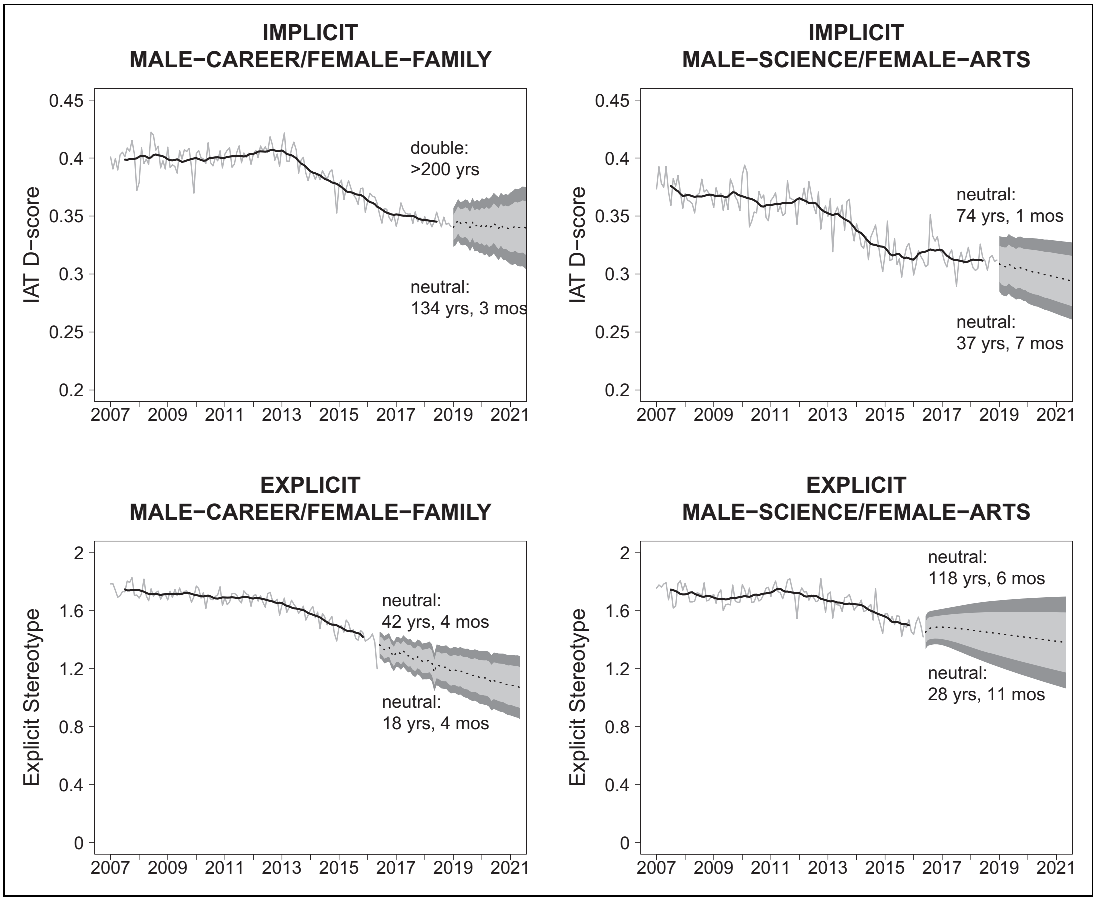

Figure 1. Implicit and explicit gender stereotype change from 2007 to 2021. *Note*. Thin lines indicate the observed monthly weighted means, thick lines indicate the decomposed trend lines of the observed monthly data (removing seasonality and noise), dark shaded areas indicate 80% confidence intervals (CIs), light shaded areas indicate 95% CIs of the autoregressive integrated moving average (ARIMA) model forecasts, dark lines inside shaded areas indicate the means of the ARIMA model forecast. Implicit stereotype data are from 2007 to 2018; explicit stereotype data are from 2007 to 2016.

Table 3. Patterns of Change Across Implicit (2007–2018), Explicit (2007–2016), and Implicit–Explicit Correlations (2007–2016) in Gender Stereotypes.

	Implicit Stereotypes				Explicit Stereotypes				Implicit–Explicit Correlation			
Stereotype	Start	End	%Δ	ARIMA	Start	End	%Δ	ARIMA	Start	End	%Δ	ARIMA
Male-career/female-family	.40	.34	−13.43	(2, 1, 1) (2, 0, 0)	1.78	1.20	−18.80	(0, 1, 1) (2, 0, 0) +drift	.16	.13	−2.59	(0, 0, 0) (0, 0, 2)
Male-science/female-arts	.37	.31	−17.12	(0, 1, 2) (1, 0, 0) +drift	1.76	1.42	−13.90	(1, 1, 3) +drift	.24	.18	−10.40	(0, 1, 1)

Note. For implicit stereotypes, starting values are from January 2007 and ending values are from December 2018; for explicit stereotypes and implicit–explicit correlations, starting values are from January 2007 and ending values are from May 2016. Starting and ending raw values are from the Implicit Association Test (*D* scores), from explicit stereotype scales, combined across two questions for each test (i.e., for gender-career: "how strongly do you associate career [or home] with males and females?" and for gender-science: "how strongly do you associate science [or arts] with males and females?"), and from correlations between these two measures. Percentage change is between the first and last values of the decomposed time series trend (removing seasonality and noise). Negative values indicate change toward neutrality (i.e., decreasing stereotypical bias); positive values indicate change away from neutrality (i.e., increasing stereotypical bias). For autoregressive integrated moving average (ARIMA) parameters, the first three parameters of the ARIMA model are nonseasonal, and the second three values are seasonal; drift is also included. In each set of parameters, *d* specifies the order of differencing parameters used to make the series stationary, *p* specifies the number of autoregressive parameters used to explain the autocorrelations in the data, and *q* specifies the number of moving average parameters used to explain the lagged forecast errors.

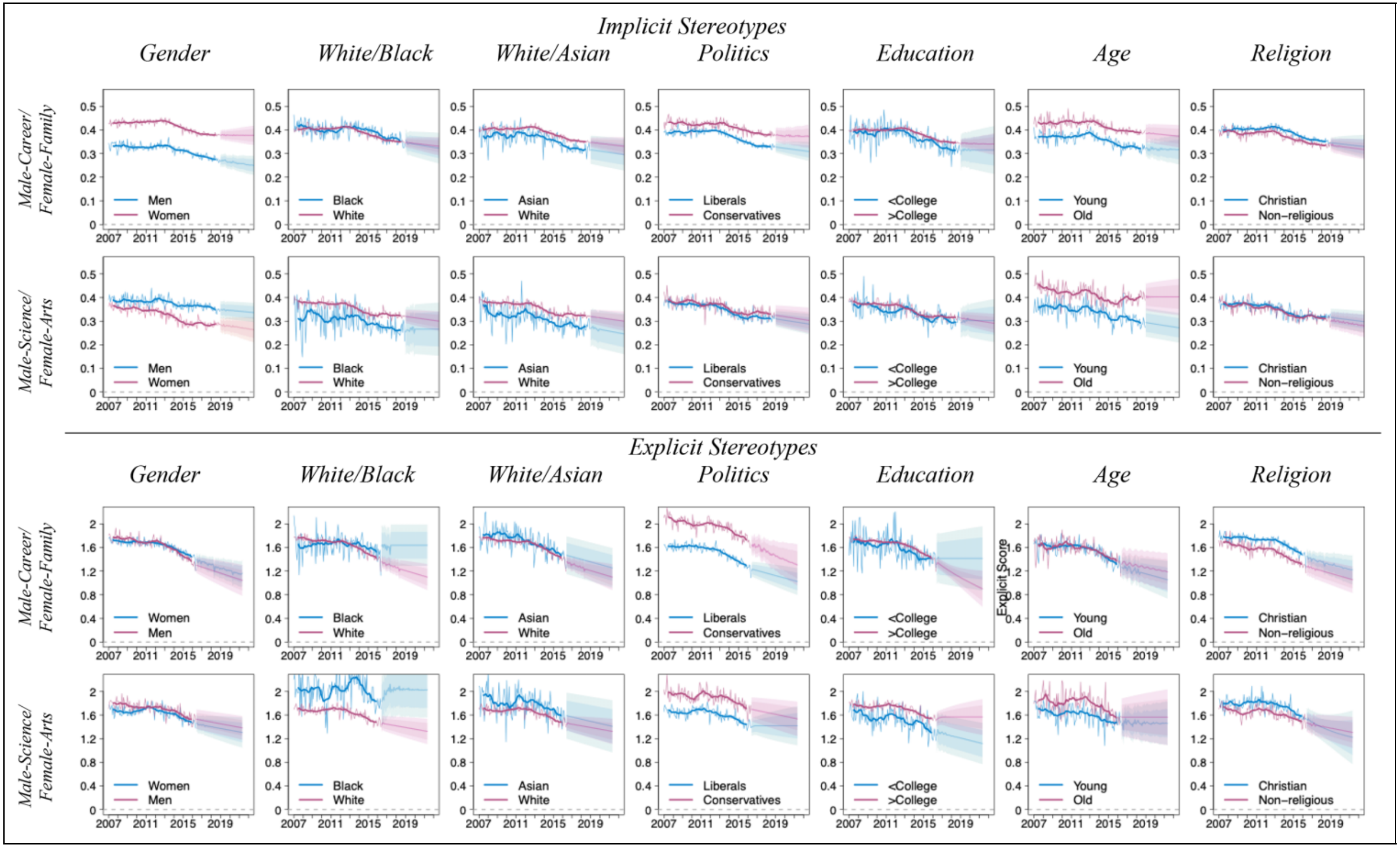

Figure 2. Demographic differences in implicit (2007–2018) and explicit (2007–2016) gender stereotype change. *Note. Y*-axis for implicit stereotypes indicates IAT *D* scores; *y*-axis for explicit stereotypes indicates combined explicit scales. Thin lines indicate the observed monthly weighted means of the two demographic subgroups as noted in the plot legend, thick lines indicate decomposed trend lines, dark and light shaded areas indicate 80% and 95% confidence intervals (CIs), respectively, of the autoregressive integrated moving average (ARIMA) model forecasts, and dark lines inside shaded areas indicate means of the ARIMA model forecasts.

Figure 3. Implicit (2007–2018) and explicit (2007–2016) gender stereotype change: Geographic differences across United Nations (UN) country regions. *Note.* Yearly averages by UN country regions, with each year-by-UN region requiring at least 100 observations. The legend in the top left plot applies to all four plots. Although individual lines of data may not be decipherable due to overlaps, we note that that is, in a sense, a key result of this article: The magnitude of stereotypes and change is generally overlapping across places. Nevertheless, in Supplemental Materials, we report the individual data for each region to make the small differences across regions more clearly interpretable (Supplemental Tables S9–S12).

or .08 IAT points, see Supplemental Table S4). The ARIMA model thus indicated demographic *divergence* between the faster changing (and less biased) women and the slower changing (and more biased) men. There are numerous possible reasons for men's relatively slower change, from men being more oblivious or resistant to persuasive messaging on gender bias (Parker et al., 2018; Régner et al., 2019) or a "backlash effect" in response to broader changes toward equity (Rudman & Glick, 2001; Rudman et al., 2012). Second, with respect to the Black–White race difference, Black respondents changed less on explicit male-career/female-family stereotypes (weakening by 7% or .12 points) than White respondents (weakening by 20% or .35 points, Supplemental Table S5). Speculatively, this difference may emerge because, historically, gender divisions in labor have been less pronounced in the Black community as lower economic status created a necessity for both parents to work outside the home (Boustan & Collins, 2013). Future research is needed to test the hypotheses for both demographic differences.

Geographic differences. At the highest level of geographic aggregation (UN geographic region), yearly trajectories indicated that *all* UN regions show a weakening of both gender stereotypes and on both implicit and explicit measures (Figure 3). On implicit male-career/female-family stereotypes, the largest change (18% decrease) was in sub-Saharan Africa, while the smallest (4% decrease) was in Western Asia; on implicit male-science/female-arts stereotypes, the largest change (25% decrease) was also in sub-Saharan Africa, while the smallest (9% decrease) was in Latin America and the Caribbean (see Supplemental Tables S9–S12 for individual regions). That sub-Saharan Africa emerged as the fastest changing region on implicit gender stereotypes is notable and may reflect fast-paced changes in women's workforce participation: In 2017, the top five countries for representation of women in the workforce were all in sub-Saharan Africa (Fetterolf, 2017).

Next, change was also widespread across individual countries: Between 82% and 91% of countries (of $N = 50$ for male-career/female-family and $N = 33$ for male-science/female-arts) moved to neutrality (Table 4; see Supplemental Figures S9 and S10 and Supplemental Tables S13–S16 for a report of all individual countries). The fastest and slowest changing countries were idiosyncratic across test topics and measures. For instance, on implicit male-career/female-

Table 4. Geographic Differences Across Countries.

Stereotype		*N Countries*	*N Median by Country*	*Percentage of Countries Weakening*	*Mean %Δ*	*Range %Δ*	*Mean Raw Δ*	*Range Raw Δ*	*Mean Overall*	*Range Overall*
Implicit	Male-career/female-family	50	1456	82	−10.09	−44.43 to 98.94	−.05	−0.20 to 0.20	0.37	0.30 to 0.44
	Male-science/female-arts	33	1170	82	−17.38	−57.59 to 15.63	−.08	−0.25 to 0.06	0.35	0.26 to 0.44
Explicit	Male-career/female-family	50	770	90	−20.00	−59.20 to 14.47	−.40	−1.19 to 0.24	1.69	1.36 to 2.05
	Male-science/female-arts	33	697	91	−21.68	−56.20 to 10.51	−.39	−1.33 to 0.12	1.56	1.18 to 1.97

Table 5. Geographic Differences by U.S. State.

Stereotype		*N States*	*N Median by State*	*Percentage of States Weakening*	*Mean %Δ*	*Range %Δ*	*Mean Raw Δ*	*Range Raw Δ*	*Mean Overall*	*Range Overall*
Implicit	Gender-career	43	11,350	98	−12.15	−22.73 to 5.09	−.05	−.11 to .02	0.38	0.36 to 0.40
	Gender-science	38	7,886	97	−15.28	−35.15 to 5.15	−.06	−.15 to .02	0.34	0.31 to 0.39
Explicit	Gender-career	43	11,350	98	−17.69	−32.13 to 18.15	−.32	−.60 to .24	1.63	1.39 to 2.12
	Gender-science	38	5,102	92	−10.12	−23.17 to 16.76	−.18	−.44 to .25	1.67	1.54 to 1.80

family stereotypes, Nigeria and Kenya moved the fastest toward neutrality (weakened by 44% and 31%, respectively), whereas on implicit male-science/female-arts stereotypes, Malaysia and Afghanistan moved the fastest toward neutrality (weakened by 58% and 56%, respectively). With this report, research can undertake correlational analyses with country-level predictors of stereotype *magnitude* (Nosek et al., 2009) but also, newly, *change*.

Across U.S. States, between 92% and 98% of states moved toward neutrality over the past decade (Table 5 and Figure 4; see Supplemental Tables S17–S20 for individual States). Despite consistency in direction, U.S. states varied in *rate* of change: from Louisiana, which changed fastest for both implicit male-career/female-family (weakened by 23%) and implicit male-science/female-arts stereotypes (weakened by 35%) to West Virginia *increasing* by 5% on implicit male-career/female-arts and District of Columbia similarly increasing by 5% on male-science/female-arts stereotypes. Nevertheless, it is again notable that the vast majority of states moved in the same direction toward neutrality.

General Discussion

Using data from 1.4 million tests of implicit and explicit gender stereotypes collected continuously between 2007 and 2018, the current study shows that, over a relatively short span of 12 years, *implicit and explicit gender stereotypes have weakened by 13%–19%*. Although gender stereotypes are sometimes argued to be particularly resistant to change because they are widely shared (e.g., Fiske, 2017), this rate of change is similar to that observed in highly discussed and more culturally variable attitudes about race and skin tone (Charlesworth & Banaji, 2019b). That gender stereotypes have indeed changed aligns with perspectives, such as social role theory (Eagly & Wood, 2012; Koenig & Eagly, 2014), that suggest the slow but steady changes in women's (and, to a lesser extent, men's) roles over the past several decades will eventually manifest in changes in the content of stereotypes. Several additional implications of this finding of change are worth emphasizing: (1) gender stereotype change is possible even on implicit measures, (2) gender stereotype change is widespread across groups, and yet (3) change toward stereotype neutrality is far from complete.

Change Is Possible Even in Implicit Stereotypes

Past work has often suggested that *implicit* measures may be particularly rigid, especially over the long term (Forscher et al., 2019; Lai et al., 2016). Indeed, even in the few cases where implicit measures have revealed long-term change, they usually change slower than explicit measures (e.g., implicit race attitudes decreased by 17%, but explicit race attitudes decreased by 37%, over the same period; Charlesworth & Banaji, 2019a). Yet, for gender stereotypes, not only have implicit gender stereotypes changed, they have changed at similar rates to explicit measures.

That implicit gender stereotypes have changed over the long term is also relevant to recent evidence (albeit drawn almost exclusively from short-term, within-individual studies)

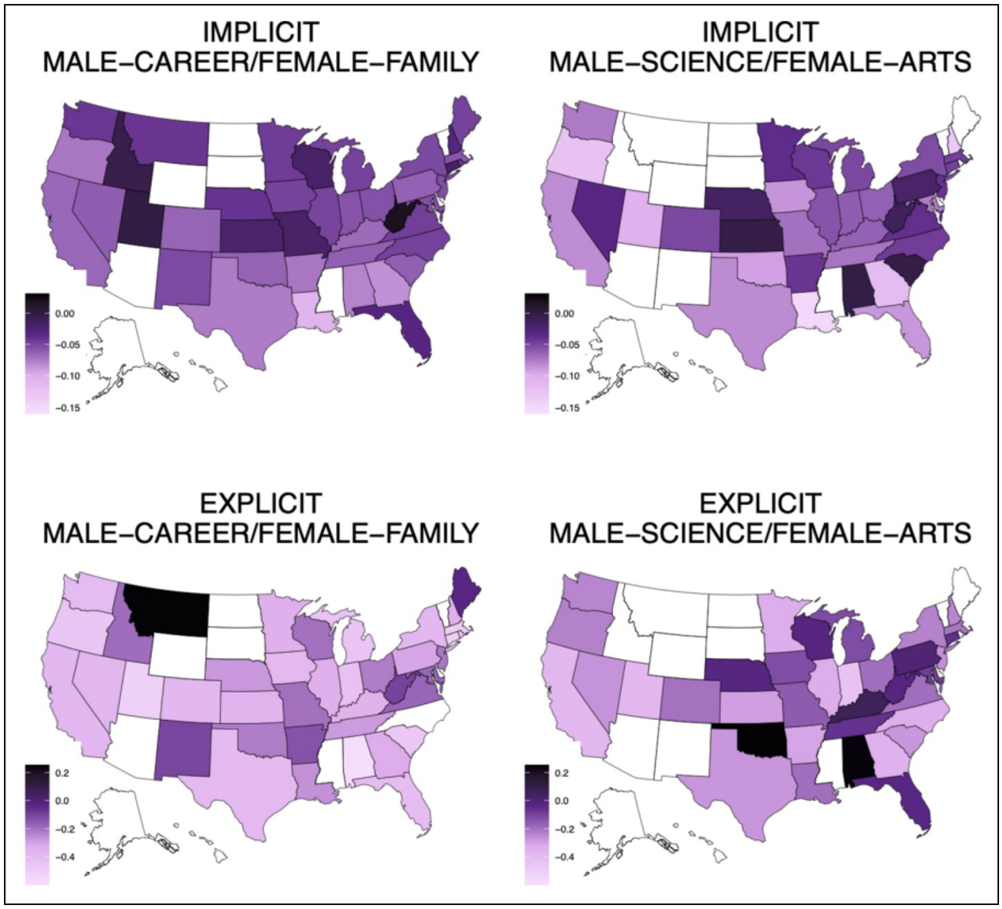

Figure 4. Implicit (2007–2018) and explicit (2007–2016) gender stereotype change: Geographic differences by U.S. states. *Note.* States are shaded according to their yearly raw change (e.g., for implicit stereotypes, [mean score in 2007 − mean score in 2018]). Darker shades indicate greater raw change away from neutrality; lighter shades indicate greater raw change toward neutrality; white indicates state did not have sufficient data by year (i.e., < 100 observations/year).

suggesting that changes in implicit attitudes and stereotypes may *not* precipitate changes in behaviors (Forscher et al., 2019). The current data could provide a new opportunity to test the link between implicit cognition and behavior *over the long term* and *at the societal level*: It is possible that, when aggregated across people and examined over extended time spans, changes in implicit gender stereotypes may precede changes in gender-based discriminatory behaviors (e.g., decreases in gender-based pay inequality or increases in reported feelings of belonging).

Implicit and Explicit Gender Stereotype Change Is Widespread

Not only do the data show that change is possible, we also find that gender stereotype change is remarkably widespread, with nearly every demographic group and geographic location moving in similar, parallel directions and rates. Such a result points to the broader, societal-level changes as the most likely lever of implicit and explicit gender stereotype change. Change has not simply been driven by some pockets of the population (e.g., young, liberals) with specific motivations for change; rather, change is permeating the whole society. Societal levers could include changes in women's representation in science and the workforce (as predicted by social role theory), social media movements like *#MeToo*, past decreases in pathogen prevalence (Varnum & Grossmann, 2017b), or loosening of social norms (Jackson et al., 2019). Ultimately, having now identified the most likely level of change, future researchers are poised to identify and test such hypotheses about *which* macro-level variables are preceding the widespread changes in implicit and explicit stereotypes.

Contemporary Change in Gender Stereotypes Is Possible but Not Complete

Although it is widely acknowledged that gender stereotypes have changed historically (e.g., Garg et al., 2018; Haines et al., 2016; Nesbitt & Penn, 2000), the current results also show that society has not reached a plateau: Contemporary change in gender stereotypes remains possible and worthy of continued focus and intervention. In fact, the current results also reinforce the importance of such continued focus. After all, society still has a very long way to go before neutrality is reached in gender stereotypes (as long as *134 years* for implicit male-career/female-family stereotypes), and many forecasts continue to include the possibility that change could even reverse away from neutrality if left unchecked.

Limitations: Sample Representativeness and International Data

Contributions notwithstanding, we recognize that the data used in this article are limited in that they are not a random sample and are therefore nonrepresentative. Previous work has considered this concern by showing that (a) patterns of change in Project Implicit data match trends observed with representative samples (Charlesworth & Banaji, 2019b; Ofosu et al., 2019), (b) overall magnitudes in Project Implicit data match biases in representative internet text (Caliskan et al., 2016), and (c) re-weighting U.S. Project Implicit data to U.S. census demographics does not substantially alter results (Charlesworth & Banaji, 2019b). Nevertheless, caution around nonrepresentativeness, especially in the reported geographic data, is warranted (Hoover & Dehghani, 2019). Although supplemental analyses revealed no differences in the conclusions between data collected in English from the U.S.-hosted website and data collected in Mandarin, Swedish, or Spanish from country-specific websites (see Supplemental Table S21 and Supplemental Figures S11 and S12), we encourage researchers interested primarily in specific international patterns to explore the data available from 41 country-specific websites (https://osf.io/kaqi5/).

Conclusion

Ever since the first use of the term "stereotype," gender stereotypes—associations of gender groups (men/women) and social attributes (e.g., science/arts, career/family)—have been largely seen as resistant to change. Yet over the past several decades, the real-world roles and representations of women (and, to a lesser extent, men) have been shifting. In this article, we bring contemporary data of 1.4 million tests collected continuously from 2007 to 2018 to show that societal implicit and explicit gender stereotypes are also changing over the long term, weakening in magnitude by 13%–19%. These changes are consistent across nearly all demographic groups and geographic regions, suggesting that gender stereotypes are collective representations shifting through widespread societal-level transformations. Far from the rigid stereotypes of a printing press, the contemporary notion of gender stereotypes may be better reflected as a dynamic representation that is responsive to the changes in society.

Authors' Note

All original data and analysis scripts reported in this article are available through OSF: https://osf.io/psru2/

Acknowledgment

We are grateful to Miao Qian for comments on this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Harvard Dean's Competitive Fund for Promising Scholarship awarded to Mahzarin R. Banaji.

ORCID iD

Tessa E. S. Charlesworth https://orcid.org/0000-0001-5048-3088

Supplemental Material

The supplemental material is available in the online version of the article.

Note

1. Further discussion of this gender difference is provided in the Supplemental Material (Figures S5 and S6). The gender difference is observed across nearly every country and year examined. In the four countries where men were more biased than women (i.e., Austria, Denmark, Finland, Sweden), men were found to have particularly strong *female-family* explicit associations. Thus, the gender difference in most other countries may be due to *men's* relatively weaker female-family associations (i.e., they also explicitly associated their own group with family).

References

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). Guilford Press.

Bolukbasi, T., Chang, K. W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 4356–4364. http://arxiv.org/abs/1607.06520

Boustan, L. P., & Collins, W. (2013). The origins and persistence of Black-White differences in women's labor force participation. *National Bureau of Economic Research*. https://doi.org/10.3386/w19040

Caliskan, A., Bryson, J. J., & Narayanan, A. (2016). Semantics derived automatically from language corpora necessarily contain human biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Charlesworth, T. E. S., & Banaji, M. R. (2019a). Gender in science, technology, engineering, and mathematics: Issues, causes, solutions. *Journal of Neuroscience*, *39*(37), 7228–7243. https://doi.org/10.1523/JNEUROSCI.0475-18.2019

Charlesworth, T. E. S., & Banaji, M. R. (2019b). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. *Psychological Science*, *30*(2), 174–192. https://doi.org/10.1177/0956797618813087

Charlesworth, T. E. S., & Banaji, M. R. (in press). Patterns of implicit and explicit attitudes II. Long-term change, regardless of group membership. *American Psychologist*.

Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of 65+ million words. *Psychological Science*. https://doi.org/10.1177/0956797620963619

Croft, A., Schmader, T., & Block, K. (2015). An underexamined inequality: Cultural and psychological barriers to men's engagement with communal roles. *Personality and Social Psychology Review*, *19*(4), 343–370. https://doi.org/10.1177/1088868314564789

Cryer, J. D., & Chan, K.-S. (2008, January). Time series analysis: With applications in R. In *Design* (2nd ed.). Springer Texts. https://doi.org/10.1007/978-0-387-75959-3

Diekman, A. B., Eagly, A. H., & Johnston, A. M. (2010). Social structure. In J. F. Dovidio, M. Hewstone, P. G. Glick, & V. M. Esses (Eds.), *The Sage handbook of prejudice, stereotyping, and discrimination* (pp. 209–224). Sage Publications.

Dorius, S. F., & Firebaugh, G. (2010). Trends in global gender inequality. *Social Forces*, *88*(5), 1941–1968. https://doi.org/10.1353/sof.2010.0040

Durkheim, E. (1924). *Sociologie et philosophie* [Sociology and philosophy]. Felix Alcan.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2019). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494.supp

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage Publications. https://doi.org/10.4135/9781446249222.n49

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*(1), 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

England, P., Levine, A., & Mishel, E. (2020). Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences*, 201918891. https://doi.org/10.1073/pnas.1918891117

Fetterolf, J. (2017, March 7). *Women make up 40% or more of workforce in many countries*. Pew Research Center. https://www.pewresearch.org/fact-tank/2017/03/07/in-many-countries-at-least-four-in-ten-in-the-labor-force-are-women/

Fiske, S. T. (2017). Prejudices in cultural contexts: Shared stereotypes (gender, age) versus variable stereotypes (race, ethnicity, religion). *Perspectives on Psychological Science*, *12*(5), 791–799. https://doi.org/10.1177/1745691617708204

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. https://doi.org/10.1037/pspa0000160

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), E3635–E3644. https://doi.org/10.1073/pnas.1720347115

Garikipati, S., & Kambhampati, U. (2020). Leading the fight against the pandemic: Does gender "really" matter? *SSRN Electronic Journal*, 1–16. https://doi.org/10.2139/ssrn.3617953

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The times they are a-changing . . . or are they not? A comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly*, *40*(3), 353–363. https://doi.org/10.1177/0361684316634081

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology*, *38*(1), 369–425. https://doi.org/0066-4308/87/0201-0369

Hoover, J., & Dehghani, M. (2019). The big, the bad, and the ugly: Geographic estimation with flawed psychological data. *Psychological Methods*. https://doi.org/10.1037/met0000240

Intellectual Property Office. (2019). Gender profiles in worldwide patenting: An analysis of female inventorship. https://www.gov.uk/ipo

Jackson, J. C., Van Egmond, M., Choi, V. K., Ember, C. R., Halberstadt, J., Balanovic, J., Basker, I. N., Boehnke, K., Buki, N., Fischer, R., Fulop, M., Fulmer, A., Homan, A. C., Van Kleef, G. A., Kreemers, L., Schei, V., Szabo, E., Ward, C., & Gelfand, M. J. (2019). Ecological and cultural factors underlying the global distribution of prejudice. *PLoS ONE*, *14*(9). https://doi.org/10.1371/journal.pone.0221953

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. *Frontiers in Psychology*, *6*(727), 1–24. https://doi.org/10.3389/fpsyg.2015.00727

Jost, J. T. (2015). Resistance to change: A social psychological perspective. *Social Research*, *82*(3), 607–636. https://muse.jhu.edu/article/603152

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, *107*(3), 371–392. https://doi.org/10.1037/a0037215

Koenig, A. M., & Eagly, A. H. (2019). Typical roles and intergroup relations shape stereotypes: How understanding social structure clarifies the origins of stereotype content. *Social Psychology Quarterly*, *82*(2), 205–230. https://doi.org/10.1177/0190272519850766

Lai, C. K., Hoffman, K. M., & Nosek, B. A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, *7*(5), 315–330. https://doi.org/10.1111/spc3.12023

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental*

Psychology: General, 145(8), 1001–1016. https://doi.org/10.1037/xge0000179

Lippmann, W. (1922). *Public Opinion*. Harcourt, Brace, and Company.

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *107*(3), 631–644. https://doi.org/http://dx.doi.org/10.1037/edu0000005

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. draw-a-scientist studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Moscovici, S. (2000). *Social representations: Explorations in social psychology*. Cambridge: Polity Press.

Nesbitt, M. N., & Penn, N. E. (2000). Gender stereotypes after thirty years: A replication of Rosenkrantz, et al. (1968). *Psychological Reports*, *87*(2), 493–511. https://doi.org/10.2466/pr0.2000.87.2.493

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., . . . Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Ofosu, E. K., Chambers, M. K., Chen, J. M., & Hehman, E. (2019). Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *Proceedings of the National Academy of Sciences*, *116*(18), 8846–8851. https://doi.org/10.1073/pnas.1806000116

Orchard, J., & Price, J. (2017). County-level racial prejudice and the Black-White gap in infant health outcomes. *Social Science and Medicine*, *181*, 191–198. https://doi.org/10.1016/j.socscimed.2017.03.036

Parker, L. R., Monteith, M. J., Moss-Racusin, C. A., & Van Camp, A. R. (2018). Promoting concern about gender bias with evidence-based confrontation. *Journal of Experimental Social Psychology*, *74*, 8–23. https://doi.org/10.1016/j.jesp.2017.07.009

Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. (2019). Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nature Human Behaviour*, 1–9. https://doi.org/10.1038/s41562-019-0686-3

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*. https://doi.org/10.1111/0022-4537.00239

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Varnum, M. E. W., & Grossmann, I. (2017a). Cultural change: The how and the why. *Perspectives on Psychological Science*, *12*(6), 174569161769997. https://doi.org/10.1177/1745691617699971

Varnum, M. E. W., & Grossmann, I. (2017b). Pathogen prevalence is associated with cultural changes in gender equality. *Nature Human Behaviour*, *1*(1). https://doi.org/10.1038/s41562-016-0003

Author Biographies

Tessa E. S. Charlesworth is a PhD candidate in the Department of Psychology at Harvard University.

Mahzarin R. Banaji is the Richard Clarke Cabot Professor of Social Ethics in the Department of Psychology at Harvard University.

Handling Editor: Margo Monteith