Northwestern | INSTITUTE FOR POLICY RESEARCH

RESEARCH EXCELLENCE • POLICY IMPACT

IPR Working Paper Series

WP-26-01

Patterns of Implicit and Explicit Attitudes V: Increase in Bias from 2021–2024

Tessa Charlesworth
Northwestern University and IPR

Meriel Doyle
Northwestern University

Mahzarin Banaji
Harvard University

Version: January 15, 2026

DRAFT

Please do not quote or distribute without permission.

Institute for Policy Research • 2040 Sheridan Rd., Evanston, IL 60208 • 847.491.3395 • ipr@northwestern.edu

Abstract

Between 2007–2020, implicit and explicit intergroup attitudes declined in bias steadily and were forecasted to continue toward attitude neutrality. New data from 2.5 million U.S. respondents (2021–2024) reveal that these encouraging trends have stalled or reversed. The largest increases in bias emerged for sexuality, transgender, race and skin-tone bias; 10–108% increases on explicit and 6–13% increases on implicit measures. Age, disability, and body weight bias also increased, but at slower rates. Exploratory breakpoint analyses showed that implicit attitudes were leading indicators of change, reversing trend earlier than explicit reports. Reversals were widespread across demographic groups for most topics, though strongest among conservatives for sexuality and transgender biases. Surprisingly, younger respondents (who had previously shown the largest decreases in bias) now showed greater increases in bias. Even after robust bias reduction spanning over 14 years, the new observed bias increases since 2021 highlight how minds get reshaped by sweeping sociocultural change.

Statement of Relevance

Recently, social commentators have warned that socio-political conditions in the U.S. are re-emboldening expressions of prejudice and discrimination. Consistent with these concerns, analyses of implicit and explicit measures of bias (on sexuality, transgender, race, skin-tone, age, disability, and body weight) from 2.5 million U.S. respondents (collected 2021-2024) reveal that long-standing declines in bias have stalled or reversed. All seven attitudes showed increasing bias, with the largest increases in race, skin-tone, sexuality, and transgender attitudes, the latter two especially among self-identifying conservatives. Surprisingly, younger respondents—previously the group showing the largest bias reductions—now showed the largest bias increases, suggesting that they are harbingers of broader societal shifts. After ruling out common alternative explanations, we suggest that this unpredicted backsliding may be explained by co-occurring societal events and elite rhetoric that activated and normalized intergroup hostilities, underscoring how macro-level trends can stall or reverse egalitarian progress in millions of individual minds.

Research Transparency Statement

General Disclosures

Conflicts of interest: All authors declare no conflicts of interest. **Funding:** No funding supported this research. **Artificial intelligence:** No artificial intelligence technologies were used in this research. **Ethics:** This research was determined to be exempt from ethics approval as it was minimum risk and constituted secondary data analysis of archival data.

Study One Disclosures

Preregistration: The associated pre-registration for data preparation and analyses can be found here: https://aspredicted.org/3sdx-8p7m.pdf (for primary analyses) and https://aspredicted.org/6m8w-9y62.pdf (for secondary analyses). **Materials:** All materials reported in this article are publicly available at the Open Science Framework: https://osf.io/3r76p/?view_only=d0a034e69d2d489699c7f9dcd4190e50. **Data:** All data are publicly available at: https://osf.io/3r76p/?view_only=d0a034e69d2d489699c7f9dcd4190e50. **Analysis scripts:** All analysis scripts are publicly available at: https://osf.io/3r76p/?view_only=d0a034e69d2d489699c7f9dcd4190e50.

Patterns of Implicit and Explicit Attitudes V. Increase in bias from 2021-2024

For more than a decade, shifts in U.S. intergroup attitudes told an optimistic story. From 2007 to 2020, explicit, consciously endorsed attitudes on six topics (age, disability, body weight, race, skin-tone, and sexuality) decreased sharply towards reduced bias (Charlesworth & Banaji, 2022b). Indeed, for the first time since data collection began, race attitudes had reached neutrality in 2020. Perhaps even more remarkable, *implicit* attitudes—more automatic and less controlled attitudes that were long thought to be resistant to change (Bargh, 1999; Greenwald & Banaji, 1995)—were also heading towards neutrality: data showed a 64% decrease in implicit anti-gay bias and 25% drop in anti-Black bias (Charlesworth & Banaji, 2022b). Although such decreases were widespread across most demographic groups, bias reduction was strongest among self-reported liberals and younger respondents. Decreasing explicit biases were also seen across this period in representative surveys (Gallup, 2025; Pew Research Center, 2019) underscoring that U.S. society appeared to be on a durable path towards greater egalitarianism.

Yet those same data (from Charlesworth & Banaji, 2022b) also provided the first evidence that change need not unidirectionally head towards neutrality; attitude trends can, at least temporarily, alter course. Around 2016, race, skin-tone, disability, and weight bias reversed direction and increased for approximately one year before returning to their decreasing trends. These reversals provided discriminant and external validation in that reversals occurred only for those topics affected by specific sociopolitical events such as the Republican national convention in 2016, or the shaming of a disabled journalist. But again, these specific perturbations were only temporary, and statistical forecasting models suggested that past trends were robust enough to predict that long-term reductions in bias should continue past 2020.

The backdrop of 2021-2024: A fraying social fabric

Despite such predictions of continued reductions in bias, it is also clear that, starting around 2020 (albeit with earlier roots) a combination of sociopolitical events in the United States have been creating a new zeitgeist of intergroup hostility. For instance, national polling and surveys document that, over this period, the U.S. experienced peak levels of sectarianism and ideological polarization (Brenan, 2025; Finkel et al., 2020), health threats from the global Covid-19 pandemic (Mathieu et al., 2020), economic insecurity and inequality (Desilver, 2024), and social media toxicity (Blumer & Kleinberg, 2025).

Empirical evidence suggests that these types of events can, in turn, shape intergroup attitudes and behaviors. For instance, political polarization leads political elites (from both sides) to use identity politics to reinforce party lines, pushing intergroup attitudes towards extremes (Iyengar, 2025; Westwood & Peterson, 2020). Perceived health threats from the pandemic have been shown to increase outgroup avoidance (Karwowski et al., 2020), existential threats of economic scarcity increase outgroup scapegoating (Bursztyn et al., 2022; Sidanius & Pratto, 1999), and social media algorithms normalize, even reinforce, toxic outgroup language (Brady et al., 2023; Mamakos et al., 2025).

Empirical and theoretical evidence of continued vs. altered attitude trends.

It is against this backdrop of co-occurring sociopolitical events that we test trends of implicit and explicit group attitudes between 2021-2024. Empirical and theoretical perspectives suggest two possibilities. First, in line with model predictions from past data (2007-2020), all explicit attitudes as well as implicit sexuality, race, and skin-tone attitudes should continue to show reduction in bias (while implicit age, disability, and weight bias should continue to remain stable). This aligns with multiple investigations showing that past perturbations of decreasing bias are short-lived (Charlesworth & Banaji, 2022b; Morehouse et al., 2025) and even that some events (e.g., Obama election) do not affect attitudes (Schmidt & Axt, 2016). Indeed, attitudes, especially implicit attitudes, are argued to reflect enduring social systems that are extremely hard and slow to change, often taking multiple decades, if not centuries, to alter (Charlesworth & Hatzenbuehler, 2025). Because of these enduring systems, collective attitudes will continue to remain stable (e.g., age, disability) or, if they had begun to reveal change (e.g., race, sexuality), then those changes must have happened because the social systems themselves changed to create a new, enduring environment. Thus, past trends, whether showing stability or change, would, under most circumstances be expected to show long-run persistence.

The Bias of Crowds model, by emphasizing that attitudes are deeply tuned to norms of the environment, suggests an alternative: in the rare cases that an environment can be changed rapidly and profoundly, such as through new legislation, policy, elections, or elite rhetoric then attitudes would be expected to also update rapidly and profoundly. Indeed, evidence has shown that strong signals of new norms, such as state and federal legalization of same-sex marriage, can prompt quick inflection points in attitude trends (Ofosu et al., 2019; Tankard & Paluck, 2017). Importantly, this new norm must be strong and repeated in society to switch an attitude trajectory: for instance, same-sex marriage legislation coincided with grassroots movements, lesbian and gay media representation, and changing demographics (Kumar et al., 2023). Thus, while a single event of one election might get washed away (e.g., Charlesworth & Banaji, 2022b), the combination of multiple, simultaneous events may be sufficient to signal a new ecosystem and alter attitudes, on both implicit and explicit measures.

To test these alternatives, we bring new evidence on the implicit and explicit attitudes from 2.5 million U.S. respondents towards seven social group topics: sexuality, race, skin-tone, age, disability, body weight and, newly also transgender identity – a central topic in today's political discussion (Lewis et al., 2022).

The current data offer several advantages: (1) continuity (starting reporting of data from 2007); (2) the necessary temporal granularity to identify temporary within-year perturbations versus more enduring trend reversals; (3) massive scale across dozens of demographic groups and all U.S. states (identifying the generalizability or potential sources of any trend alteration); (4) comparisons across implicit and explicit measures (argued to change via different mechanisms both individually and societally; Charlesworth & Banaji, 2022b; Gawronski & Bodenhausen, 2006); and (5) comparisons across topics that were previously changing (sexuality, race, skin-tone) versus stable (age, disability), revealing how past trajectories shape new trends.

Method

Open Practices Statement.

All data and analysis scripts are available from the Open Science Framework: https://osf.io/3r76p/?view_only=d0a034e69d2d489699c7f9dcd4190e50. In addition, analyses for this project were pre-registered through two pre-registrations. First, we preregistered primary data analyses for the six original tests at https://aspredicted.org/3sdx-8p7m.pdf. Second, after observing these initial results, we also preregistered (https://aspredicted.org/6m8w-9y62.pdf) analyses to inspect trends across a new test—the Transgender IAT, which has been collecting data since January 2020—as well as to compare results with covariates of public interest (i.e., Google searches, Republican voting). When necessary, we conducted additional exploratory analyses that were previously unconsidered and note any deviations from pre-registrations below. All data and analytic strategies also follow those introduced and validated in prior work (Charlesworth & Banaji, 2019, 2021, 2022b); this provides another layer of pre-registration since we are directly testing any new data and trends against already-published model forecasts.

Data Source.

We used the open data from the Project Implicit (PI) demonstration website at https://osf.io/px8h3/. The PI demonstration website hosts Implicit Association Tests (IATs) and a variety of explicit measures of attitudes and beliefs about various social group topics including the seven key topics we include here (sexuality, race, skin-tone, age, disability, body weight, and transgender identity). All participants are volunteers who arrive at the PI demonstration website (see SM for additional details about sample sources).

Pre-planned analyses of these data are conducted every 4-5 years. For this cycle, the full sample of 9.5 million respondents spans January 1, 2007, to December 31, 2024, with new data (not appearing in any previous publication) of 2.5 million respondents spanning four years from January 1, 2021, to December 31, 2024. Participants were excluded if: (1) their IAT D scores fell outside of the conditions in the revised scoring algorithm (Greenwald et al., 2003); (2) they were not residents of the United States; or (3) they did not have complete explicit measures or demographic information on age, gender, race, political ideology, and education. On average, we retained 85% of complete U.S. sessions across tests, which is on par with previous publications using these data (e.g., Charlesworth & Banaji, 2022b; SM provides test-specific retentions).

Sample Demographics.

Table 1 reports the combined sample demographics for all data from the ~9.5 million respondents from 2007-2024; demographics for just the new data from 2021-2024 are reported in the SM (Table S3.3). Overall, across all tests and all 18 years, the modal sample was female-identifying (67% of sample), White (71%), college-educated (81%), liberal (61%), and Christian (54%), consistent with previous years of this data. Although we found some changes in the sample demographics from 2007-2024 (e.g., the sample generally became older and less White; see Table S6 in the SM) we controlled for any such sample changes through raking and weighting approaches described below. This ensures a comparable timeseries because each monthly attitude estimate comes from the same (weighted) demographic composition over all 18 years of data.

Table 1
Sample Demographics Across All Attitudes in 2007-2024 Data

Attitude	N	Mean Age	SD Age	% Female	% Male	% White	% Black	% Asian	% > College	% Liberal	% Neutral	% Conserv.	% Christian	% Other Rel.	% Non-Rel.
Age	1189624	29.60	13.69	69.19	29.96	72.97	8.86	6.43	85.40	38.23	35.72	26.06	59.35	10.21	28.31
Disability	559586	31.32	13.46	73.39	24.76	76.09	8.07	5.24	88.48	46.67	29.93	23.41	56.65	9.84	30.95
Skin-tone	994068	30.41	13.10	69.62	29.24	62.40	16.19	6.27	87.26	51.29	29.44	19.27	54.76	11.27	31.73
Race	3875307	30.62	13.75	62.90	35.83	70.26	11.67	5.84	79.87	51.14	27.35	21.50	53.71	9.82	33.55
Sexuality	1438260	27.76	12.42	66.02	31.30	73.12	8.94	5.62	85.79	54.14	27.11	18.75	48.49	10.40	38.57
Weight	1233826	29.46	12.96	71.60	27.36	74.55	7.69	6.14	80.33	44.90	30.17	24.92	54.68	9.50	33.19
Transgender[a]	232053	33.24	13.80	63.93	26.73	74.64	8.48	4.38	83.38	60.87	20.48	18.65	44.77	9.43	43.34
Total	9522724	29.99	13.42	66.63	31.76	71.22	10.57	5.89	80.85	49.16	28.93	21.91	53.80	10.05	33.53

Note. [a.] Transgender task spanned 2020-2024; all other tasks spanned 2007-2024. See Tables S3.1, S3.2, and S3.3 in the supplemental materials for sample demographics across separated data sets (2007-2016 versus 2017-2020 versus 2020-2024). Sample demographic representations were generally consistent over time (see Tables S6 in SM for correlations of demographics with time).

Materials

Implicit Association Test (IAT). To assess implicit attitudes, we used the IAT (Greenwald et al., 1998), which is the most widely used measure of implicit cognition, with extensive psychometric validation (Kurdi & Banaji, 2021). Although debates continue about the psychometric properties of using the IAT as an *individual* difference measure, there is clear evidence that, at a *population* level (as we use the test here), the IAT is correlated moderately-to-strongly with meaningful outcomes, and consistently passes benchmarks for construct validity (Calanchini et al., 2022; Charlesworth & Banaji, 2022a; Gawronski, 2025; Hehman et al., 2019).

The IAT is a computerized reaction time task that compares participants' speed of categorizing image and word stimuli together in "congruent" blocks (e.g., Thin=good/Fat=bad) against the speed of categorizing image and word stimuli in "incongruent" blocks (e.g., Fat=good/Thin=bad). The psychological assumption behind the test is that faster sorting in the congruent blocks indicates that the congruent pairs (e.g., Thin=good/Fat=bad) are more strongly and commonly associated in a participants' mind and culture. IAT D scores are calculated by taking the difference in average speed of categorizing in congruent versus incongruent blocks and then dividing (normalizing) by the pooled standard deviation across all trials. All image and word stimuli used in the IATs are available from the archived Project Implicit website data at https://osf.io/px8h3/.

Explicit preference. To assess explicit attitudes, we used a 7-point scale ranging from -3 reflecting strong counter-cultural preferences, such as "I strongly prefer Fat people to Thin people" to +3, reflecting strong normative-cultural preferences, such as "I strongly prefer Thin people to Fat people". The neutral point (0) indicates that participants self-report having "no preference" e.g., "I like Thin people and Fat people equally." Although other explicit questions are available on Project Implicit (i.e., feeling thermometer measures), we chose this relative explicit question to mirror the relative associations captured in the IAT.

Analytic Strategy.

Primary Analyses: Overall (2007-2024) and new (2021-2024) trends. Timeseries data, as we have here, are characterized by temporal autocorrelations (i.e., measures close in time are closely related), nonlinear trends, and even sub-yearly seasonal trends (e.g., systematic rises in bias in winter). For these reasons, common modeling approaches, such as multiple regressions, are inappropriate statistical methods. Instead, following past work (Charlesworth & Banaji, 2019, 2022b), we rely on a timeseries model class called ARIMA models, or Autoregressive Integrated Moving Average models (Cryer & Chan, 2008; Hyndman & Athanasopoulos, 2013). All ARIMA models were estimated using an automated approach through the *auto.arima()* function implemented in the *forecast* package (version 8.24.0) in the R computing environment (version 4.4.1).

Briefly, ARIMA models work by separating the underlying timeseries process into specific seasonal (i.e., repeated yearly patterns) and non-seasonal components, as well as a possible slope (or "drift") capturing any constant linear trend in the data. For each seasonal/non-seasonal component, ARIMA models first ensure that the timeseries process is stationary (i.e., the process itself is not changing or drifting over time) by differencing the series. Next, the ARIMA model tests whether the (differenced) process is best captured by moving average processes (i.e., experiencing "shocks" in the system) and/or by autoregressive processes (i.e., reliable lagged relationships). This results in an ARIMA model with up to seven parameters $(p,d,q)(P,D,Q)$+drift, capturing the autoregressive (p), differencing (d) and moving average (q) parameters first for the non-seasonal, then the seasonal process, and finally any slope.

The number of ARIMA parameters indicate how *many* of these autoregressive, differencing or moving average parameters are needed to fully capture the observed timeseries process. However, the exact order and values of these parameters are typically not the main interpretation. Instead, ARIMA models are more commonly interpreted by focusing on: (1) the forecasts of future trends, based on the ARIMA models description of past trends; and (2) summaries of both the raw and percent change from decomposed trends after removing the seasonal and noise components of the timeseries process.

Comparing 2021-2024 trends to ARIMA forecasts. A key advantage of ARIMA models is that they offer forecasts of how future attitude trends may unfold. Because this is the fifth paper in a series with continuous monitoring of attitude trends since 2007, we are able to offer a comparison between forecasted and observed data. In fact, exactly this approach was taken in a previous paper (Charlesworth & Banaji, 2022b), in which we showed that observed data from 2016-2020 had robustly and accurately followed the forecasts based on 2007-2015 data.

As before, we rely on three standard forecast accuracy metrics (Hyndman & Athanasopoulos, 2013): (1) the mean average error (MAE); (2) the root mean squared error (RMSE); and (3) the mean average scaled error (MASE). MAE reports the actual difference between the observed and forecasted mean or $mean(|e|)$, but it is scale-dependent, meaning that it must be interpreted relative to the actual values (i.e., the range of IAT D scores or explicit preference scores). RMSE, also scale-dependent, is calculated as $\sqrt{mean(|e^2|)}$ and reflects the square root of the amount that the mean of the forecast was "off" from the observed mean. For both MAE and RMSE, larger values indicate more error in the forecast.

Finally, MASE is the only non-scale-dependent metric that can therefore be more easily compared across tasks and across implicit/explicit attitude outcomes. MASE compares against a standard naïve forecast (i.e., from a simple prediction of y_t from y_{t-1}) and calculates relative accuracy of the actual forecasts using the full ARIMA model. MASE values greater than 1 indicate that errors from the ARIMA forecast are larger than the errors from the naïve forecast (i.e., the naïve forecast would have been just as efficient). We had also pre-registered assessing accuracy with the Mean Absolute Percentage Error (MAPE). However, after further investigation, we noted that this was not appropriate for our setting because it is often unreliable when the data includes small or zero values, which characterizes many of the timeseries in the current work since the attitudes are near neutrality.

Robustness analyses: Sample demographic change. One concern when using cross-sectional data to infer change is that the observed change could be driven not by true attitude change but by changes in the underlying population of respondents (see Table S6 in SM). For example, if the average for a given attitude in January 2007 is calculated from a sample that is 50% women, but the average in December 2024 is calculated from a sample that is 70% women, we cannot directly compare the two averages. To resolve this concern, we use a raking and weighting approach implemented through the *anesrake* package (version 0.80) implemented in R. Target "weights" are set as the average demographic representation across the entire 18-year timespan (Table 1) and then each year is re-weighted to match those targets.

Specifically, we use raking and weighting across the marginal population representations of five key demographics: gender, race, education, age, and politics. So, for example, we match the proportion of respondents who are women, White, college-educated, young, and liberal (and every other factor combination) for all years and then calculate the weighted monthly mean. As noted above, this ensures that each mean is calculated from the same (weighted) composition of respondents. Notably, as we report in the open code on the project's OSF page, all trends and conclusions are consistent across both weighted and unweighted trajectories, underscoring that sample change alone is unlikely to account for the observed changes.

Robustness analyses: Differences in change by generational cohorts. Another potential concern for interpreting cross-sectional timeseries data is that the underlying process of change could be due not to attitudinal change but, instead, to cohort replacement, whereby older cohorts (e.g., Silent generation, Baby boomers) are replaced gradually by younger cohorts (e.g., Millennials, Gen Z). Change that is entirely attributable to cohort replacement would be seen as no within-cohort change. Instead, we would observe a main effect of different overall bias between cohorts (e.g., Millennials/Gen Z may have lower bias than Baby boomers) with Millennials gradually forming a larger portion of the sample, but without any within-cohort shift.

Alternatively, a "period effect" is a change process that reflects more genuine attitude change, whereby every cohort shows similar, parallel change over time (i.e., Millennials and Baby boomers both drop in bias by similar amounts), thus reflecting a widespread shift in society. Finally, there is the possibility of interactions between cohort and period sources of change. For instance, younger cohorts may show within-cohort change that is faster than the within-cohort change of older cohorts (e.g., Millennials show large drops in bias, Baby boomers show mostly steady bias). This indicates that, although there is a cultural period effect that is genuinely changing the minds of some people, the strength of this cultural effect may first be filtered through one's cohort. Indeed, in both U.S. and international data collected before 2020, we found that the most common source of attitude change is often cohort-by-period interactions, with younger cohorts typically reflecting stronger period effects than older cohorts (Charlesworth & Banaji, 2019, 2022b; Kurdi et al., 2025).

Note that age and cohorts are linearly related to one another (i.e., one's age is the year of test-taking minus one's birth year cohort), and thus we chose to focus our primary reporting below on participants' age at test-taking because age (unlike cohorts) doesn't require relatively arbitrary cut-offs based on birth years (S. M. Campbell et al., 2017). Nevertheless, in the SM, we also report results from cohort differences by dividing the sample according to participants' birth years, using four cohort cut-offs: Baby boomers/Silent Generation (combined for sample size;

born before 1964); Generation X (1965-1980); Millennials (1981-1996); Generation Z/Alpha (combined for sample size; born after 1997). Conclusions are consistent across both age-based analyses and cohort-based analyses.

Robustness analyses: Repeat test-takers, and sample sources. A final threat to inferences from cross-sectional data is the possibility that any observed change may be driven simply by practice effects or regression to the mean from participants who repeatedly return to the Project Implicit website. To address this concern, we recompute all trends with only those novice test-takers who report never having taken the IAT before. Although there are main effects (novice test-takers show larger biases overall), as in previous reports (Charlesworth & Banaji, 2019, 2021, 2022b), the conclusions from analyses of attitude change are consistent across both novice and repeat test-takers (see SM).

Furthermore, participants may come to the Project Implicit website for different reasons over time. As Project Implicit became more mainstream in education and workforce programs, a greater proportion of participants may arrive through such education programs rather than through individually-motivated interest or word of mouth. As with demographic population change, any change in the source of participants could hinder comparability of trends across time. We address this concern by computing separate trends for participants coming to the website due to assignments for work/school versus from word of mouth. Again, although main effects emerge (assignment participants showed higher biases overall than did word-of-mouth participants), the conclusions regarding the rate and direction of change remained robust across both sources of participants (see SM).

Secondary Analyses: Breakpoint detection.

To pinpoint *when* any trend reversal happens (if at all), we performed exploratory breakpoint detection. Breakpoint detection is used to estimate possible inflection points in otherwise continuous trends. The approach determines the optimal number of segments to describe a timeseries trend, with each segment reflecting a different slope of change. Thus, a breakpoint is detected when the timeseries trend changes substantially enough (e.g., flips direction or rapidly increases) to require an entirely different slope estimate. We perform breakpoint detection using an automated approach implemented in the *breakpoints()* function in the *strucchange* package (version 1.5.4 in R). We identified breakpoints for the overall series of implicit and explicit attitudes (from 2007-2024), as well as for subsets of liberal and conservatives' trajectories, to detect when (if ever) these sub-groups altered attitude trajectories.

Secondary Analyses: Demographic differences.

For demographic differences, we compare trajectories across all groups with sufficient data across tasks and years. Specifically, we compared six demographic categories: (1) politics (liberal, conservative); (2) age (< 20-year-olds, >40-year-olds); (3) gender (men, women); (4) race (White, Black, Asian); (5) education (no college, college); and (6) religion (Christian, Jewish, Non-religious, Other religion). We compared the trends and ARIMA models/forecasts across all pairs of subgroups to identify whether groups were moving in parallel, converging, or diverging in their trends (see full methodological development in Charlesworth & Banaji, 2021).

Notably, we took one more step for pre-processing this demographic data because it is theoretically possible that the samples from individual demographic groups (e.g., women, men) could change in different ways over time (e.g., women could become more liberal, men could become more conservative). Thus, to account for the possible demographic differences in sample changes, we performed a second weighting approach, different from our overall weighting approach described above. Specifically, we first subset the data along the key demographic differences (e.g., women, men) and then reweight each of these subsets to match the same target weights from the overall population. For instance, imagine that the full sample is 49% liberal, but the sample of women is 60% liberal in 2007 and 65% liberal in 2024, while the sample of men is 38% liberal in 2007 and 33% liberal in 2024. To create an accurate and comparable interpretation of this hypothetical data, we would construct weights so that liberals were consistently down-weighted in women's data and up-weighted in men's data. This approach makes the two gender samples more directly comparable across time and across all other demographic groups (because they are reweighted to have similar political orientation, race, education, and so on), thus allowing us to focus more directly on only the demographic difference of interest (e.g., gender).

Secondary Analyses: Geographic differences and correlations with covariates.

Modeling geographic differences across U.S. states is important not only to visualize the variability of trajectories but also to link data across multiple attitude tasks and covariates. As preregistered, and as done in previous analyses (Charlesworth & Banaji, 2021), we first subset the data according to U.S. state and then computed the weighted mean (using the target weights of the full sample population) for each state in each year. We used preregistered cut-offs of at least 50 observations per year (i.e., any state had to have at least 50 observations in a year to be included as a mean in the final data). Additionally, any trajectory that was missing more than 25% of its data (due to not passing the cut-offs) was excluded from analysis.

As pre-registered, we also recomputed trajectories for liberals and conservatives in each state separately to investigate the interaction between participant demographics and place. To ensure that the focus of the paper remained on understanding and explaining recent changes in attitude trends, we chose to perform geographic analyses only on the largest tasks that also have shown past changes – namely, sexuality and race attitudes – as well as the new task that has yet to be extensively reported or investigated – namely, attitudes towards transgender people. We also preregistered computing state-monthly trajectories and trajectories for metropolitan statistical areas (MSAs). However, after investigation, these trajectories were too data-sparse (and did not pass pre-registration thresholds). This would introduce noisy and even potentially inaccurate conclusions and, as such, we chose to focus on the larger samples from state-year trajectories.

In addition to computing U.S. state variation in attitudes, we also pre-registered secondary correlational analyses with: (1) voting patterns (votes for Trump vs. Biden in the 2020 federal election); (2) Google searches for the moral panic related to "sexual grooming" (Cohen & Galloway, 2025), downloaded from Google Trends, using "sexual grooming" as a topic search and limiting to searches from January 2007 – December 2024; and (3) cross-attitude relationships between transgender and sexuality attitudes, as we have previously theorized a potential spillover between these attitude topics (Charlesworth & Hatzenbuehler, 2025)

First, we calculated simple Pearson's correlations between voting, Google Searches, and attitudes within the same state/time. Second, we tested possible lagged relationships using Granger causality analyses with the *grangertest()* function in the *lmtest* package (version 0.9.40) in R. Granger models test whether the monthly rises/falls in a given timeseries (e.g., sexuality attitudes) are better explained when we include a lagged second timeseries (e.g., monthly rises/falls in, for example, Google searches for the previous month). We compare the effect sizes and significance of Granger causality in a forward direction (e.g., Google searches preceded sexuality attitudes) versus a backward direction (e.g., sexuality attitudes preceded Google searches) at lags of 1-6 months. If only one direction (e.g., forward) is significant, then this is interpreted as evidence for Granger causality. If both directions are significant then that suggests a third variable likely causes rises/falls in both series. And if neither direction is significant then that suggests the two variables are not temporally related to each other. We also explored additional approaches for modeling cross-lagged relationships (e.g., panel Granger causality) as reported in the SM; results were generally consistent across all approaches.

Results

Results address three primary questions. First, what are the overall biases and the overall rates of change from 2007 (the first year for which analyzable data are available) to 2024? Second, how do these trends vary when looking only at recent trends (2021-2024)? That is, have recent trends followed the predicted forecasts (of change or stability), or have they deviated from past trajectories? If so, when do such deviations start? Third, how do these trends vary across people and place? And does such variation correlate across topics (i.e., sexuality and transgender attitudes) and with other covariates of Republican voting and anti-gay Google searches?

RQ1: What are the overall biases and overall rates of change since 2007?

Overall biases. Across all time, every explicit and implicit attitude showed significant preferences for the culturally dominant group (abled, thin, young, white, light-skinned, straight, and cisgender), with the strongest biases emerging for disability, weight, and age attitudes, and the weakest biases for sexuality, transgender, race, and skin-tone attitudes (Table 2).

Table 2.
Means and Correlations of Six Implicit and Explicit Social Group Attitudes (2007-2024)

	Implicit (IAT D score)				Explicit (7-point scale)				Implicit-Explicit	
Attitude	**Mean**	**SD**	**95% CI**	**d**	**Mean**	**SD**	**95% CI**	**d**	**r**	**95% CI**
Sexuality	0.20	0.50	0.19,0.20	0.39	0.29	1.26	0.28,0.29	0.23	0.45	0.45,0.46
Race	0.27	0.44	0.27,0.27	0.62	0.12	1.03	0.12,0.12	0.12	0.31	0.31,0.31
Skin-tone	0.29	0.43	0.29,0.29	0.68	0.16	0.91	0.15,0.16	0.17	0.23	0.22,0.23
Age	0.44	0.39	0.44,0.44	1.12	0.38	1.17	0.37,0.38	0.32	0.12	0.11,0.12
Disability	0.51	0.45	0.51,0.51	1.13	0.36	0.88	0.36,0.36	0.41	0.14	0.13,0.14

	Implicit (IAT D score)				Explicit (7-point scale)				Implicit-Explicit	
Attitude	**Mean**	**SD**	**95% CI**	**d**	**Mean**	**SD**	**95% CI**	**d**	**r**	**95% CI**
Weight	0.47	0.41	0.47,0.47	1.14	0.76	1.06	0.76,0.77	0.72	0.20	0.20,0.20
Transgender[a]	0.12	0.45	0.11,0.12	0.26	0.44	1.17	0.44,0.45	0.38	0.35	0.34,0.35

Note. [a.] Transgender task spanned 2020-2024, all other tasks spanned 2007-2024. All means and correlations significantly different from zero, $p < .001$

Overall change (2007-2024). Turning to overall change across the 18-year timespan (2007-2024), we found that sexuality, race, and skin-tone attitudes have consistently decreased in bias (Table 3; Figure 1 and Figure 4). Explicit attitudes decreased by 86% (sexuality), 102% (race), and 66% (skin-tone), while implicit attitudes decreased by slower but still notable magnitudes of 61% (sexuality), 24% (race), and 12% (skin-tone). In contrast, body-related topics of age, disability, and weight attitudes changed at less than half the rate on explicit measures (dropping by 33%, 42%, and 33%, respectively) and implicit age and disability attitudes hardly changed (1% and 8% increases, respectively), with implicit weight attitudes even increasing by 13% since 2007.

At the most general level, these new data replicate reports from previous papers (Charlesworth & Banaji, 2019, 2022b) showing a distinction between faster change on more sociodemographic topics of sexuality, race, skin-tone attitudes versus slower change on body-related topics. And yet, we also foreshadow new observed trends by noting that these overall rates of change over 18 years (2007-2024) are generally *smaller* than the overall rates of change previously reported over just 14 years (2007-2020). Implicit sexuality attitudes, for example, dropped by 61% over all 18 years, but previous reports showed they already dropped by 65% over an earlier 14-year span. Already, this suggests that the newest attitude trends are not following the rates of forecasted change.

Table 3.
Overall Patterns of Change Across All Data (2007 – 2024) and future ARIMA Forecasts

Attitude	*Start*	*End*	*Raw* Δ	%Δ	ARIMA	Forecast Lower 95% CI	Forecast Mean	Forecast Upper 95% CI
				Implicit Attitudes (IAT D scores)				
Sexuality	0.31	0.12	-0.19	-61%	(0,1,2)(2,0,0)+drift	4 yrs, 4 mos (n)	14 yrs, 8 mos (n)	51 yrs, 11 mos (n)
Race	0.33	0.25	-0.08	-24%	(1,1,1)(2,0,0)	35 yrs, 3 mos (n)	>200 yrs (n)	64 yrs, 11 mos (d)
Skin-tone[a]	0.32	0.28[a]	-0.04	-12%	(2,1,1)	107 yrs, 4 mos (n)	>200 yrs (n)	101 yrs, 1 mos (d)
Age	0.44	0.44	+0.01	+1%	(0,1,2)(2,0,0)	>200 yrs (n)	>200 yrs (n)	147 yrs, 6 mos (d)
Disability	0.49	0.52	+0.04	+8%	(2,1,3)	143 yrs, 2 mos (n)	>200 yrs (n)	113 yrs, 10 mos (d)
Body Weight[b]	0.42[b]	0.48	+0.06	+13%	(0,1,1)(2,0,0)	36 yrs, 3 mos (n)	>200 yrs (n)	16 yrs, 4 mos (d)
Transgender[c]	0.11[c]	0.13	+0.02	+20%	(0,1,3)(1,0,0)	138 yrs, 5 mos (n)	>200 yrs (n)	110 yrs, 3 mos (d)
				Explicit Attitudes (7-point scale)				
Sexuality	0.61	0.08	-0.53	-86%	(0,1,2)(2,0,0)	Already neutral	>200 yrs (n)	26 yrs, 8 mos (d)
Race	0.31	-0.01	-0.32	-102%	(0,1,2)(1,0,1)+drift	Already neutral	Already neutral	29 yrs, 4 mos (n)
Skin-tone[a]	0.27	0.09[a]	-0.18	-66%	(0,1,1)(1,0,1)	2 yrs, 2 mos (n)	>200 yrs (n)	16 yrs, 9 mos (d)
Age	0.53	0.36	-0.17	-33%	(1,0,1)(0,1,1)	11 yrs, 6 mos (n)	>200 yrs (n)	47 yrs, 2 mos (d)
Disability	0.56	0.32	-0.24	-42%	(3,1,1)(1,0,0)+drift	11 yrs, 7 mos (n)	18 yrs (n)	25 yrs, 4 mos (n)
Body Weight[b]	0.96[b]	0.65	-0.31	-33%	(0,1,2)(2,0,0)	19 yrs, 5 mos (n)	>200 yrs (n)	11 yrs, 9 mos (d)
Transgender[c]	0.43[c]	0.47	+0.04	+8%	(0,1,1)	34 yrs, 6 mos (n)	>200 yrs (n)	18 yrs, 1 mo (d)

[a] Assessment of skin-tone attitudes changed in 2023 from a general test (i.e., about light versus dark skin, in general) to four specific tasks that crossed race and skin tone (e.g., light versus dark skin White faces), as we describe in detail in the SM. However, some data were still collected with the previous task; we use the data from the consistent task for all analyses, although inferences are offered more cautiously due to the large changes in sample size.
[b] The IAT stimuli for implicit body weight attitudes changed in April 2010 from face images to body silhouettes; to facilitate inferences, however the data around 2007 showed similar average bias. As such, for simplicity, we chose to report the data continuously (combining across the two tasks) from 2007-2024.
[c] The Transgender task was introduced in January 2020. As such, the starting values and overall change values are not directly comparable to the other tasks.

Note. Start and end values, as well as percent change (%Δ) and raw change (Δ), are calculated from the start and end points of the decomposed trend line (removing seasonality and noise). Raw change is calculated as the difference between start – end values; percent change is the raw change, divided by the starting value (start-end / start). As such, even given the same amount of raw change (e.g., -0.18 and -0.17 in explicit skin-tone and age attitudes), percent changes will differ because of different attitude start points, such that smaller starting values (smaller denominators) will correspond to larger percent changes. For ARIMA: The first three parameters of the autoregressive-integrated-moving-average (ARIMA) model are nonseasonal, the second three values are seasonal, and, in some cases, a drift parameter (slope) is included. In each set of parameters, p specifies the number of autoregressive parameters used to explain the autocorrelations in the data, d specifies the number of differencing parameters necessary to make the series stationary, and q specifies the number of moving-average parameters used to explain the lagged forecast errors. Forecasts indicate the number of months and years it could take to reach attitude neutrality (n) or to double in magnitude (d) from January 2025, if past trends were to continue.

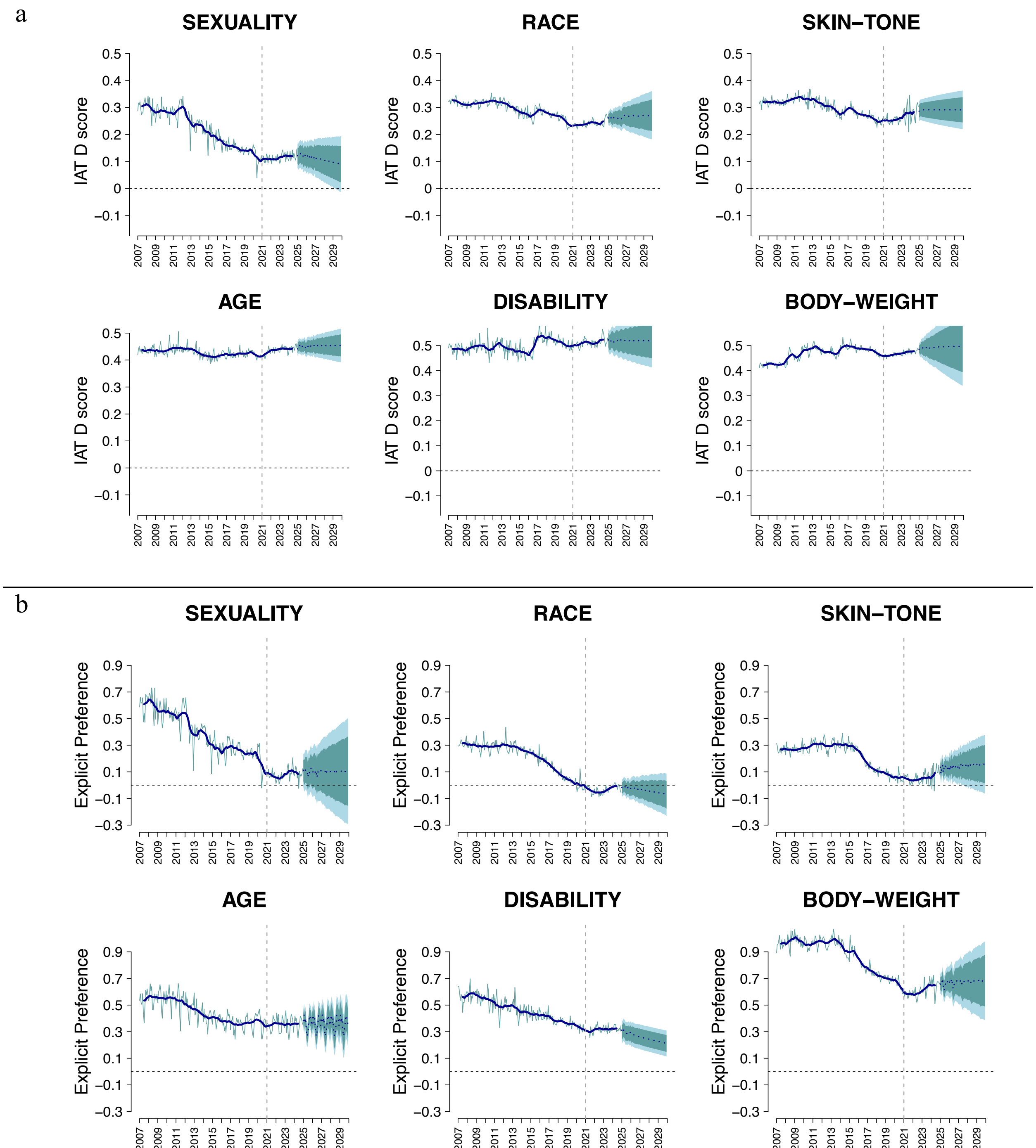


Figure 1. *Implicit (a) and explicit (b) attitude trends and forecasts over all time (2007-2024).* Vertical dashed gray line indicates the onset of the new, 2020-2024 data. The overlayed thick dark blue lines indicate the decomposed trend (removing seasonality and noise) for the raw observed data (thin light blue lines). Dark shaded areas indicate 80% forecast confidence interval (CI), light shaded areas indicate 95% forecast CI. Implicit body weight attitudes included two tests (differentiated by stimuli), with the early test (using stimuli of faces) plotted in a dashed line and the later test (using stimuli of silhouettes) plotted in a solid line. Note that all transgender attitudes are visualized in Figure 4 due to differences in the time scale.

RQ2: Have trends persisted or reversed from 2021-2024?

These overall trends mask a critical new result in the data collected since January 2021: *every single attitude shows some evidence of increased bias away from neutrality on both implicit and explicit measures* (Table 4; Figure 1, data after the dashed lines). The largest reversal emerged in anti-gay/pro-straight (sexuality) attitudes, which, based on model forecasts from past trends, should have continued decreasing by 44% for implicit attitudes and even reached neutrality for explicit attitudes. Instead, since 2021, anti-gay/pro-straight attitudes increased by 9% (implicit attitudes) and 10% (explicit attitudes). Similar reversals were found for both implicit and explicit anti-Black/pro-White (race) and anti-dark/pro-light skin (skin-tone) attitudes. Despite forecasts predicting continued drops towards neutrality, the observed data showed increases of 6% (implicit anti-Black/pro-White), 13% (implicit anti-dark/pro-light skin), 79% (explicit anti-Black/pro-White), and 108% (explicit anti-dark/pro-light skin).

Even anti-old/pro-young (age), anti-disabled/pro-abled (disability) and anti-fat/pro-thin (body weight) attitudes, which were forecasted to remain relatively stable, have shown increased biases recently, increasing by 3-11% over just four years (Table 4). Additionally, the new data from transgender attitudes showed that, over this same time period of 2021-2024, both implicit and explicit anti-transgender/pro-cisgender attitudes have increased by 13% and 16%, respectively, putting them on par with increases seen in all other social group attitudes.

These reversals produced large forecasting errors (e.g., MASE generally greater than 1), and most of the trends fell outside the 80% confidence intervals of past forecasts (Table 4). So, although most of the observed data (especially in early months) still fell within the most liberal 95% confidence intervals, it appears that the new trends are not closely following any of the past model predictions.

Breakpoint detection: When did trends reverse? Across the full timeseries (2007-2024), models for implicit attitudes consistently identified a final breakpoint that started in mid-2020 (see SM, Table S25). That is, five implicit attitudes (out of the six trends, not including transgender attitudes due to shorter timeseries) showed a breakpoint and reversal around April-August 2020. Only age attitudes showed an earlier breakpoint in December 2019. Thus, using a bottom-up automated approach across the entire timespan, it is clear that the period around mid-2020 generated a large-scale reversal in implicit attitudes, regardless of topic.

Interestingly, results suggested that implicit attitude breakpoints preceded the breakpoints for explicit attitudes: explicit race, skin-tone, disability, and body-weight attitudes reversed about a year after, generally starting to reverse in April-December 2021 (age attitudes again showed no breakpoint discovered in the last 4 years; Table S25). This ordering—where implicit precedes explicit attitude reversals—was also robust even among the individual series of liberals and conservatives alone (Table S26). At first glance, this ordering may be unexpected: explicit attitudes are often thought to be easier to change (Kurdi & Charlesworth, 2024) and therefore any societal shift would be expected to affect explicit attitudes first. However, these data align with the emerging perspective that implicit attitudes are more attuned to cultural phenomena (Payne et al., 2017) and therefore may be more sensitive and early detectors of social change.

The only attitude that did not fit the pattern of implicit attitude change leading explicit attitude change were anti-gay/pro-straight attitudes. Instead, these attitudes showed nearly simultaneous change on both implicit and explicit attitudes (reversing course within one month of each other). Perhaps, in cases of such extreme anti-gay legislative changes and public backlash that provide overwhelming evidence of a new norm, both implicit and explicit attitudes may reverse simultaneously.

Table 4.

Recent Trends from 2021-2024: Forecasted Versus Observed Trends

	Forecast Predictions (2021-2024)				***Observed Data (2021-2024)***				***Accuracy of Predictions***				
	Start	***End***	***Raw Δ***	***%Δ***	***Start***	***End***	***Raw Δ***	***%Δ***	***MAE***	***RMSE***	***MASE***	***%in 80%***	***%in 95%***
Implicit Attitudes (IAT D scores)													
Sexuality	0.10	0.06	-0.05	-44	**0.11**	**0.12**	**+0.01**	**+9**	0.04	0.04	1.25	63%	96%
Race	0.23	0.21	-0.02	-7	**0.23**	**0.25**	**+0.01**	**+6**	0.03	0.03	1.54	83%	100%
Skin-tone[a]	0.25	0.25	-<0.01	-<1	**0.25**	**0.28**	**+0.03**	**+13**	0.02	0.03	1.19	73%	94%
Age	0.41	0.41	-0.01	-1	**0.43**	**0.44**	**+0.02**	**+4**	0.03	0.03	2.08	33%	90%
Disability	0.50	0.50	-<0.01	-<1	**0.50**	**0.52**	**+0.02**	**+5**	0.02	0.02	0.54	98%	100%
Body weight	0.46	0.46	-<0.01	-<1	**0.46**	**0.48**	**+0.02**	**+4**	0.01	0.01	0.53	100%	100%
Transgender	*N/A: Not enough past data*				**0.12**	**0.13**	**+0.01**	**+13**	*N/A: Not enough past data to compare*				
Explicit Attitudes (7-point scale)													
Sexuality	0.12	-0.01	-0.13	->100	**0.08**	**0.08**	**+0.01**	**+10**	0.06	0.08	0.80	94%	100%
Race	-0.04	-0.12	-0.08	->100	**-0.04**	**-0.01**	**+0.03**	**+79**	0.05	0.07	1.34	63%	90%
Skin-tone[a]	0.04	0.02	-0.02	-44	**0.04**	**0.09**	**+0.05**	**+108**	0.05	0.06	1.17	83%	100%
Age	0.35	0.33	-0.01	-4	**0.35**	**0.36**	**+0.01**	**+3**	0.03	0.04	0.73	100%	100%
Disability	0.32	0.29	-0.03	-8	**0.30**	**0.32**	**+0.03**	**+9**	0.03	0.03	0.58	98%	100%
Body weight	0.60	0.56	-0.04	-11	**0.58**	**0.65**	**+0.07**	**+11**	0.06	0.07	0.87	92%	100%
Transgender	*N/A: Not enough past data*				**0.41**	**0.48**	**+0.07**	**+16**	*N/A: Not enough past data to compare*				

Note. Forecast predictions indicate what was expected to happen based on the ARIMA models from past trends from 2007-2020. Observed data indicate what actually happened in the data. See explanation of start, end, percent change, raw change in the note to Figure 3 above. As noted, percent change values are calculated from (start – end)/start, and thus smaller starting values will translate to larger proportional changes, even for the same amount of raw change. All explanations of accuracy statistics are provided in the main text. MAE = Mean average error; RMSE = Root mean squared error; MASE = mean average scaled error; %in 80% = the percentage of the 48 months (from 2021-2024) that were included in the 80% confidence interval of past forecasts; %in 95% = the same but for 95% confidence interval of past forecasts.

RQ3.1: How do trends vary by demographics?

Finally, to gain a deeper understanding of the spread and possible sources of these recent reversals we explore demographic and geographic differences. Such differences can tell us whether the sources of change are likely to come from events and rhetoric targeting specific groups/places or whether they reflect more general, macro-level effects. For simplicity, we maintain focus on the two demographic differences that have previously been shown to matter in long-term attitude trends (Charlesworth & Banaji, 2021, 2022b): political differences (liberals vs. conservatives) and age differences (younger <25-year-olds vs. older >45-year-olds), as well as explore their intersections (between politics, age, and gender). Differences by religion, education, race, and gender are summarized briefly below and reported in detail in the SM.

Demographic differences: Politics. Recent trends of sexuality and transgender attitudes showed large political differences. Specifically, the rise in anti-gay/pro-straight implicit attitudes was concentrated in self-identified conservative respondents (Figure 2), who increased by 17% since 2021, whereas self-identified liberal respondents remained stable around neutral attitudes (slightly pro-gay/anti-straight). Similar political differences were seen in implicit anti-transgender/pro-cisgender attitudes (Table 5).

Such political differences were even more notable on explicit attitudes, where conservative respondents increased in their reported anti-gay/pro-straight bias by 35% in just 4 years. Respondents who identified as liberal, however, continued to decrease towards stronger reported *pro-gay/*anti-straight attitudes by 20% over the same period. Similarly, conservative respondents increased in their reported anti-transgender/pro-cisgender attitudes by 31%, whereas liberals decreased by 52%, and forecasts from the ARIMA models for liberals' attitudes already included neutrality.

To be clear, previous data had also shown that conservatives and liberals differed in their rate of change in sexuality attitudes, with liberals often changing relatively faster. But, in that past data, conservatives had also been consistently (albeit more slowly) decreasing towards less bias on both implicit and explicit measures. This data is the first, to our knowledge, that shows such robust backsliding of conservatives' previous change towards tolerance.

This difference between liberal and conservatives' sexuality and transgender attitudes is striking because, for most other implicit and explicit attitudes in the recent data (2021-2024), both liberals and conservatives showed small and parallel reversals in preferences. For instance, on implicit race attitudes, self-identifying liberals increased in bias by 8% and self-identifying conservatives increased by 5%. Similarly, on explicit race attitudes, liberals increased in their reported anti-Black/pro-White bias by 25%, while conservatives increased in anti-Black/pro-White bias by a similar 32%. The only additional difference between liberal and conservative respondents emerged on explicit disability and body weight attitudes. For those two tests, conservatives increased in their reported anti-disabled/pro-abled attitudes (by 15%) and anti-fat/pro-thin attitudes (by 16%), whereas liberals did not change and held steady to their previous levels of bias.

Table 5.
Political Differences in Implicit and Explicit Attitude Trends From 2021-2024

		Implicit Attitudes (IAT D-score)					***Explicit Attitudes (7-point scale)***				
Attitude	***Group***	***Start***	***End***	***%Δ***	***Raw Δ***	***Interpretation (2007-2024)***	***Start***	***End***	***%Δ***	***Raw Δ***	***Interpretation (2007-2024)***
Sexuality	Conserv	0.33	0.39	+17.13	+0.06	Diverging	0.82	1.11	+35.20	+0.29	Diverging
	Liberals	-0.04	-0.04	+4.25	+0.00	**(C ↗ > L →)**	-0.38	-0.45	-20.02	-0.08	**(C ↗ > L ↘)**
Race	Conserv	0.32	0.33	+4.59	+0.01	Parallel	0.17	0.22	+32.27	+0.05	Parallel
	Liberals	0.19	0.21	+8.29	+0.02	**(C ↗ = L ↗)**	-0.14	-0.10	+25.37	+0.03	**(C ↗ = L ↗)**
Skin-tone	Conserv	0.32	0.34	+6.63	+0.02	Parallel	0.21	0.27	+26.10	+0.06	Parallel
	Liberals	0.22	0.24	+10.06	+0.02	**(C ↗ = L ↗)**	-0.02	0.00	+99.92	+0.02	**(C ↗ = L ↗)**
Age	Conserv	0.44	0.45	+2.97	+0.01	Parallel	0.27	0.29	+8.94	+0.02	Parallel
	Liberals	0.42	0.45	+6.09	+0.03	**(C ↗ = L ↗)**	0.53	0.54	+2.56	+0.01	**(C ↗ = L ↗)**
Disability	Conserv	0.57	0.58	+1.71	+0.01	Parallel	0.36	0.41	+14.75	+0.05	Diverging
	Liberals	0.45	0.48	+5.07	+0.02	**(C ↗ = L ↗)**	0.31	0.30	-2.45	-0.01	**(C ↗ > L →)**
Body weight	Conserv	0.51	0.54	+5.55	+0.03	Parallel	0.69	0.80	+15.86	+0.11	Diverging
	Liberals	0.42	0.44	+3.45	+0.01	**(C ↗ = L ↗)**	0.61	0.62	+1.68	+0.01	**(C ↗ > L →)**
Transgender	Conserv	0.30	0.37	+21.77	+0.07	Diverging	1.20	1.58	+31.39	+0.38	Diverging
	Liberals	0.04	0.02	-32.67	-0.01	**(C ↗ > L ↘)**	0.16	0.08	-51.61	-0.08	**(C ↗ > L ↘)**

Note. Start and end values, as well as percent change (**%Δ**) and raw change (Δ) are calculated from the start and end points of the decomposed trend line (removing seasonality and noise). Using these decomposed trend values rather than raw monthly estimates eliminates results that may emerge from an outlier month that was unusually high or low. The interpretation indicates whether the individual subgroup trends have moved in parallel or non-parallel (diverging or converging) patterns based on criteria outlined in Charlesworth & Banaji (2021). The "=" symbol refers to the two groups moving at similar rates; the ">" symbol refers to one group having a faster trend, in that the first listed group showed a faster trend than the second listed group; and the "&" refers to two groups moving in opposite directions. Arrows indicate direction of the trend (downward to neutral, ↘ ; no trend, →; or upwards to neutral, ↗).

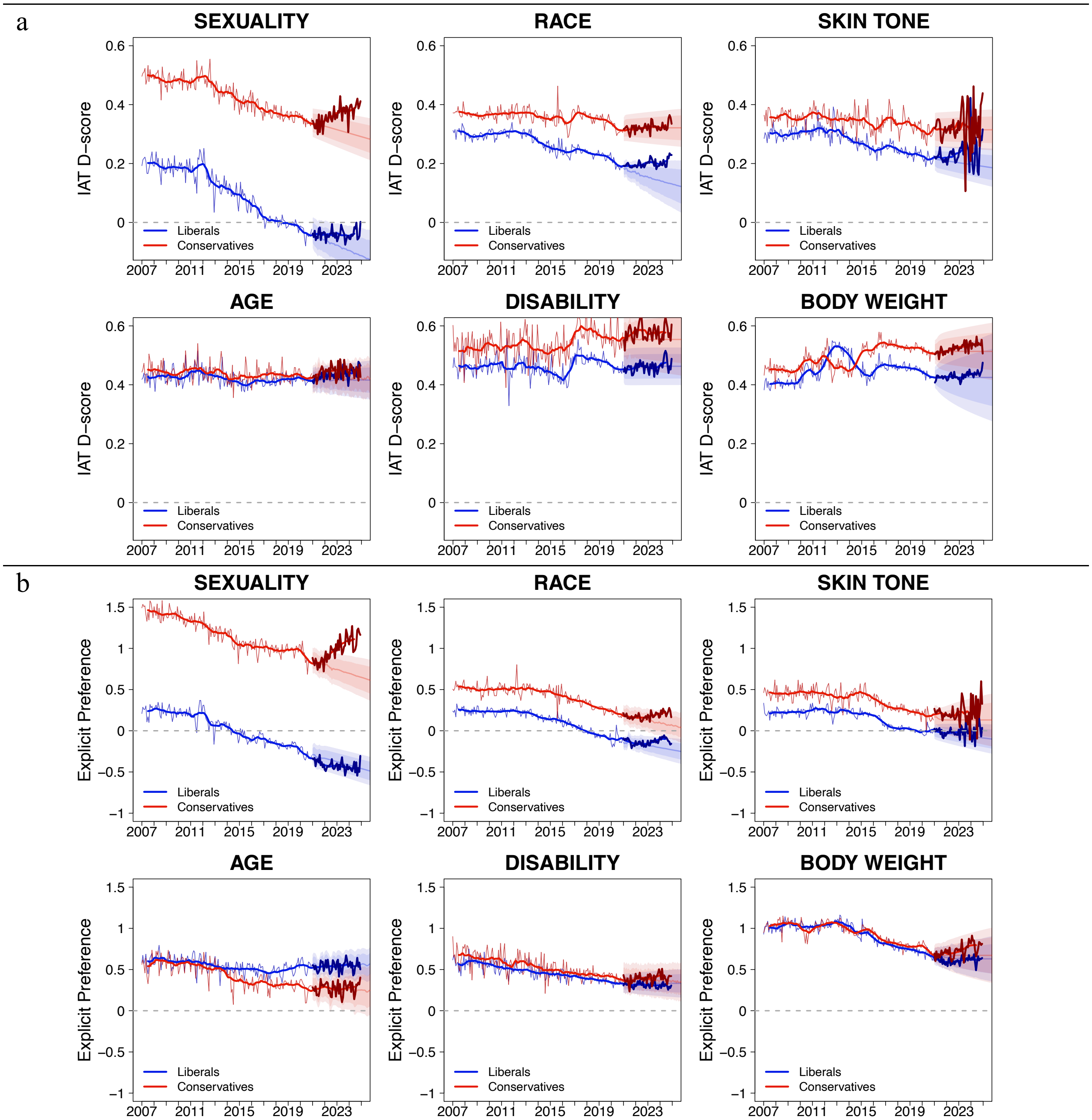


Figure 2. **Political differences (liberals versus conservatives) on key implicit (a) and explicit (b) attitudes from 2007-2025.** Thin blue or red lines indicate the monthly weighted means (from 2007-2020); the thick blue or red lines overlaying those thin lines indicate the decomposed trends (removing seasonality and noise) for the 2007-2020 data. Darker shaded areas indicate 80% confidence interval (CI), light shaded areas indicate 95% CI of the ARIMA model forecasts, and thin lines within the shaded areas indicate the forecasted monthly means, all based on the 2007-2020 data. Thick lines within the shaded areas indicate the observed monthly weighted means of new data (from 2020-2024).

Demographic differences: Age. Past work showed that, from 2007-2020, younger people were generally showing the fastest decreases in bias, compared to older people (Charlesworth & Banaji, 2022b). In sharp contrast to this finding, the newer trends from 2021-2024 show that younger respondents have generally shown the fastest relative *increases in bias* on sexuality, transgender and race attitudes (Figure 3). On sexuality attitudes, anti-gay/pro-straight implicit bias nearly tripled among young people (moving from 0.02 to 0.07, corresponding to a 242% increase over just 4 years), despite remaining relatively stable among older people (increase of 5%). Parallel findings were seen for implicit anti-transgender/pro-cisgender attitudes, with larger rises seen in young respondents (Table 6). Explicit sexuality attitudes and, to a lesser extent explicit transgender attitudes, also showed larger increases in bias among younger respondents. For instance, for sexuality attitudes, younger respondents started at moderate pro-gay/anti-straight preferences (M = -0.17) but showed backsliding towards neutrality (M = -0.01), or a rise of 93%. The bias of older respondents increased relatively slower (increase of 9%).

Race attitudes showed similarly faster reversals among younger respondents: implicit anti-Black/pro-White attitudes rose by 14% among young people over the four years (2021-2024). In contrast, anti-Black/pro-White biases of older people remained stable (decreasing by only 4%). And for explicit race attitudes, younger respondents increased by 64%, reversing from pro-Black/anti-White preferences (M = -0.11) in 2021 to neutrality in 2024 (M = -0.04), whereas again, older respondents did not change (only 3% increase in explicit race attitudes).

Such age differences were also apparent for disability and body weight attitudes. On explicit attitudes, younger respondents increased by 20% and 31% (for disability and weight attitudes, respectively) while older respondents remained more stable (increase of 6% and 2%). And, on implicit attitudes, young people increased slowly (10% and 8% for disability and body weight, respectively) but still at twice the rate than older respondents (3% and 4%). Interestingly, the only attitude that showed no age differences in trends were implicit and explicit anti-old/pro-young attitudes, where both younger and older respondents remained largely stable, with virtually identical raw change over the 4-year period.

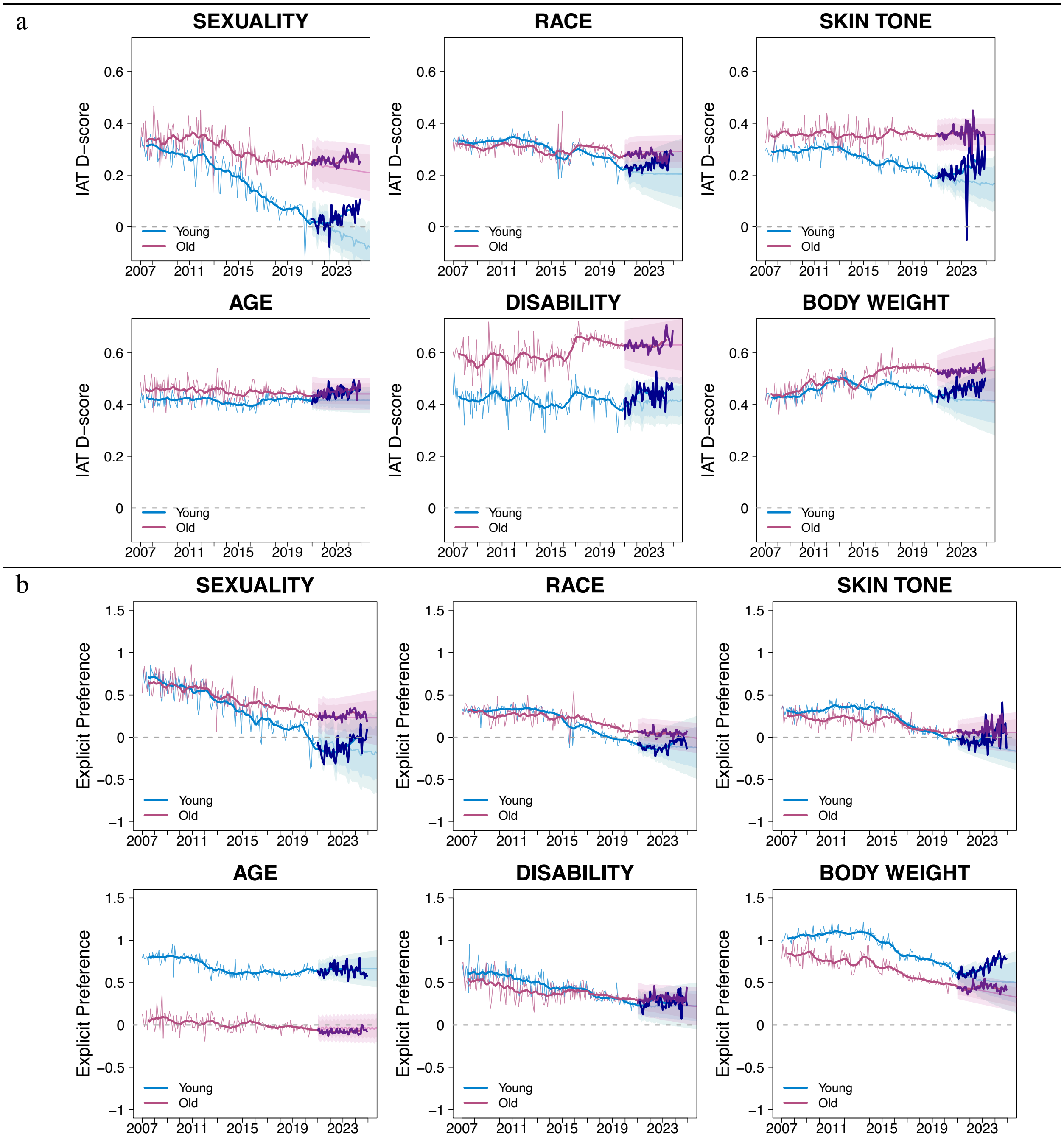


Figure 3. **Age differences (younger <20 years versus older > 40 years) on all implicit (a) and explicit (b) attitudes from 2007-2025.** Thin blue or purple lines indicate the monthly weighted means (from 2007-2020); the thick blue or purple lines overlaying those thin lines indicate the decomposed trends (removing seasonality and noise) for the 2007-2020 data. Darker shaded areas indicate 80% confidence interval (CI), light shaded areas indicate 95% CI of the ARIMA model forecasts, and thin lines within the shaded areas indicate the forecasted monthly means, all based on the 2007-2020 data. Thick lines within the shaded areas indicate the observed monthly weighted means of new data (from 2020-2024).

Table 6.

Age Differences in Implicit and Explicit Attitude Trends in Overall Data (2007-2024)

		Implicit Attitudes					***Explicit Attitudes***				
Attitude	***Group***	***Start***	***End***	***%Δ***	***Raw Δ***	***Interpretation (2007-2024)***	***Start***	***End***	***%Δ***	***Raw Δ***	***Interpretation (2007-2024)***
Sexuality	Younger	0.02	0.07	+241.80	+0.05	Converging	-0.17	-0.01	+92.83	+0.16	Converging
	Older	0.26	0.27	+5.35	+0.01	(**Y** ↗ > **O** ↗)	0.24	0.27	+9.25	+0.02	(**Y** ↗ > **O** ↗)
Race	Younger	0.23	0.26	+13.63	+0.03	Converging	-0.11	-0.04	+63.91	+0.07	Converging
	Older	0.28	0.27	-3.78	-0.01	(**Y** ↗ > **O** ↗)	0.05	0.05	-3.49	-0.00	(**Y** ↗ > **O** ↗)
Skin-tone	Younger	0.20	0.25	+23.29	+0.05	Converging	-0.04	0.09	+312.87	+0.13	Converging
	Older	0.35	0.37	+4.31	+0.02	(**Y** ↗ > **O** ↗)	0.06	0.11	+86.22	+0.05	(**Y** ↗ > **O** ↗)
Age	Younger	0.43	0.45	+5.87	+0.03	Parallel	0.62	0.63	+1.78	+0.01	Parallel
	Older	0.44	0.46	+4.96	+0.02	(**Y**→ = **O** →)	-0.08	-0.07	+18.26	+0.01	(**Y**→ = **O** →)
Disability	Younger	0.41	0.45	+10.42	+0.04	Converging	0.23	0.28	+20.42	+0.05	Converging
	Older	0.63	0.65	+3.47	+0.02	(**Y** ↗ > **O**→)	0.29	0.31	+6.08	+0.02	(**Y** ↗ > **O** ↗)
Body weight	Younger	0.44	0.47	+7.64	+0.03	Converging	0.60	0.79	+31.32	+0.19	Diverging
	Older	0.52	0.54	+4.04	+0.02	(**Y** ↗ > **O**→)	0.42	0.43	+1.54	+0.01	(**Y** ↗ > **O** ↗)
Transgender	Younger	0.01	0.06	+389.94	+0.05	Converging	0.27	0.35	+30.02	+0.08	Parallel
	Older	0.24	0.18	-23.97	-0.06	(**Y** ↗ > **O** ↘)	0.53	0.61	+15.40	+0.08	(**Y** ↗ = **O** ↗)

See note to Table 5.

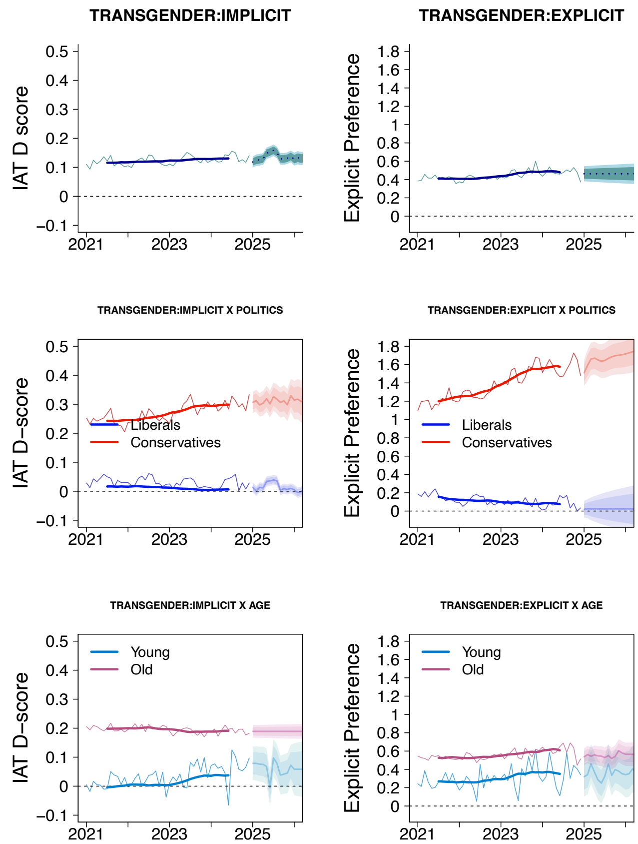


Figure 4. **Overall implicit and explicit transgender attitudes from 2021-2024, as well as political and age differences (younger <20 years versus older > 40 years).** See caption to Figure 2.

Demographic differences: Intersection of politics, age, and gender. There are numerous possible intersectional demographics that could be explored and, thus, we provide all data openly at the Open Science Framework to do just that. However, for our purposes, we chose to focus our exploration on the three-way intersection of age, politics, and gender (e.g., young, conservative, men) because this intersection has been recently shown to shape political opinions in recent polling. Specifically, the opinions and voting patterns of young, conservative men are polarizing away from other age-political-gender groups (R. Campbell & Cowper-Coles, 2025; Harvard Kennedy School Institute of Politics, 2024). Older liberal women have also shown unique increases in anti-gay and anti-transgender biases in line with the emergence of trans-exclusionary radical feminism or "TERF" ideologies (Pearce et al., 2020). We thus focus on age-political-gender differences in sexuality and race attitudes which have the largest and most consistently collected data for intersectional analysis (Figure 5; Table 7).

First, for sexuality attitudes, we find that six (out of eight) of the intersectional age-political-gender groups show increased in implicit bias, with the largest raw and proportional increases in older liberal women (increased by +0.07 IAT D score points, or 82%), followed by young conservative women and men (+0.05 point or 16% increase and +0.03 point or 8% increase, respectively). Similarly, five (out of eight) intersectional groups increased in explicit sexuality biases, again with the largest proportional changes among older liberal women (+0.05 points, or 98% increase) and the largest raw increases (in terms of magnitude) among young conservative men and women (+0.27 and +0.23 points). Notably, due to the higher initial biases of young conservative men and women, these raw increases amounted to percentage increases of only 21% and 28% on explicit anti-gay bias, but the raw increases still reflect that these groups showed particularly large backsliding. In sum, increased bias in sexuality attitudes appears to arise most strongly in both: (1) older liberal women (perhaps from "TERF"-like ideology); and (2) young conservative men and women (perhaps from online rhetoric like conservative "red pill" and "manosphere" spaces).

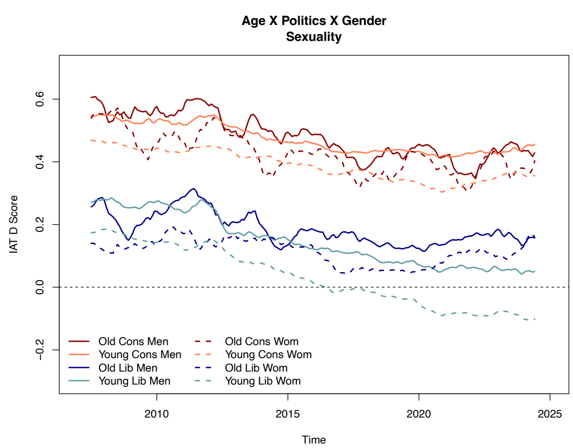
Second, for implicit race attitudes, all subgroups of young people showed large raw increases (8-18% increases), with young conservative men and women both increasing by +0.03, alongside young liberal men and women (+0.04, +0.03-point increases). Similarly, on explicit attitudes, all four subgroups of young people showed parallel movement away from pro-Black attitudes. Specifically, young liberal men and women started with pro-Black attitudes but backslid to neutral by +0.05 and +0.06 explicit attitude points, respectively. Similarly, young conservative men and women started out with anti-Black attitudes and backslid to even stronger anti-Black attitudes by +0.09 and +0.07 points, respectively. By contrast, older liberal women and men, as well as older conservative women all showed *decreases* towards less explicit bias. Overall, this underscores that: (1) the young vs. old age difference on race attitudes is robust and spans intersections with politics and gender; while (2) the unique pattern of older liberal women increasing in bias is specific to sexuality attitudes, reinforcing a sexuality-specific source of change (like the "TERF" ideology).

Table 7.
Age, Politics, Gender Intersectional Differences in Implicit and Explicit Sexuality and Race Attitude Descriptive Trends in Recent Data (2021-2024)

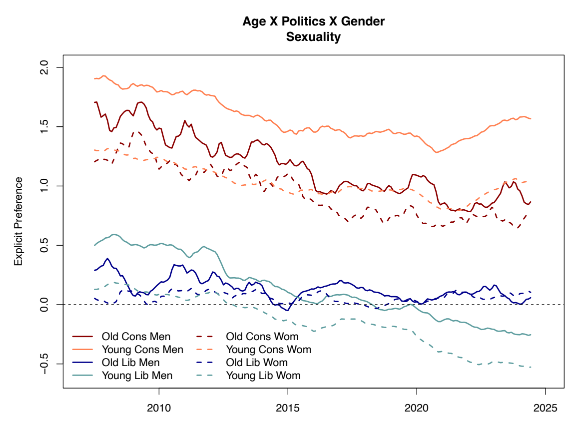
	Implicit: Sexuality			***Explicit: Sexuality***			***Implicit: Race***			***Explicit: Race***		
Group	***Start***	***Raw Δ***	***%Δ***	***Start***	***Raw Δ***	***%Δ***	***Start***	***Raw Δ***	***%Δ***	***Start***	***Raw Δ***	***%Δ***
Young lib men	0.06	-0.01	-11%	-0.13	-0.12	-89%	0.23	+0.04	+16%	-0.06	+0.05	+95%
Older lib men	0.14	+0.02	+15%	0.08	-0.02	-25%	0.29	+0.01	+3%	0.18	-0.04	-22%
Young lib women	-0.09	-0.02	-18%	-0.39	-0.13	-34%	0.15	+0.03	+18%	-0.23	+0.06	+26%
Older lib women	0.09	+0.07	+82%	0.05	+0.05	+98%	0.25	-0.01	-6%	0.07	-0.06	-86%
Young cons men	0.42	+0.03	+8%	1.30	+0.27	+21%	0.36	+0.03	+8%	0.31	+0.09	+29%
Older cons men	0.42	+0.01	+3%	0.85	+0.01	+2%	0.36	+0.01	+2%	0.35	+0.01	+4%
Young cons women	0.31	+0.05	+16%	0.81	+0.23	+28%	0.34	+0.03	+8%	0.25	+0.07	+28%
Older cons women	0.38	+0.02	+5%	0.66	+0.12	+18%	0.36	-0.01	-3%	0.25	-0.07	-28%

Note. Start and end values, as well as percent change (%Δ) and raw change (Δ) are calculated from the start and end points of the decomposed trend line (removing seasonality and noise). Using these decomposed trend values rather than raw monthly estimates eliminates results that may emerge from an outlier month that was unusually high or low.

A.



B.



C. D.

Figure 5. **Age by Politics by Gender Intersections in Overall Trends (2021-2024).** Panel A shows implicit sexuality attitudes; Panel B shoes explicit sexuality attitudes; Panel C shows implicit race attitudes; and Panel D shows explicit race attitudes. Dashed lines indicate data from women; solid lines indicate data from men. Dark red lines indicate older conservatives; orange lines indicate younger conservatives; dark blue indicate older liberals; light blue indicate younger liberals

Other demographic differences. Other demographic differences along gender, race, and religion were smaller and less consistent across measures and attitude topics. That is, there were mostly no meaningful gender differences in the increases since 2020 (see SM): both men and women have been increasing slightly or stable in implicit attitudes for sexuality (men: 4%, women: 10%), race (men: 9%, women: 5%), skin-tone (men: 11%, women: 4%), age (men: 5%, women: 4%), disability (men: 3%, women: 6%) and weight attitudes (men: 7%, women: 3%). Similarly, men and women have both shown similar (small to moderate) increases for explicit sexuality, race, and age attitudes (see SM); although men have increased more quickly than women for explicit disability, body weight, and skin-tone attitudes. Critically, neither men nor women were ever found to *decrease* in bias, underscoring the widespread backsliding of attitudes regardless of respondents' gender.

There were also no race differences on recent implicit attitude trends: White, Black, and Asian participants all showed similar stability (or small increases) across all attitude topics (see

SM). On explicit attitudes, there were a few more racial group differences in trends, often with Black respondents increasing in bias faster than White respondents (for sexuality, disability, and body weight attitudes), except for on explicit race and skin-tone attitudes where White respondents increased in bias more quickly than Black respondents.

Finally, we observed no consistent religion differences in implicit attitude trends from 2020-2024, with both non-religious and Christian respondents generally showing small increases or stability in parallel (see SM). As above, there were a few more differences on explicit attitudes, such as in explicit sexuality attitudes where Christian respondents remained relatively stable and strongly biased while non-religious respondents decreased in anti-gay/pro-straight bias slightly. But, critically, no single religious group always moved faster than the others (see SM), underscoring that religion is not as central as politics or age for explaining recent trend reversals.

RQ3.2: How do reversals in attitudes vary by geography and geographic correlates?

We now turn to geographic differences across U.S. states, focusing on sexuality attitudes because of their large sample size and large relative reversal in bias. Previous data showed that sexuality attitudes were decreasing in implicit and explicit bias across all states (Charlesworth & Banaji, 2021). In stark contrast, with the new data from 2021-2024, we found that now 32 out of 50 states (64%) on implicit attitudes and 27 out of 50 states (54%) on explicit attitudes had *increased* in bias (see Table S13.1).

These reversals towards increasing bias were even more notable when considering only conservative respondents within the states: conservative respondents increased in implicit attitudes across 80% (33/41) of available states, and in explicit attitudes across 90% (37/41) of states. In contrast, liberal respondents increased in implicit attitudes for only 40% (21/50) of states and in explicit attitudes for 14% (7/50) of available states. This result underscores the robust political difference in sexuality bias trends, with conservative respondents increasing in bias largely regardless of where they live.

Correlation of sexuality attitudes with Trump votes in 2020. Next, we consider whether geographic variation in state-level politics (operationalized as Trump vs. Biden vote margin in the 2020 federal U.S. election) may help explain recent backsliding. We first found that, as expected, state-level variation in the overall *magnitude* of anti-gay/pro-straight attitudes was significantly correlated with votes for Trump in 2020: states with higher Trump 2020 votes also had stronger anti-gay/pro-straight attitudes, $r = .62$ [.41, .77], $p <.001$ for explicit attitudes, and $r = .61$ [.41, .76], $p <.001$ for implicit attitudes. However, the amount of recent *change* (i.e., the increase from 2021-2024) was not correlated with higher Trump 2020 votes, $r = .01$ [-.27, .29], $p =.95$ for explicit attitudes and $r = -.13$ [-.39, .15], $p =.37$ for implicit attitudes. Thus, while place-based politics may help explain which places are, *on average,* higher or lower in bias, place-based politics do not appear to explain which places are more likely to *rise* in bias. That is, geography (e.g., the conservativism of a state) plays less of a role than demography (e.g., the conservativism of a person) in shaping the recent patterns of bias increase.

Correlation of sexuality attitudes with transgender attitudes. We also considered the possibility of prejudice spillover between sexuality and transgender attitudes (Charlesworth & Hatzenbuehler, 2024). Both spatial and temporal correlations across U.S. states showed that the *magnitude* of sexuality and transgender attitudes are closely related. For instance, states with higher anti-gay attitudes also had higher anti-transgender attitudes at $r = .85$, $p < .001$ for implicit attitudes, and $r = .87$, $p < .001$ for explicit attitudes. Additionally, a monthly rise in anti-gay attitudes corresponded to a simultaneous rise in anti-transgender attitudes in that same month for both implicit, $r = .39$, $p = .006$ and explicit attitudes, $r = .51$, $p < .001$.

However, Granger causality analyses suggested no consistent evidence for a lagged relationship between sexuality and transgender attitudes, even across various analytic strategies and model specifications (see SM, pages 49-55). Exploratory analyses (see OSF) suggested there may be a dynamic relationship such that rises in sexuality attitudes significantly preceded rises in transgender attitudes early in the timeseries (2021-2023) but not later in the timeseries (2023-2024). Perhaps, early on (~2021-2023), when transgender issues were still relatively less discussed, anti-transgender bias may have emerged from anti-gay bias. Critically however, we never found evidence that a rise in anti-gay attitudes may have been preceded by a rise in anti-transgender attitudes, as had been hypothesized. Thus, a single causal explanation of prejudice spillover (from anti-transgender legislation and sentiment to the reappearance of anti-gay bias) is not supported in the current data.

Correlation of sexuality attitudes with "sexual grooming" rhetoric. Finally, we consider the impact of "sexual grooming" rhetoric—the moral propaganda that older gay and lesbian people are grooming young children to become gay/lesbian themselves (Cohen & Galloway, 2025). As expected, Google search trends showed that the relative rate of searching for "sexual grooming" topics increased drastically in 2020/2021, temporally coinciding with the breakpoint identified in reversed attitude trends. Additionally, we found significant monthly temporal correlations with explicit sexuality attitudes, $r = .33$, $p = .03$, implying that a monthly spike in Google searches for "sexual grooming" coincided with a monthly spike in anti-gay attitudes. However, we found no such correlations for implicit attitudes, suggesting that moral panics of "sexual grooming" may only be related to explicit cognition.

Moreover, we again found no consistent evidence for any lagged relationships (see SM, pages 56-59 for Granger causality models). Thus, as above, we rule out the explanation that rising sexuality attitudes are exclusively attributable to one topic of moral panic. Rather, as we will discuss below, the unexpected backsliding of all explicit and implicit attitudes seems to be attributable to a more complex additive compilation of social events and cultural changes.

Summary of results.

There are many insights to be gained from this compiled dataset of 9.5 million responses now covering seven attitude topics over 18 years. For simplicity, we focus on three new empirical conclusions to guide future theory development. First, across all topics, and on both implicit and explicit measures, we see widespread stalling—even reversals—in intergroup attitudes. These observed trends deviated from past trends and model forecasts that had largely predicted continued movement to neutrality. Additionally, exploratory breakpoint detection suggested that these reversals may happen first among implicit attitudes and, approximately a year later, in explicit attitudes, suggesting that anti-egalitarian norms may first affect the more automatic, culturally tuned, implicit attitudes (Payne et al., 2017).

Second, for those tasks that showed the largest backsliding (sexuality, transgender, and race attitudes), the greatest increases were surprisingly concentrated among younger respondents and, for sexuality and transgender attitudes, among conservatives. Some intersectional age-politics-gender differences also emerged, with younger conservative men and women often showing the most consistent and large raw changes. And yet, the new trends were generally widespread across most demographics for most tasks, leading us to conclude that large-scale attitude changes are likely driven by macro-level events that affect a broad range of demographic groups.

Third, correlational analyses showed that sexuality attitudes were, as expected, significantly correlated with conservative voting, transgender attitudes, and sometimes with

Google searches for anti-gay rhetoric of “sexual grooming”. However, there was little evidence for causal, lagged relationships between these variables. Thus, we suggest that common explanations for rising anti-gay/pro-straight bias (e.g., from a single conservative event, from prejudice spillover, or from sexual grooming rhetoric) are unlikely to be the single operating variables. Instead, the results imply a more complex network of events are activating negative group associations and re-emboldening the expression of prejudice.

General Discussion

For nearly a decade, surveys of long-term attitude change have shown hopeful evidence of steadily declining biases in the U.S. and internationally, especially for some highly discussed topics like race and sexuality (Charlesworth & Banaji, 2019, 2022b; Kurdi et al., 2025). The current data on attitudes measured 2021-2024 complicate that picture. Instead of expected declines in bias, attitudes stalled or reversed in their movement towards neutrality for all topics on both implicit and explicit attitudes.

For example, forecasts from past data predicted that implicit sexuality attitudes should decrease by 44% but we found that, in reality, the attitudes increased by 9%. Without such data, many people may have assumed that “all is well,” and attitudes remained on track toward equality. When people assume such continued decreases in prejudice, they may stop being concerned about maintaining or furthering those changes (Wright & Baray, 2012). Thus, in showing that change halted and even reversed, these data help motivate researchers, and the public, to act towards renewed efforts for attitude change.

Proposing (and ruling out) sources of increasing bias.

Reversals in attitude trends are not without precedent: In previous analyses (Charlesworth & Banaji, 2022b), small reversals in attitudes occurred around 2016. These reversals showed discriminant and predictive validity: they were specific to only implicit race, disability, and body weight attitudes and each coincided with an event targeting that group such as racist rhetoric in the Republican National Convention, public shaming of a disabled journalist, and viral fat-shaming comments. But critically, within a year, these reversals had halted and returned to their previous trends of decreasing bias.

The reversals observed starting in 2021 are not of this type. For one, they are not isolated to one topic but are widespread across every topic, both implicit and explicit measures, and nearly every demographic group. And, for another, the current reversals are longer lasting, with breakpoints detected in mind-2020 and mid-2021 suggesting enduring new slopes. Such endurance and spread suggest that no one targeted event (e.g., one piece of legislation, one election) is the likely source of recent change. Indeed, direct tests of common hypotheses about rising sexuality biases such as moral panics of "sexual grooming" show that, although the backsliding in 2021 coincides with recent rises in anti-trans and anti-gay rhetoric, there is no evidence of causal relationships.

Instead, we propose that increasing bias reflects a complex process whereby a combination of destabilizing events activated existential concerns that were translated into intergroup hostility. As discussed in the Introduction, numerous events coincided around 2020-2024 including: threats from the COVID pandemic; rising costs of living and economic inequality; and, most obviously, unprecedented political sectarianism (Brenan, 2025) and challenges to electoral processes (Bright Line Watch, 2025). This perspective helps explain why the reversals happened around 2020/2021, even after the election of President Biden. Even though Biden might have been expected to symbolize liberal, egalitarian values, the 2020 election was vocally challenged, exacerbating partisan divides and fracturing civil discourse.

We suggest that this compilation of events and erosion of public trust activated existential psychological concerns around health, economics, and security that, in turn, were converted into intergroup hostility and scapegoating (Bursztyn et al., 2022). Such intergroup hostility was also mainstreamed by elite political rhetoric (Finkel et al., 2020) and amplified by online social media algorithms (Brady et al., 2023), ultimately creating a widespread normative ecosystem that emboldens intergroup prejudice.

Rising bias among younger people.

Although this explanation is post-hoc, we underscore that it not only explains the timing and spread of the reversals across all topics but also helps explain the surprising demographic difference whereby younger people have reversed relatively faster than older people. This result was surprising because younger respondents are typically at the vanguard of egalitarian progress (Charlesworth & Banaji, 2022b). However, consider that young people in the U.S. today express the strongest existential concerns about their future economic, environmental, and political situation (Helliwell et al., 2024). These existential concerns, in turn, get funneled towards intergroup hostility—especially in young-dominated “manosphere” and “red pill” online communities weaponizing anti-LGBTQ and anti-Black discourse (Ging, 2019; Ribeiro et al., 2021). Thus, young people today appear primed towards attitude backsliding. Critically, the current work suggests such toxic ideologies don’t remain hidden in online corners of sub-populations (e.g., young conservative men) but spread to entire populations of young people, including young liberals and young women, who are also grappling with existential concerns.

Of course, young people may also be changing faster towards greater bias because they are particularly attuned to social norms and peer influence (Somerville, 2013) making them most likely to witness and adopt new cultural trends. Indeed, related work identifies young people as early adopters of new trends (e.g., in technology; Morris et al., 2005). This may also be the case for intergroup attitudes, such that young people are both the fastest to change towards neutrality (when egalitarian norms dominate) but also the fastest to reverse towards hostility (when anti-egalitarian norms emerge).

Limitations.

Causal inference is notoriously difficult in observational data. While we speculate that the observed reversals likely stem from a compilation of destabilizing events alongside pro-prejudice media, future work should attempt to test this post-hoc explanation and isolate causal variables of (1) events (e.g., COVID, election challenges), (2) existential concerns (e.g., about health, economics), and (3) amplified toxic online rhetoric. For instance, difference-in-difference approaches could examine how local or international variation in the timing and co-occurrence of events and online rhetoric explain variation in the timing and degree of attitude backlash.

Relatedly, our reliance on Project Implicit data, which is large but non-representative, limits any claims about representative population attitudes. At the same time, we note that the current findings align with representative surveys (Gallup, 2025), and we ruled out deflationary concerns of sample change, repeat test-takers, or cohort replacement. Still, both the Project Implicit data and representative polls under-represent extremist groups, despite these groups' disproportionate impact in shaping norms (Bursztyn et al., 2020). Future work identifying the attitudes of more extremist groups using naturalistic media such as Gab, 4chan, or Twitter/X, will help identify whether these groups may be early sources of attitude backsliding.

Conclusion.

Long-term attitude change is fragile – trends that were previously declining in bias can stall or even entirely reverse and increase in bias. In the U.S.—a culture that still explicitly and widely endorses values of “diversity” and “inclusion” (Isenberg & Brauer, 2024)—these trends are concerning. We suggest that the reversals arise from a combination of events, beginning around 2020, that eroded the societal fabric of the U.S., activated existential concerns, and transformed into intergroup hostility. While this hostility spread across all attitude topics and nearly every demographic group, the effects were amplified among conservatives and young people—groups most affected by existential concerns and most exposed to online toxic rhetoric. Ultimately, as anti-democratic and anti-egalitarian threats unfold in U.S. society, continued tracking of implicit and explicit attitudes will remain critical to quantify, understand, and intervene on the devolution of intergroup harmony.

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.

Blumer, K., & Kleinberg, J. (2025). Tracking patterns in toxicity and antisocial behavior over user lifetimes on large social media platforms. *Scientific Reports*, *15*(25369). https://doi.org/10.1038/s41598-025-07086-3

Brady, W. J., Jackson, J. C., Lindström, B., & Crockett, M. J. (2023). Algorithm-mediated social learning in online social networks. *Trends in Cognitive Sciences*, *27*(10). https://doi.org/10.1016/j.tics.2023.06.008

Brenan, M. (2025). *U.S. Political Parties Historically Polarized Ideologically*. https://news.gallup.com/poll/655190/political-parties-historically-polarized-ideologically.aspx

Bright Line Watch. (2025). *Violence, redistricting, and democratic norms in Trump's America | Bright Line Watch*. https://brightlinewatch.org/violence-redistricting-and-democratic-norms-in-trumps-america/

Bursztyn, L., Egorov, G., Fiorin, S., Banerjee, A., Cantoni, D., Duflo, E., Enke, B., Fisman, R., Hassan, T., List, J., Kamenica, E., Perez-Truglia, R., Schilbach, F., Shleifer, A., Voth, H.-J., Yuchtman, N., Han, R., Kang-Landsberg, A., Kao, A., … Whit, P. (2020). From Extreme to Mainstream: The Erosion of Social Norm. *American Economic Review*, *110*(11), 3522–3548. https://doi.org/10.1257/aer.20171175

Bursztyn, L. ;, Egorov, G. V. ;, Haaland, I. ;, Rao, A. ;, & Roth, C. (2022). Scapegoating during crises. *AEA Papers and Proceedings*, *112*, 151–155. https://doi.org/10.1257/pandp.20221069

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

Campbell, R., & Cowper-Coles, M. (2025). From a gap to a gulf? Gen Z and the gender gap in political attitudes in post-industrial societies. *European Journal of Politics and Gender*, 1–11. https://doi.org/10.1332/25151088Y2025D000000093

Campbell, S. M., Twenge, J. M., & Keith Campbell, W. (2017). Fuzzy But Useful Constructs: Making Sense of the Differences Between Generations. *Work, Aging and Retirement*, *3*(2), 130–139. https://doi.org/10.1093/WORKAR/WAX001

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/ASSET/IMAGES/LARGE/10.1177_0956797618813087-FIG1.JPEG

Charlesworth, T. E. S., & Banaji, M. R. (2021). Patterns of Implicit and Explicit Attitudes II. Long-Term Change and Stability, Regardless of Group Membership. *American Psychologist*, *76*(6), 851–869. https://doi.org/10.1037/amp0000810

Charlesworth, T. E. S., & Banaji, M. R. (2022a). Evidence of Covariation Between Regional Implicit Bias and Socially Significant Outcomes in Healthcare, Education, and Law Enforcement. *Handbook on Economics of Discrimination and Affirmative Action*, 1–21. https://doi.org/10.1007/978-981-33-4016-9_7-1

Charlesworth, T. E. S., & Banaji, M. R. (2022b). 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

Charlesworth, T. E. S., & Hatzenbuehler, M. L. (2025). The Stigma Stability Framework: An Integrated Theory of How and Why Society Transmits Stigma Across History. *Social and Personality Psychology Compass*, *19*(3), e70051. https://doi.org/10.1111/SPC3.70051

Cohen, A., & Galloway, S. R. (2025). Grooming Authoritarianism: Anti-Trans/ Queer Panic as Pedagogy for Democratic Decline. *PS: Political Science & Politics*, *58*(3: Contemporary and Future LGBTQ+ Scholarship in Political Science), 456–460. https://doi.org/10.1017/S1049096524001197

Cryer, J. D., & Chan, K.-S. (2008). *Time Series Analysis*. Springer New York. https://doi.org/10.1007/978-0-387-75959-3

Desilver, D. (2024). *Prices up across US metros, but in some more than others | Pew Research Center*. https://www.pewresearch.org/short-reads/2024/10/07/prices-are-up-in-all-us-metro-areas-but-some-much-more-than-others/

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., & Druckman, J. N. (2020). Political sectarianism in America: A poisonous cocktail of othering, aversion, and moralization poses a threat to democracy. *Science*, *370*(6516), 533–536. https://doi.org/10.1126/science.abe1715

Gallup. (2025). *LGBTQ+ Rights*. https://news.gallup.com/poll/1651/gay-lesbian-rights.aspx

Gawronski, B. (2025). Implicit Bias: What Is It and How Does It Matter for Diversity, Equity, and Inclusion. In V. M. Esses, J. F. Dovidio, J. Jetten, D. Sekaquaptewa, & K. West (Eds.), *The Sage Handbook of Psychological Perspectives on Diversity, Equity and Inclusion*. Sage.

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

Ging, D. (2019). *Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere*. *22*(4), 638–657. https://doi.org/10.1177/1097184X17706401

Greenwald, A. G., & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-295X.102.1.4

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

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

Harvard Kennedy School Institute of Politics. (2024). *Spring 2024 Harvard Youth Poll*. https://iop.harvard.edu/youth-poll/47th-edition-spring-2024

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

Helliwell, J. F., Layard, R., Sachs, J. D., Neve, J.-E. De, Aknin, L. B., & Wang, S. (2024). *World Happiness Report 2024*. *World Happiness Report*, 1–158. https://worldhappiness.report/ed/2024/

Hopkins, D. J., & Washington, S. (2020). THE RISE OF TRUMP, THE FALL OF PREJUDICE? TRACKING WHITE AMERICANS' RACIAL ATTITUDES VIA A PANEL SURVEY, 2008-2018. *Public Opinion Quarterly*, *84*(1), 119–140. https://doi.org/10.1093/poq/nfaa004

Hyndman, R. J., & Athanasopoulos, G. (2013). *Forecasting : principles and practice*. OTexts. https://www.otexts.org/fpp

Isenberg, N., & Brauer, M. (2024). Diversity and inclusion have greater support than most Americans think. *Scientific Reports 2024 14:1*, *14*(1), 1–9. https://doi.org/10.1038/s41598-024-76761-8

Karwowski, M., Kowal, M., Groyecka, A., Białek, M., Lebuda, I., Sorokowska, A., & Sorokowski, P. (2020). When in Danger, Turn Right: Does Covid-19 Threat Promote Social Conservatism and Right-Wing Presidential Candidates? *Human Ethology*, *35*(1), 37–48. https://doi.org/10.22330/HE/35/037-048

Kumar, V., Kodipady, A., & Young, L. (2023). A psychological account of the unique decline in anti-gay attitudes. *Philosophical Psychology*. https://doi.org/10.1080/09515089.2023.2249500

Kurdi, B., & Banaji, M. R. (2021). Implicit social cognition: A brief (and gentle) introduction. In A. S. Reber & R. Allen (Eds.), *The cognitive unconscious: The first half-century*. Oxford University Press.

Kurdi, B., Charlesworth, T. E. S., & Mair, P. (2025). International Stability and Change in Explicit and Implicit Attitudes: An Investigation Spanning 33 Countries, Five Social Groups, and 11 Years (2009–2019). *Journal of Experimental Psychology: General*. https://doi.org/10.1037/XGE0001746

Lewis, D. C., Flores, A. R., Haider-Markel, D. P., Miller, P. R., & Taylor, J. K. (2022). Transitioning Opinion? Assessing the Dynamics of Public Attitudes Toward Transgender Rights. *Public Opinion Quarterly*, *86*(2), 343–368. https://doi.org/10.1093/POQ/NFAC014

Mamakos, M., Charlesworth, T. E. S., & Finkel, E. J. (2025). Moralizing partisanship when surrounded by copartisans versus in mixed company. *Proceedings of the National Academy of Sciences: Nexus*. https://doi.org/10.1093/pnasnexus/pgaf105

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Gavrilov, D., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2020). COVID-19 Pandemic. *Our World in Data*. https://ourworldindata.org/coronavirus

Morehouse, K. N., Charlesworth, T. E. S., Devos, T., & Banaji, M. R. (2025). Who is American? A comprehensive analysis of the American = White/Foreign = Asian stereotype (2007-2023). *Scientific Reports*, *15*(1), 3368. https://doi.org/10.1038/S41598-024-83794-6,

Morris, M. G., Venkatesh, V., & Ackerman, P. L. (2005). Gender and age differences in employee decisions about new technology: An extension to the theory of planned behavior. *IEEE Transactions on Engineering Management*, *52*(1), 69–84. https://doi.org/10.1109/TEM.2004.839967

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

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

Pearce, R., Erikainen, S., & Vincent, B. (2020). TERF wars: An introduction. *Sociological Review*, *68*(4), 677–698. https://doi.org/10.1177/0038026120934713;PAGEGROUP:STRING:PUBLICATION

Pew Research Center. (2019). *Views on Race in America 2019*. https://www.pewsocialtrends.org/2019/04/09/race-in-america-2019/

Ribeiro, M. H., Blackburn, J., Bradlyn, B., Cristofaro, E. De, Stringhini, G., Long, S., Greenberg, S., & Zannettou, S. (2021). The Evolution of the Manosphere across the Web. *Proceedings of the International AAAI Conference on Web and Social Media*, *15*, 196–207. https://doi.org/10.1609/ICWSM.V15I1.18053

Schmidt, K., & Axt, J. R. (2016). Implicit and Explicit Attitudes Toward African Americans and Barack Obama Did not Substantively Change During Obama's Presidency. *Social Cognition*, *34*(6), 559–588. https://doi.org/10.1521/soco.2016.34.6.559

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression.* Cambridge University Press. https://doi.org/10.2307/2655372

Somerville, L. H. (2013). The Teenage Brain: Sensitivity to Social Evaluation. *Current Directions in Psychological Science*, *22*(2), 121–127. https://doi.org/10.1177/0963721413476512;PAGE:STRING:ARTICLE/CHAPTER

Tankard, M. E., & Paluck, E. L. (2017). The Effect of a Supreme Court Decision Regarding Gay Marriage on Social Norms and Personal Attitudes. *Psychological Science*, 095679761770959. https://doi.org/10.1177/0956797617709594

Wright, S. C., & Baray, G. (2012). Models of social change in social psychology: collective action or prejudice reduction? Conflict or harmony? In J. DIxon & M. Levine (Eds.), *Beyond Prejudice* (pp. 225–247). Cambridge University Press. https://doi.org/10.1017/CBO9781139022736.014