Social Cognition, Vol. 42, No. 4, 2024, pp. 317–333
© 2024 The Guilford Press. https://doi.org/10.1521/soco.2024.42.4.317

THE TEMPORAL AND DIRECTIONAL RELATIONSHIP BETWEEN GROUP-LEVEL IMPLICIT AND EXPLICIT GENDER BIAS

Yun Tang
Department of Psychology, University of Utah

Eric Hehman
Department of Psychology, McGill University

Jacqueline M. Chen
Department of Psychology, University of Utah

Explicit and implicit gender-science and gender-career biases have shifted toward neutrality in the past decade. Researchers speculate that these changes result from women's increased visibility in the science field and job market, but little is known about how the changes in group-level explicit and implicit gender-science and gender-career bias relate to one another over time. Building on contemporary models of group-level bias, this study investigates the temporal and directional relationship between group-level implicit and explicit gender bias between 2007 to 2016 using multivariate multilevel modeling. We found that lower group-level implicit bias in a previous month predicts lower group-level explicit bias in the following month. We also found evidence that group-level explicit bias in a previous month was positively associated with group-level implicit bias in the following month. These findings have practical and theoretical implications for understanding the bidirectional relationship between group-level implicit and explicit biases over time.

Keywords: social cognition, gender bias, quantitative models

Individuals' negative attitudes about social groups can result in preferential treatment that produces social inequalities (e.g., FitzGerald & Hurst, 2017; Rice & Barth, 2017; Rosen et al., 2021; Staats, 2015–2016). People's attitudes about a social group can be implicit or explicit (Gawronski & Bodenhausen, 2011; Rydell et al., 2006; Strack & Deutsch, 2004). Some theories posit that implicit attitude measures are thought to capture individuals' mental associations with a target group (e.g., associating women with the home), while explicit attitude measurements are thought

Address correspondence to Yun Tang, College of Social and Behavioral Sciences, University of Utah, 260 S. Central Campus Dr., Salt Lake City, UT 84112. E-mail: yun.tang@psych.utah.edu

to capture people's propositional reasoning processes (e.g., believing women should stay at home; Gawronski & Bodenhausen, 2006). Awareness of the limitations of individual-level implicit measures (Greenwald et al., 2022) and increased access to large-scale data collection have led some researchers to shift focus to investigating group-level attitudes, or attitudes held by large groups of individuals in a shared context, such as a geographic region or organization. Compared to individual-level implicit attitudes, group-level implicit attitudes have superior reliability, stronger correlations with explicit attitudes, and better predictive validity for behavioral outcomes (Calanchini et al., 2022; Hehman et al., 2019). Because these group-level estimates, which can be understood as the collective average attitudes shared by individuals in theoretically relevant situations (e.g., locations and time frame with shared culture), are associated with meaningful and theoretically consistent health, education, and criminal justice outcomes (Hehman et al., 2018; Leitner et al., 2018; Miller et al., 2015; Orchard & Price, 2017), they are not a mere statistical artifact of aggregation (Hannay & Payne, 2022). This research aims to build on contemporary theories of group-level attitudes to understand how group-level implicit and explicit attitudes influence one another over time.

Three theories have emerged to understand the bases of group-level attitudes and their real-world effects. First, although not a model explicitly focusing on group-level attitudes, the Prejudice-in-Places model argues that places can be prejudiced when they create disparate outcomes for social groups in those places (Murphy & Walton, 2013), indicating that the institutional and cultural influence of a place contributes to levels of bias. Second, the Bias of Crowds model reconciles the paradox between the fluctuation of implicit bias at the individual level and its stability at the group level. The Bias of Crowds model argues that group-level prejudice reflects systematic inequalities (Payne et al., 2017). Third, Calanchini et al. (2022) argued that regional-aggregated implicit and explicit bias may largely represent shared culture, including beliefs and norms in a geographic region. Although these contemporary models proposed distinct explanations of group-level implicit and explicit bias, their core ideas overlap in suggesting that group-level biases index a group's shared context with respect to an intergroup bias (cf. Connor & Evers, 2020, for another view). Consistent with these views, the current research investigates group-level bias shared within a temporal context.

Indeed, recent empirical work indicates that group-level attitudes not only reflect the social environment but also can be interpreted as a barometer for cultural shifts over the years. Just as how comparing the group-level bias in geographic regions could give us a sense of the regions' cultural environment, examining group-level bias in different time periods provides insight into the cultural shift driven by media representation and social movement. Researchers have proposed that the increasingly positive and complex representation of gay and lesbian people in the media contributes to the decrease in group-level implicit and explicit sexuality bias (Westgate et al., 2015). Moreover, changes in group-level implicit and explicit attitudes are associated with societal-level change, for example, legislation such as the legalization of same-sex marriage (Aksoy et al., 2020; Ofosu et al., 2019), and social movements, such as the Black Lives Matter movement (Sawyer & Gampa, 2018). Despite the growing evidence for the link between societal-level change and changes in group-level implicit and explicit bias, little attention has been paid to the *process* of these changes, specifically the directional relationship between the group-level implicit and explicit attitudes across time.

The current study examines the process of longitudinal changes in group-level implicit and explicit bias with respect to the association between gender and science and gender and career. To date, very few studies have directly examined the relationship between changes in group-level implicit and explicit bias. Charlesworth and Banaji (2019) found mixed results on the relationship between the change of group-level implicit and explicit bias in five different aggregated implicit and explicit attitude data sets on race, skin tone, age, disability, and body weight. They found evidence that implicit attitude change precedes explicit attitude change for race and skin-tone attitudes but not for age, disability, and body-weight attitudes, although exactly why this may be the case remains uncertain. Two recent studies by the same authors suggested that people's implicit and explicit gender-science and gender-career biases have been decreasing in the past decades across various demographic groups, and they proposed that the decrease is related to the increase in women's visibility in science and in the job market (Charlesworth & Banaji, 2021, 2022). While these studies focus on implicit and explicit bias levels over time, they examine the two constructs separately and cannot speak to how group-level implicit and explicit gender biases relate to one another during the decline in the past decade. We are aware of the idea that group-level implicit and explicit bias may be two measurements capturing a single construct in circumstances when they are highly correlated, in this case reflecting a region's strong cultural consensus (Calanchini et al., 2022). Yet, there is considerable heterogeneity in correlations between group-level implicit and explicit bias depending on the type of bias (Calanchini et al., 2022), and empirical evidence shows that the change in implicit and explicit bias does not always move in the same direction (e.g., explicit body-weight bias decreased while implicit body-weight bias increased in the past decade; Charlesworth & Banaji, 2019). It is safe to conclude that group-level implicit and explicit biases are not always the same thing, and it is plausible that one may precede the other. Even when they are highly correlated, other factors influence group-level implicit and explicit bias in addition to the potential overlapping culture they represent. Unraveling the processes underlying the decline of group-level gender implicit and explicit bias contributes to the understanding of the theoretical relationship between these two constructs through offering valuable insights into how their unique components are interrelated and have practical implications for interventions to reduce gender bias.

In this study, we analyzed the temporal and directional relationship between group-level implicit and explicit gender-science and gender-career bias changes between 2007 and 2016[1] using the Project Implicit data sets. Although many stud-

1. There was an immediate mean drop of two standard deviations in explicit gender-science and gender-career bias in 2016, most likely due to changes in the format of the explicit bias measures (Charlesworth & Banaji, 2021; N. Frost, personal communication, June 26, 2023). The current study focuses on the data before the change in measurement.

ies have analyzed the Project Implicit data sets, and some of them focused on the change in bias over time (Ravary et al., 2023; Somo et al., 2021), this study will be the first to employ a multilevel modeling approach to understand the causal relationship between group-level gender implicit and explicit bias during the change over time. Having multiple independent teams of researchers analyze the same data set using different analytical techniques and choices to test similar research questions is essential for achieving scientific reproducibility (Botvinik-Nezer et al., 2020; Silberzahn et al., 2018).

The current study used multivariate multilevel modeling to examine the cross-lagged relationship between the monthly aggregated mean of implicit and explicit gender bias while controlling for participants' demographics. We implemented a random intercept cross-lagged panel model (RI-CLPM), which accounts for both the temporal and the time-invariant stability of the two variables, so that the estimates of the cross-lagged effects can accurately capture how the changes in the level of the two variables influence one another above and beyond the shared contribution of them (Hamaker et al., 2015; Mulder & Hamaker, 2021). Specifically, our model tested if group-level explicit bias in a previous month is associated with the group-level implicit bias in the following month or vice versa. The decision to use a month as an aggregated unit balanced the considerations of choosing a time interval that can capture the change and gaining enough group-level data points to run the multivariate multilevel model. It also aligns with the approach to depict change over time in Charlesworth and Banaji (2019).

On the basis of the observed decreasing trend of implicit and explicit group-level gender-science bias and gender-career bias in the past decade (Charlesworth & Banaji, 2019), we hypothesized that group-level implicit bias is significantly associated with the previous group-level explicit bias on top of the previous group-level implicit bias. We also hypothesized that the change in group-level implicit bias is significantly related to the subsequent change in group-level explicit bias on top of the previous group-level explicit bias. In addition, because past research has found that men employed in the fields of science, technology, engineering, and mathematics (STEM) hold stronger gender-science implicit and explicit bias than women in STEM (Carli et al., 2016; Cundiff et al., 2013; Smyth & Nosek, 2015), we examined whether the relationship between group-level implicit and explicit bias may be different for men in STEM versus women in STEM. We also conducted robustness checks by excluding control variables and aggregating the time unit differently.

METHODS

SOURCES OF DATA

Gender-Science Implicit and Explicit Bias. We used publicly available data from Project Implicit (https://implicit.harvard.edu/implicit/takeatest.html) collected between 2007 and 2016. During this time, the gender-science Implicit Association Test (IAT) data sets collected 501,560 responses in the United States. Participants were excluded if more than 10% of their responses were faster than 300 ms (Greenwald et al., 2003) or if they did not finish the IAT task or the survey questions, resulting in 291,164 (32.7% male) observations included in primary analyses. The data sets include measures of implicit and explicit gender-science bias, people's attitudes and experiences with science, and demographic information.

In the gender-science IAT, participants were asked to categorize stimuli that are related to either female or male with stimuli that are related to either the science or the liberal arts field. Participants' responses were then calculated into IAT *D* scores, showing the direction and strength of the implicit association (Greenwald et al., 2003). A positive direction suggests a stronger association of women with liberal arts and men with science compared to the association of women with science and men with liberal arts (Greenwald et al., 2022).

Explicit gender-science bias is measured by two items, which asked participants how much they associate science or liberal arts with males and females (e.g., Please rate how much you associate the following domains with males or females: Science). The items are on a 7-point scale, from 1 (*strongly female*) to 7 (*strongly male*). To align with the relative nature of IAT, and consistent with previous work, we calculated a relative bias score by subtracting the liberal arts item from the science item. The higher the score, the more participants reported congruent gender-STEM bias (i.e., associating males with science and/or females with liberal arts).

In addition, considering how science identity is closely related to implicit and explicit gender-science bias (Smyth & Nosek, 2015; Zitelny et al., 2017), we used two items in the data sets that asked participants how much they liked science and averaged their responses into one score to get participants' science identity ($r = .802$, $p < .001$). The options ranged from 1 (*strongly disliked*) to 5 (*strongly liked*).

Demographics of the participants (gender, age, education attainment, political orientation, etc.) were also collected. The data sets also include information on the participants' major(s). Participants who reported majoring in biological/life sciences, computer and information sciences, engineering, mathematics, physical sciences/science technologies, health professions or related sciences, and psychology are considered as majoring in the STEM field. About 53% of the participants reported that they majored in the STEM field.

Gender-Career Implicit and Explicit Bias. The gender-career IAT data set includes the results that reflect participants' implicit association of males with career and females with family ($n = 707{,}120$; 30.8% male). Identical to the gender-science data set, participants' implicit attitude is operationalized as the IAT *D* score. Participants' explicit bias is measured by two items: how strongly they associate the career with males and females (career item), and how strongly they associate the family with males and females (family item). The items are on a 7-point scale, from 1 (*strongly female*) to 7 (*strongly male*). Consistent with the gender-STEM data set, explicit bias was operationalized as the difference between the home and work items, with higher scores indicating that participants associated women more with family and men more with work. The gender-career IAT data set also includes participants' demographics (gender, age, education attainment, political orientation, etc.). The gender-career data set we used included participants' responses collected in the United States from 2007 to 2016.

ANALYTIC APPROACH

To capture the changes in group-level implicit and explicit bias in the past years, we aggregated the Project Implicit raw data for each data set to calculate the monthly mean implicit and explicit bias across participants. We aggregated and restructured these data by month, along with the monthly mean of participants' age, education attainment, political orientation, and proportion of participants' gender. In the restructured data, each row was the average implicit and explicit bias per month, the previous month, and aggregated demographics for those months (see Figure 1).

To test our hypothesis, we conducted analyses using the nmle package in R (Pinheiro et al., 2022) in a multilevel framework with data clustered within months. We regressed a month's average explicit and implicit bias on the explicit and implicit biases of the previous month (Figure 2). The "a" paths can be considered as the measurement of the autoregressive effect of the previous month's mean bias on the mean bias of the following month. The two diagonal "b" paths represent the effect of group-level implicit bias in the previous month on group-level current explicit bias and the effect of group-level explicit bias in the previous month on current group-level implicit bias. The model fits with our research focus on the impact of relative change in group-level implicit bias on group-level explicit bias and vice versa (Orth et al., 2021). In addition, considering the potential nonindependence between monthly aggregated implicit bias and explicit bias based on past findings (Calanchini et al., 2022; Hehman et al., 2019), the model allows for the correlation between these two variables (indicated by the two curved lines in Figure 2), which plays an important role in controlling for the effect of each other when predicting the outcome variables (Kenny, 2018). The model also allows for different error variances for the implicit bias and explicit bias outcomes. In order to test if the relationship between group-level implicit and explicit bias may be different for men versus women in STEM, we added participants' gender and interaction with the other bias variables and ran the model in a subset including only participants who self-reported majoring in the STEM fields. In addition to testing this model in the restructured data set of the group-level implicit and explicit bias based on the calendar month (i.e., beginning on the first day of the month), we tested whether our conclusions were robust to this arbitrary clustering by repeating analyses based on clustering months from the 16th of the previous month to the 15th of the subsequent month and from the 21st of the previous month to the 20th of the subsequent month.

RESULTS

To explore the relationship between group-level implicit and explicit gender-science and gender-career attitudes, we used multivariate multilevel modeling to

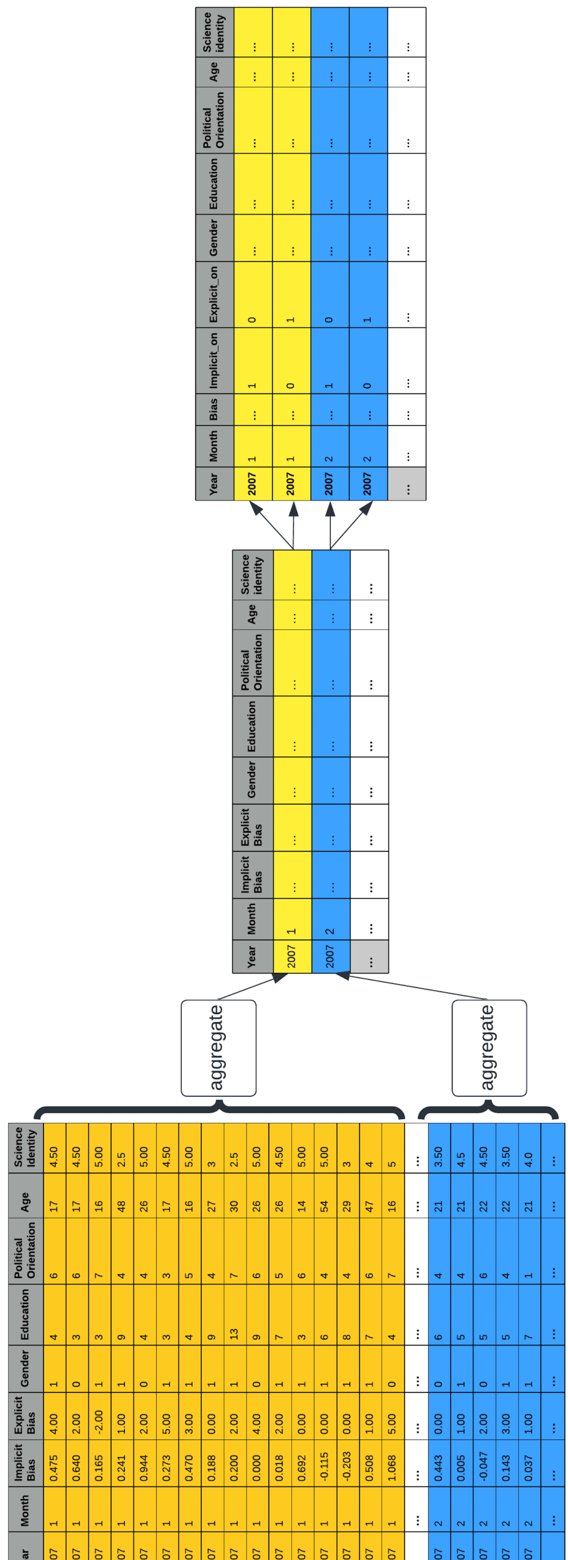


FIGURE 1. The process of restructuring the data set for multivariate multilevel modeling.

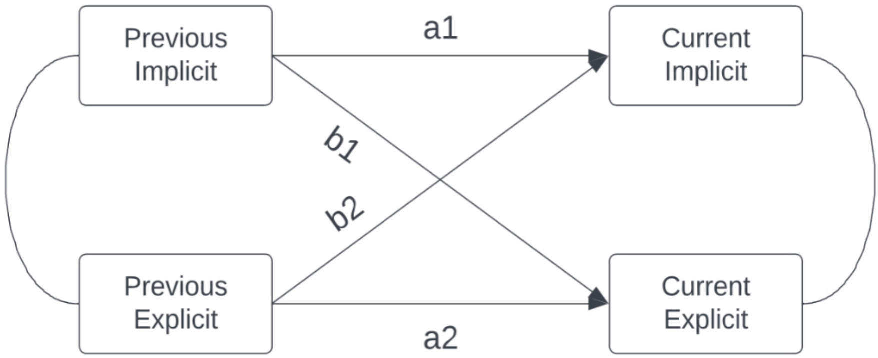


FIGURE 2. Proposed model of the relationship between monthly aggregated implicit and explicit gender bias.

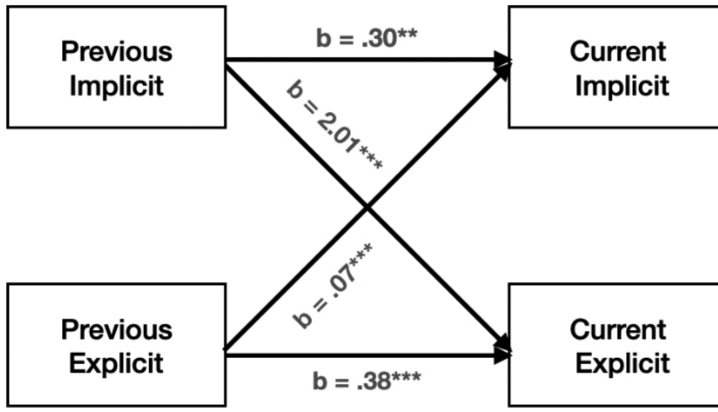
disentangle the cross-lagged relationship while taking the correlational nature of the variables into consideration. We included demographic variables because they may be influential in predicting people's implicit and explicit gender-science and gender-career biases.

GENDER-SCIENCE BIAS

Model 1A: Entire Sample (2007–2016)

We first examined the null model to test the degree of nonindependence in the data. The Rho = 0.84, indicating a strong positive correlation across implicit bias and explicit bias in the same month. We first examined the cross-lagged relationship between group-level implicit bias and group-level explicit bias in the general population. When controlling for demographics and science identity, the group-level implicit bias in the previous month is significantly associated with the group-level implicit bias in the following month, $b = .30$, $SE = .10$, 95% CI [.11, .49], $\beta = .08$, $p = .002$. Group-level explicit bias in the previous month was also significantly associated with the group-level explicit bias in the following month, $b = .38$, $SE = .11$, 95% CI [.17, .59], $\beta = .06$, $p < .001$. These significant carryover effects are consistent with the overall stability of group-level bias reported in previous literature (Vuletich & Payne, 2019).

More critical to our research questions, results indicated a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, $b = .07$, $SE = .02$, 95% CI [.03, .11], $\beta = .07$, $p < .001$ (see Figure 3). For each one-point decrease in the group-level explicit bias in the previous month, the expected group-level implicit bias decreased by .07 points in the subsequent month, which is more than two times the standard deviation of group-level implicit bias ($SD = .023$). In addition, we found a significant cross-lagged effect between the group-level implicit bias in the



Note. Paths show unstandardized coefficients.
$^{**}p < .01$ $^{***}p < .001$.

FIGURE 3. The temporal and directional relationship between group-level gender-science implicit and explicit bias. Paths show unstandardized coefficients. $^{**}p < .01$. $^{***}p < .001$.

previous month and the group-level explicit bias in the following month, $b = 2.01$, $SE = .48$, 95% CI [1.08, 2.94], $\beta = .08$, $p < .001$. For each one-point decrease in the group-level implicit bias in the previous month, the expected group-level explicit bias decreased by 2.01 points in the subsequent month, which is more than 20 times the standard deviation of group-level explicit bias ($SD = .11$).

Robustness and Replication. We tested for robustness by running the same model, absent control variables, and obtained the same pattern of results (see Table 1). Another robustness check was to cluster our data differently, aggregating the data from the 16th of the previous month to the 15th of the subsequent month and from the 21st of the previous month to the 20th of the subsequent month. This replicated the results reported here except for one case that previous explicit bias was not significantly related to implicit bias in the following month when we aggregated the month from the 15th of the previous month to the 15th of the subsequent month and ran the model without control variables.

Overall, we consider results consistent, and provide evidence that our conclusions were not driven by the arbitrary clustering of biases by month.

Model 2: Moderation by Gender Within STEM Fields (2007–2016)

We also examined whetther the cross-lagged relationship varied by gender between those who are or were in the STEM field. With gender (contrast coded as −1 = Male, 1 = Female) we found that when controlling for other demographics (age, educational attainment, political orientation) and science identity, gender did not moderate the relationship between previous implicit bias and explicit bias,

TABLE 1. Coefficients of the Autoregressive and Cross-Lagged Paths in Model 1

Gender-science data set	With control variables					Without control variables				
	Unstandardized B	Standardized Beta	*SE*	*T* value	*p* value	Unstandardized B	Standardized Beta	*SE*	*T* value	*p* value
Previous explicit on following explicit	0.38	0.06	0.11	3.51	.001	0.38	0.06	0.11	3.53	.001
Previous implicit on following implicit	0.3	0.08	0.1	3.13	.002	0.59	0.1	0.1	6.21	<.001
Previous implicit on following explicit	2.01	0.07	0.48	4.23	<.001	2.3	0.08	0.47	4.88	<.001
Previous explicit on following implicit	0.07	0.07	0.02	3.58	.001	0.07	0.07	0.02	3.23	.001

Gender-career data set	With control variables					Without control variables				
	Unstandardized B	Standardized Beta	*SE*	*T* value	*p* value	Unstandardized B	Standardized Beta	*SE*	*T* value	*p* value
Previous explicit on following explicit	0.69	0.13	0.07	9.67	<.001	0.69	0.13	0.07	9.59	<.001
Previous implicit on following implicit	0.49	0.04	0.08	6.23	<.001	0.54	0.06	0.08	6.63	<.001
Previous implicit on following explicit	1.68	0.04	0.56	2.99	.003	1.745	0.04	0.57	3.06	.003
Previous explicit on following implicit	0.04	0.14	0.01	4.34	<.001	0.04	0.14	0.01	4.13	.001

$b = .29$, $SE = .31$, 95% CI [−.31, .89], $\beta = .03$, $p = .346$, or previous explicit bias and implicit bias, $b = .002$, $SE = .02$, 95% CI [−.03, .036], $\beta = -.08$, $p = .920$. Interestingly, the effect of previous implicit bias on the implicit bias in the subsequent month differed by gender, $b = .14$, $SE = .26$, 95% CI [.003, .27], $\beta = -.08$, $p = .046$, suggesting that implicit bias is less stable for men than women in STEM. Replicating the result of Model 1A, previous monthly aggregated explicit bias had a significant main effect on the monthly aggregated implicit bias, $b = .04$, $SE = .02$, 95% CI [.003, .07], $\beta = .08$, $p = .035$. Yet, unlike the result of Model 1A, previous monthly aggregated implicit bias was not significantly associated with the monthly-aggregated explicit bias in the following month, $b = .29$, $SE = .26$, 95% CI [−.21, .79], $\beta = .07$, $p = .255$. When we ran the same model without control variables, we obtained results in the same pattern, with one exception (see Supplementary Material).

GENDER-CAREER BIAS

Model 1B: Entire Sample (2007–2016)

Data from the gender-science models suggest that previous explicit bias causes changes in future implicit bias. But given the complicated model, we would have more confidence in the relationship to the extent it is observed in another similar data set. To this end, we examined the gender-career data set for converging evidence. Similar to the gender-science data, null models indicated the Rho = 0.81, indicating a strong positive correlation across implicit bias and explicit bias in the same month. We found similar patterns between the relationship between group-level implicit and explicit bias as in the gender-science data set, such that there were significant cross-lagged effects between group-level explicit bias in the previous month and group-level implicit bias in the following month, and the reverse was also true. In addition, the results held regardless of whether control variables were included in the model.

Robustness and Replication. We also conducted the same robustness check process in the gender-career data set as we did in the gender-science data sets. Again, the significant cross-lagged relationship between group-level explicit bias in the previous month and group-level implicit bias in the following month emerged, and the group-level implicit bias in the previous month and group-level explicit bias in the following month was not significant (see Supplementary Material). Across the different studies and model specifications, the group-level explicit bias in the previous month and group-level implicit bias in the following month relationship appears far more robust than the reverse.

DISCUSSION

The current research provides us with a greater understanding of the temporal and directional changes in group-level implicit and explicit gender bias. Across two data sets regarding gender-science and gender-career relationships, we found a consistent cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month. This relationship was robust to a number of different specifications of the samples and demographics included. We also found some regular, but less consistent, evidence of the reverse effect: Group-level implicit gender bias in a previous month is associated with group-level explicit gender bias in the following month. While there is some evidence for this relationship, it was more dependent on the sample and model specifications. The findings are able to delineate the amplified shift in attitude consensus in the past decade and, more importantly, provide insight into how group-level explicit and implicit gender-science and gender-career attitude changes unfold over time.

Our findings may have several underlying causes. The significant relationship between the change in group-level explicit bias and group-level implicit bias is consistent with a norm-based account of collective attitude change (Tankard & Paluck, 2016). Social changes (e.g., social movements) could signal group-level attitude change, which produces shifts in cultural norms and media change that catalyze heated and widespread discussion on gender-science bias that goes along with the shift in people's attitudes (Sawyer & Gampa, 2023). And indeed, recent work has found greater implicit-explicit correlations among biases receiving more public attention (and thus sending stronger normative signals), such as sexuality or race-based biases, relative to disability or weight-based biases (Calanchini et al., 2022). Also, research has documented relatively immediate shifts in attitudes as a result of major, societally impactful events such as the highly publicized murder of George Floyd by police (Primbs et al., 2022) and same-sex marriage legislation passed in various states (Ofosu et al., 2019). Yet it is less clear what strong and salient gender-related event might have happened in the times covered by the present data (2007–2016) that signals strong public norms. Instead, we speculate that what might be happening here is a slow but consistent change in gender attitudes in North America over time, driving the present effects. Future research could work on identifying major shifts that may contribute to the decreased group-level implicit and explicit gender bias between 2007 and 2016 and further explore more types of biases and how social events affect the temporal relationship of group-level explicit and implicit biases.

Another possible mechanism is that as implicit associations change (e.g., stronger science-female association) in accordance with environmental stimuli (e.g., shift in gender ratio in the science field), people may reflect on their propositional reasoning of what they previously believed, which leads to changes in their reported explicit bias (Gawronski & Bodenhausen, 2006). Future research could capitalize on impending social changes by examining how individuals' and groups' implicit and explicit attitudes change associated with those changes and investigating the likelihood of bidirectional influences between group-level implicit and explicit bias. Future studies may investigate the impact of contextual factors on the sequential change of group-level explicit bias and implicit bias, for instance, taking advantage of the variation in the speed of increasing women's representation across different STEM subfields to explore how the rising visibility of women affects the temporal relationship of the group-level explicit and implicit bias (National Science Foundation, 2013). In addition to societal-level changes, future endeavors on bias reduction interventions in organizations may also investigate the potential of diminishing implicit bias by first tackling explicit bias and fostering inclusive cultural change within the organization. This is advantageous because, whereas implicit bias has been known to be difficult to change with interventions (Lai et al., 2014), explicit bias is thought to be more malleable and subject to interventions than implicit bias, through approaches such as educational strategies that promote increasing understanding and appreciation of diverse groups (e.g., multicultural education programs), and strong communications emphasizing the norm that prejudice is unacceptable (Dovidio et al., 2013).

LIMITATIONS

Although our theoretical model is predicated on directional relationships, because of the observational nature of the data, we are cautious in interpreting the relationships between previous-month explicit bias and subsequent-month implicit bias and between previous-month implicit bias and subsequent-month explicit bias as causal. The longitudinal nature of the data does ensure temporal precedence, providing much stronger evidence for causal claims. Yet, like any observational data, unaccounted-for confounders or the "third variable problem" might have led to these spurious relationships. Overall, the current study is consistent with the idea of the potential reciprocal relationship between the change in group-level explicit and implicit bias, but evidence from other approaches with different limitations is necessary to buttress this conclusion.

Another limitation of this study is the nonrepresentative sample used. The sample had more women, was younger, and was more liberal than the general population. Also, the respondents self-selected to participate in the IAT tests. Because the sample is not representative of the general population nor random, the generalizability of our conclusions is questionable, and we hope it will be addressed by future research (although we note that previous research using Project Implicit data has found similar results with representative samples; Ofosu et al., 2019). In addition, the relative nature of IAT does not enable us to know if the weakening of the stereotypical association or the strengthening of the nonstereotypical association played a dominant role in the change of group-level gender bias. Future work could use no-relative measures (e.g., Implicit Relational Assessment Procedure [IRAP]; Farrell & McHugh, 2017) or the high-validity single-category IAT (Axt et al., 2022) to further disentangle the complexity of the group-level gender bias change. Finally, in the gender-science data set, implicit and explicit bias were decreasing over time. If group-level explicit and implicit biases were increasing or remain relatively stable, would change in explicit bias still lead to change in implicit bias and vice versa? Future research may work on replicating the findings in other IAT data sets that have also been observed to have decreased group-level explicit and implicit bias in the past years, such as the skin-tone IAT (Charlesworth & Banaji, 2019).

CONCLUSION

The current research used a multivariate multilevel modeling approach to examine the relationship between group-level implicit and explicit bias change in the gender-science and gender-career Project Implicit data sets. Evidence consistently indicates that group-level explicit bias has a unique contribution in influencing subsequent group-level implicit bias above and beyond the measurement equivalence of group-level implicit bias across time. We also found some evidence that the decrease in group-level explicit bias also depends on the temporal level of group-level implicit bias at the previous occasion. These results not only have theoretical implications for the association between group-level implicit and explicit bias change but also introduce unique prospects for future studies on the long-term reduction of group-level bias, including bias reduction interventions.

REFERENCES

Aksoy, C. G., Carpenter, C. S., De Haas, R., & Tran, K. D. (2020). Do laws shape attitudes? Evidence from same-sex relationship recognition policies in Europe. *European Economic Review*, *124*, 103399. https://doi.org/10.1016/j.euroecorev.2020.103399

Axt, J., Buttrick, N., & Feng, R. Y. (2022). A comparative investigation of the predictive validity of four indirect measures of bias and prejudice. *Personality and Social Psychology Bulletin*, *50*(6), 871–888. https://doi.org/10.1177/01461672221150229

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., . . . Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Calanchini, J., Hehman, E., Ebert, T., Esposito, E., Simon, D., & Wilson, L. (2022). Regional intergroup bias. *Advances in Experimental Social Psychology*, *66*, 281–337. https://doi.org/10.1016/bs.aesp.2022.04.003

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly*, *40*(2), 244–260. https://doi.org/10.1177/0361684315622645

Charlesworth, T. E. S., & Banaji, M. R. (2019). 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. (2021). Patterns of implicit and explicit stereotypes III: Long-term change in gender stereotypes. *Social Psychological and Personality Science*, 13(1), 14–26. https://doi.org/10.1177/1948550620988425

Charlesworth, T. E. S., & Banaji, M. R. (2022). Patterns of implicit and explicit attitudes: IV. Change and stability from 2007 to 2020. *Psychological Science, 33*(9), 1347–1371. https://doi.org/10.1177/09567976221084257

Connor, P., & Evers, E. R. K. (2020). The bias of individuals (in crowds): Why implicit bias is probably a noisily measured individual-level construct. *Perspectives on Psychological Science, 15*(6), 1329–1345. https://doi.org/10.1177/1745691620931492

Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education*, *16*(4), 541–554. https://doi.org/10.1007/s11218-013-9232-8

Dovidio, J. F., Gaertner, S. L., & Thomas, E. L. (2013). Intergroup relations. In J. M. Levine (Ed.), *Group processes* (pp. 323–349). Psychology Press.

Farrell, L., & McHugh, L. (2017). Examining gender-STEM bias among STEM and non-STEM students using the Implicit Relational Assessment Procedure (IRAP). *Journal of Contextual Behavioral Science*, *6*(1), 80–90.

FitzGerald, C., & Hurst, S. (2017). Implicit bias in healthcare professionals: A systematic review. *BMC Medical Ethics*, *18*(1), 19. https://doi.org/10.1186/s12910-017-0179-8

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2011). The associative–propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59–127. https://doi.org/10.1016/B978-0-12-385522-0.00002-0

Greenwald, A. G., Brendl, M., Cai, H., Cvencek, D., Dovidio, J. F., Friese, M., Hahn, A., Hehman, E., Hofmann, W., Hughes, S., Hussey, I., Jordan, C., Kirby, T. A., Lai, C. K., Lang, J. W. B., Lindgren, K. P., Maison, D., Ostafin, B. D., Rae, J. R., . . . Wiers, R. W. (2022). Best research practices for using the Implicit Association Test. *Behavior Research Methods*, *54*(3), 1161–1180. https://doi.org/10.3758/s13428-021-01624-3

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hannay, J. W., & Payne, B. K. (2022). Effects of aggregation on implicit bias measurement. *Journal of Experimental Social Psychology*, *101*, 104331. https://doi.org/10.1016/j.jesp.2022.104331

Hehman, E., Calanchini, J., Flake, J. K., & Leitner, J. B. (2019). Establishing construct validity evidence for regional measures of explicit and implicit racial bias. *Journal of Experimental Psychology: General*, *148*(6), 1022–1040. https://doi.org/10.1037/xge0000623

Hehman, E., Flake, J. K., & Calanchini, J. (2018). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science*, *9*(4), 393–401. https://doi.org/10.1177/1948550617711229

Kenny, D. A. (2018). Reflections on the actor–partner interdependence model. *Personal Relationships*, *25*(2), 160–170. https://doi.org/10.1111/pere.12240

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology. General*, *143*(4), 1765–1785. https://doi.org/10.1037/a0036260

Leitner, J. B., Hehman, E., & Snowden, L. R. (2018). States higher in racial bias spend less on disabled Medicaid enrollees. *Social Science & Medicine*, *208*, 150–157. https://doi.org/10.1016/j.socscimed.2018.01.013

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/10.1037/edu0000005

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling*, *28*(4), 638–648. https://doi.org/10.1080/10705511.2020.1784738

Murphy, M. C., & Walton, G. M. (2013). From prejudiced people to prejudiced places: A social-contextual approach to prejudice. In C. Stangor & C. S. Crandall (Eds.), *Stereotyping and prejudice* (pp. 181–203). Psychology Press.

National Science Foundation. (2013). *Women, minorities, and persons with disabilities in S&E*. Retrieved May 12, 2022, from https://www.nsf.gov/statistics/women/

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 of the United States of America*, *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 & Medicine*, *181*, 191–198. https://doi.org/10.1016/j.socscimed.2017.03.036

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. https://doi.org/10.1037/pspp0000358

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*(4), 233–248. https://doi.org/10.1080/1047840X.2017.1335568

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Development Team. (2022). *nlme: Linear and nonlinear mixed effects models.*

Primbs, M., Holland, R., Maatman, F. O., Lansu, T., Faure, R., & Bijlstra, G. (2022). *The effects of the 2020 BLM protests on racial bias in the United States*. PsyArXiv. https://osf.io/preprints/psyarxiv/x7uch

Ravary, A., Bartz, J. A., & Baldwin, M. W. (2023). Variability across time in implicit weight-related bias: Random noise or meaningful fluctuations? *Journal of Personality and Social Psychology*, *125*(5), 991–1017. https://doi.org/10.1037/pspa0000345

Rice, L., & Barth, J. M. (2017). A tale of two gender roles: The effects of implicit and explicit gender role traditionalism and occupational stereotype on hiring decisions. *Gender Issues*, *34*(1), 86–102. https://doi.org/10.1007/s12147-016-9175-4

Rosen, E., Garboden, P. M., & Cossyleon, J. E. (2021). Racial discrimination in housing: How landlords use algorithms and home visits to screen tenants. *American Sociological Review*, *86*(5), 787–822. https://doi.org/10.1177/00031224211029618

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science*, *17*(11), 954–958. https://doi.org/10.1111/j.1467-9280.2006.01811.x

Sawyer, J., & Gampa, A. (2018). Implicit and explicit racial attitudes changed during Black Lives Matter. *Personality and Social Psychology Bulletin*, *44*(7), 1039–1059. https://doi.org/10.1177/0146167218757454

Sawyer, J. E., & Gampa, A. (2023). Social movements as parsimonious explanations for implicit and explicit attitude change. *Personality and Social Psychology Review, 27*(1), 28–51.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Cervantes, I. F., . . . Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Smyth, F. L., & Nosek, B. A. (2015). On the gender–science stereotypes held by scientists: Explicit accord with gender-ratios, implicit accord with scientific identity. *Frontiers in Psychology*, *6*, 415. https://doi.org/10.3389/fpsyg.2015.00415

Somo, A., Sadler, M., & Devos, T. (2021). Implicit Black-weapon associations weakened over time in increasingly multiethnic metropolitan areas. *Analyses of Social Issues and Public Policy*, *21*(1), 520–540. https://doi.org/10.1111/asap.12228

Staats, C. (2015–2016). Understanding implicit bias: What educators should know. *American Educator*, *39*(4), 29–33, 43.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review, 10*(1), 181–211. https://doi.org/10.1111/sipr.12022

Vuletich, H. A., & Payne, B. K. (2019). Stability and change in implicit bias. *Psychological Science*, *30*(6), 854–862. https://doi.org/10.1177/0956797619844270

Westgate, E. C., Riskind, R. G., & Nosek, B. A. (2015). Implicit preferences for straight people over lesbian women and gay men weakened from 2006 to 2013. *Collabra, 1*(1), 1. https://doi.org/10.1525.collabra.18

Zitelny, H., Shalom, M., & Bar-Anan, Y. (2017). What is the implicit gender-science stereotype? Exploring correlations between the gender-science IAT and self-report measures. *Social Psychological and Personality Science, 8*(7), 719-735. https://doi.org/10.1177/1948550616683017

SUPPLEMENTARY MATERIAL

Gender-STEM bias (2007–2016)

Moderation by gender within STEM fields (2007–2016) without covariates

Unlike what is found in model 2 with control variables included, there was a significant cross-lagged effect found between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = .61, *SE* = .26, 95% CI [.10, 1.13], β = .11, *p* = .02. All of the other paths replicated. The significant interaction effect of previous implicit bias on the implicit bias in the subsequent month again differed by gender, b = .21, *SE* = .08, 95% CI [.06, .36], β = .08, *p* = .007; when controlling for other demographics and science identity, gender did not moderate the relationship between previous implicit bias and explicit bias, b = .42, *SE* = .32, 95% CI [–.22, 1.04], β = .05, *p* = .19, or previous explicit bias and implicit bias, b = –.02, *SE* = .02, 95% CI [–.06, .01], β = –.10, *p* = .22, *SE* = .019, 95% CI [.01, .087], β = .09, *p* = .01.

Robustness Check aggregating the data from the 16th of the previous month to the 15th of the subsequent month

With covariates

In order to perform a robustness check for our findings for Model 1A in the gender-STEM dataset, we re-cluster our aggregation unit by "16th to 15th" (e.g., Jan. 16th to Feb. 15th is considered as a cluster unit) instead of the calendar month, which is what we did in the main analysis. We ran the same model and found the exact same pattern. The results suggested that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .236, *SE* = .11, 95% CI [.13, .46], β = .05, $p = .04$; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .43, *SE* = .12, 95% CI [.20, .66], β = .06, $p < .001$. We also find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .06, *SE* = .02, 95% CI [.01, .11], β = .07, $p = .01$. As we find in the other models, we also found a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = 1.38, *SE* = .51, 95% CI [.39, 2.38], β = .05, $p = .01$.

Without covariates

We ran the same model, absent control variable. Evidence showed that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .67, *SE* = .11, 95% CI [.45, .88], β = .08, $p < .001$; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .41, *SE* = .12, 95% CI [.17, .64], β = .06, $p < .001$. We also find a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = 1.81, *SE* = .52, 95% CI [.80, 2.83], β = .06, $p = < .001$. Yet, unlike what we find in the other models, we did not find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .04, *SE* = .03, 95% CI [–.01, 2.38], β = .07, $p = .14$.

Robustness Check aggregating the data from the 21st of the previous month to the 20th of the subsequent month

With covariates

Additionally, we re-cluster our aggregation unit by “21st to 20th” (e.g., Jan. 21st to Feb. 20th is considered as a cluster unit). We ran the same model and found the exact same pattern as what we found in the main analysis. The results suggested that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .48, *SE* = .11, 95% CI [.27, .70], β = .08, $p < .001$; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .43, *SE* = .12, 95% CI [.20, .66], β = .07, $p < .001$. We also find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .07, *SE* = .03, 95% CI [.02, .12], β = .08, $p < .001$. As we find in the other models, we also found a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = 1.75, *SE* = .52, 95% CI [.73, 2.75], β = .06, $p < .001$.

Without covariates

Again, we ran the same model in the “21st to 20th” aggregated dataset, absent control variable. Evidence showed that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .62, *SE* = .11, 95% CI [.41, .83], β = .09, $p < .001$; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .41, *SE* = .12, 95% CI [.18, .64], β = .06, $p < .001$. We find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = 1.88, *SE* = .51, 95% CI [.88, 2.89], β = .07, *p* < .001. In addition, we also found a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = 1.88, *SE* = .51, 95% CI [.88, 2.89], β = .07, *p* < .001.

Gender-Career bias (2007–2016)

Robustness Check aggregating the data from the 16th of the previous month to the 15th of the subsequent month

With covariates

Just as what we did for the robustness check for our findings in the gender-science dataset, we first re-clustered our aggregation unit by "16th to 15th" (e.g., Jan. 16th to Feb. 15th is considered as a cluster unit). We ran the Model 1b and found the exact same pattern except for the insignificance of the relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = –.17, *SE* = .62, 95% CI [–1.38, 1.04], β = –.003, *p* = .78. As we consistently find in the U.S. gender-science dataset, the results suggested that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .30, *SE* = .10, 95% CI [.10, .51], β = .02, *p* = .004; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .87, *SE* = .09, 95% CI [.70, 1.05], β = .12, *p* < .001. We also find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .06, *SE* = .02, 95% CI [.02, .10], β = .13, *p* = .002.

Without covariates

We ran the same model, absent control variable. Evidence showed that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .35, *SE* = .11, 95% CI [.13, .58], β = .08, *p* = .002; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .88, *SE* = .09, 95% CI [.70, 1.05], β = .06, *p* < .001. We also find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .06, *SE* = .09, 95% CI [.03, .09], β = .06, *p* = .004. Also, identical to what we find when running the model the with the control variables in the "16th to 15th" clustered gender-career dataset, we did not find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = –.12, *SE* = .62, 95% CI [–1.35, 1.11], β = .07, *p* = .85

Robustness Check aggregating the data from the 21st of the previous month to the 20th of the subsequent month

With control variable

Additionally, we re-cluster our aggregation unit by "21st to 20th" (e.g., Jan. 21st to Feb. 20th is considered as a cluster unit). The results suggested that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .37, *SE* = .08, 95% CI [.22, .52], β = .02, p < .001; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .84, *SE* = .19, 95% CI [.72, .95], β = .07, *p* < .001. We also find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .05, *SE* = .01, 95% CI [.03, .06], β = .20, *p* < .001. Yet, we did not find a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = .53, *SE* = .55, 95% CI [–.55, 1.59], β = .01, *p* = .34.

Without control variable

Again, we ran the same model in the "21st to 20th" aggregated dataset, absent control variable. Evidence showed that group-level implicit and explicit bias has high stability, such that the group-level implicit bias in a month is related to the group-level implicit bias in the following month, b = .61, *SE* = .08, 95% CI [.45, .78], β = .03, *p* < .001; the group-level explicit bias in a month is related to the group-level explicit bias in the following month, b = .81, *SE* = .06, 95% CI [.70, .93], β = .19, *p* < .001. We find a significant cross-lagged relationship between the group-level explicit bias in the previous month and the group-level implicit bias in the following month, b = .02, *SE* = .01, 95% CI [.01, .04], β = .18, *p* = .01. We did not find a significant cross-lagged relationship between the group-level implicit bias in the previous month and the group-level explicit bias in the following month, b = .77, *SE* = .55, 95% CI [–.31, 1.85], β = .02, *p* = .16.