Online Supplement for

# International stability and change in explicit and implicit attitudes: An investigation spanning 33 countries, five social groups, and 11 years (2009–2019)

Benedek Kurdi[1]*, Tessa E. S. Charlesworth[2]*, and Patrick Mair[3]

[1] Department of Psychology, University of Illinois Urbana–Champaign, Champaign, IL (kurdi@illinois.edu)

[2] Management and Organizations Department, Kellogg School of Management, Northwestern University, Evanston, IL (tessa.charlesworth@kellogg.northwestern.edu)

[3] Department of Psychology, Harvard University, Cambridge, MA

* B.K. and T.E.S.C. contributed equally to this work.

## Table of Contents

## Additional Details on Sample Composition I: Change in Country Sample Sizes

The primary data used in this project were first introduced by Charlesworth, Navon and colleagues (2023). That paper reports all country sample sizes and the distributions of demographic variables by country. Readers interested in country-level sample sizes should refer to Table 3 in Charlesworth, Navon et al. (2023); those interested in the distribution of demographic variables by country (e.g., percent liberal or percent female-identifying participants) should refer to Table 5. Given that our focus in the present project is less on the specific cross-country differences, we instead provide a high-level summary visualization of the average country-level sample size for each of the five tasks.

The popularity of the Project Implicit websites has grown rapidly over the past decade around the world. Thus, as expected, the sample sizes within the current paper also show changes across time, with rising trends in the sample sizes for all five tasks. Specifically, between 2009 and 2019 the mean country-level sample size for the age task increased from 285 to 1,081, for the race task from 1,008 to 2,459, for the sexuality task from 535 to 1,404, for the skin tone task from 276 to 918, and for the body weight task from 334 to 1,060 (see Figure S1).

Of note, this rising sample size could explain, at least in part, the observed convergence of attitudes across countries. That is, as the data increase in size, the estimates get more precise and more stable, thus spuriously suggesting that countries have become more similar over time.

To empirically test this idea, we turned to Monte Carlo simulations. We used the country-level annual means and standard deviations of explicit sexuality attitudes as the starting point given that this is the test characterized by the highest level of convergence over time. On each of 1,000 iterations, we then drew random samples for each country assuming a normal distribution with the empirically observed means and standard deviations using two different sample sizes:

(*a*) for the unequal sample size simulations, we used the actually observed — unequal — country-by-year sample sizes, and (*b*) for the equal sample size simulations, we artificially set the country-by-year sample sizes to 1,000.

Once the samples were obtained, on each iteration we fit country-level INLA models as described in the main text and stored the country-level standard deviation of the fitted values for the beginning and the end of the time series. We then calculated a difference score by subtracting the time-1 standard deviation from the time-101 standard deviation. As such, positive values of the difference score correspond to divergence (larger standard deviation at the end than at the beginning of the time series), whereas negative values of the difference score correspond to convergence (smaller standard deviation at the end than at the beginning of the time series).

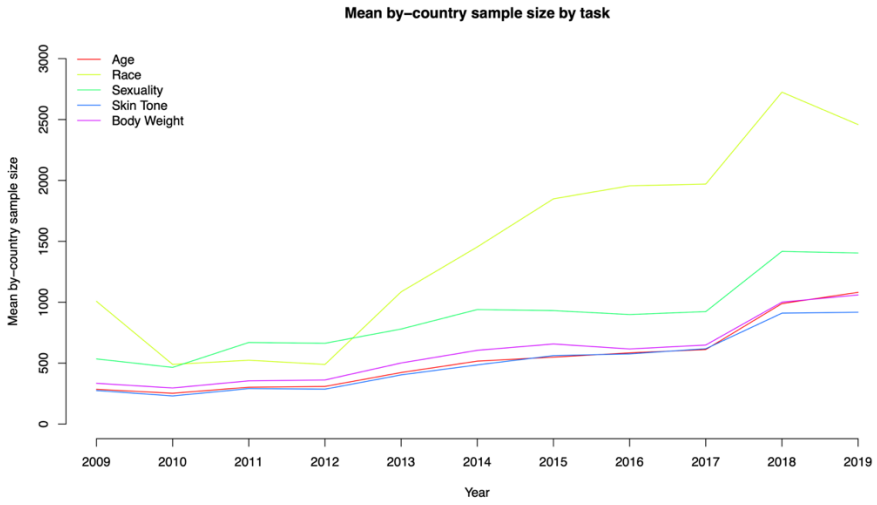
The results of the simulations are shown in Figure S2. A few patterns are notable. First, country-level divergence of attitudes is the more likely outcome under both sample size assumptions given that the mean of each distribution is positive (mean = 0.008 for the unequal simulations and mean = 0.012 for the equal simulations). Second, the standard deviation of each distribution is notably different, with 0.022 for the unequal simulations and 0.008 for the equal simulations, suggesting a broader range of potential outcomes under unequal sample sizes. Third, and critically, whereas virtually all iterations show divergence under the equal sample size assumption (940/1,000 or 94%), this is the case for only 665/1,000 (66.5%) iterations under the unequal sample size assumption. In the remaining 33.5% of cases, we observed convergence over time.

To summarize, convergence over time — the pattern observed in the empirical data — was close to six times more likely to occur under the unequal than under the equal sample size assumption. As such, we can conclude that the unequal country-level sample sizes across years

may well have resulted in a spurious finding of country-level convergence in the analyses reported in the main text. Therefore, we believe that these results should be interpreted with the appropriate caution.

**Figure S1**

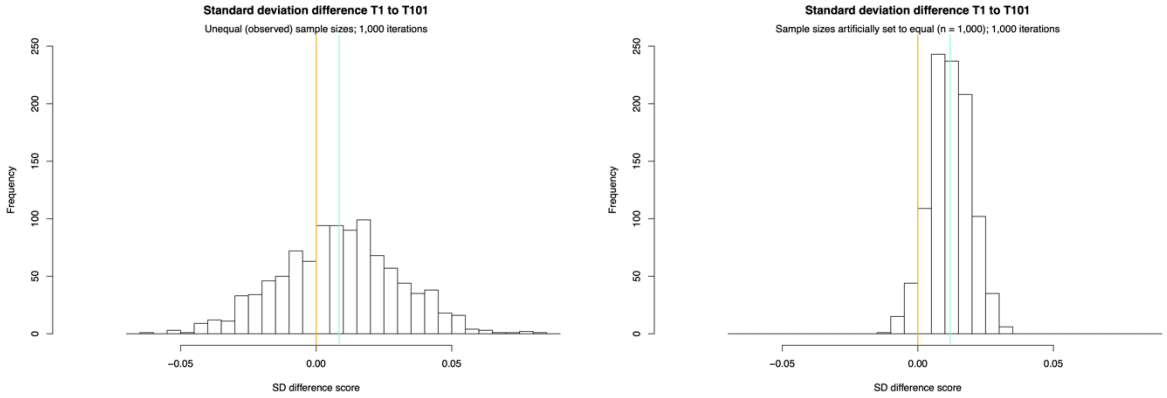
*Mean Country Sample Sizes by Task*




*Note.* Mean by-country sample size over the duration of the Project Implicit International dataset (2009–2019) for each task. Age is shown in red, race in green, sexuality in teal, skin tone in blue, and body weight in purple.

**Figure S2**

*Simulations of Country-Level Convergence Assuming Unequal Sample Sizes (Left Pane) vs. Equal Sample Sizes (Right Pane)*




*Note.* Each distribution is based on 1,000 independent simulations. Positive scores correspond to divergence and negative scores correspond to convergence over time. The orange lines mark zero (no divergence or convergence over time); the aqua lines show the mean of each distribution.

## Additional Details on Sample Composition II: Reasons for Visiting the Project Implicit Websites

A commonly stated perception of the participants visiting the Project Implicit websites is that they are mostly composed of volunteer respondents who seek out and self-select the website and are therefore uncommonly motivated to understand and challenge their implicit biases. As in Charlesworth and Banaji (2019), we address this objection by examining the reasons that participants state for arriving at the Project Implicit websites. However, given that the Project Implicit International websites did not collect information on respondents' reasons for coming to the websites, we analyze the data collected for this question on the U.S. website.

Among respondents arriving at the U.S. website, the majority are not coming from self-selection or self-directed searches. Rather, 44.82% come for assignments for school and 18.28% come for assignments for work, for a total of 63.10%. Only a relative minority come to the website for more self-directed reasons, with 19.39% coming from mentions in a news story, Internet search, or from a recommendation from a friend, and 5.64% report coming from "other" sources. The remaining participants, ~11%, did not answer.

Given that the international websites are less well-known than the U.S. website, and less frequently covered in news stories or word-of-mouth, we can reasonably expect that, even on the international websites, the majority of respondents come to complete assignments for work or school. Thus, although the sample is a convenience sample of volunteers, it does not necessarily overrepresent those respondents particularly motivated to learn about or personally interested in implicit bias.

## Additional Details on Sample Composition III: Cultural Differences in Response Styles

Participants from different cultures are known to vary in how they approach self-report scales (Baumgartner & Steenkamp, 2001; Chen et al., 1995; Hui & Triandis, 1989; Johnson et al., 2005). Specifically, some cultures (e.g., typically individualist, highly masculine, or high-power-distance countries such as the United States and Canada) tend to use more extreme responses, whereas other cultures (e.g., typically collectivist, low masculine, or low-power-distance countries such as Japan, Turkey, or Portugal) tend to use the middle responses. Notably, these are small effects and typically do not affect comparisons of means across countries (Chen et al., 1995), but they can affect the dispersion and standard deviation of responses.

We emphasize again that the primary aim of the present paper is not to evaluate specific country-level differences (which could be affected by country-level response styles) but rather to provide a general summary of attitude trajectories across countries over time. Nevertheless, understanding potential response style differences across countries can still add nuance to our interpretation of whether the general trend across countries may be masking meaningful variation in how respondents from different countries approach the self-report scales. As such, below we explore whether or not respondents from different countries indeed show different distributions on the explicit attitude measures.

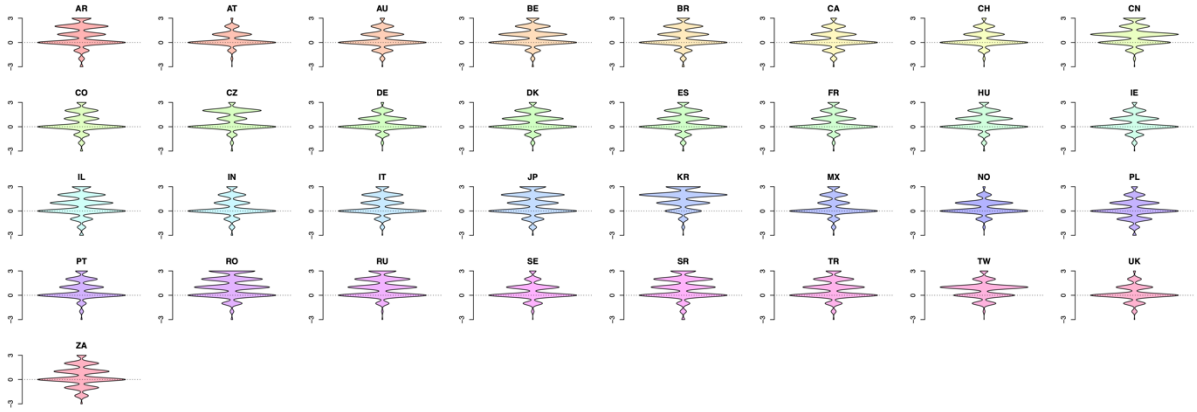
The distribution of self-report responses for each of the five attitude targets (age, race, sexuality, skin tone, and body weight) is shown in Figures S3–S7 below. As expected, given the country-level differences in means, some countries have more responses > 0 than others (indicating an explicit preference for the privileged over the stigmatized group). However, it is also clear that nearly all countries and all tests have a mode at neutrality, with the exception of China, Korea, and Taiwan (which have modes of +1 for the age, race, and skin tone tests), and the body

weight test on which several countries show higher modes (including Belgium, China, the Czech Republic, Denmark, Spain, Israel, Norway, Romania, Russia, Sweden, and Taiwan at +1 and Japan, Korea, and Serbia at +2). No countries show obvious patterns of using only the extreme high or low ends of the scales.

To more formally confirm these visual inspections, we also quantitatively test the differences using Kolmogorov–Smirnov tests between the average (overall) sample distribution (i.e., the average proportion of respondents responding with each value) and each individual country's distribution. Across the 33 tests comparing each country to the average distribution, no significant differences emerged, all $p$s > .212, with $D$s (the statistic for Kolmogorov-Smirnov tests) ranging from [0.14, 0.29] for age, [0.14, 0.57] for race, [0.14, 0.43] for sexual orientation, [0.14, 0.29] for skin tone, and [0.14, 0.29] for body weight. Together, these results indicate that no country-specific distribution for any test is meaningfully different from the overall distribution (given that $D$s of < .05 are customarily considered to indicate meaningful differences).

**Figure S3**

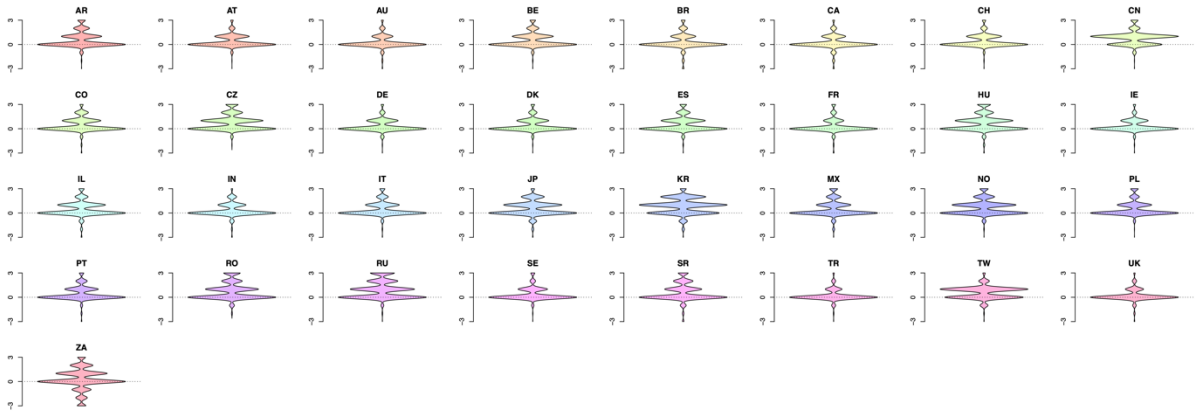
*The Distribution of Explicit Age Attitudes Across 33 Countries*




*Note.* Higher scores correspond to more bias in favor of the privileged over the stigmatized group. AR = Argentina; AT = Austria; AU = Australia; BE = Belgium; BR = Brazil; CA = Canada; CH = Switzerland; CN =China; CO = Colombia; CZ = Czech Republic; DE = Germany; DK = Denmark; ES = Spain; FR = France; HU = Hungary; IE = Ireland; IL = Israel; IN = India; IT = Italy; JP = Japan; KR = South Korea; MX = Mexico; NO = Norway; PL = Poland; PT = Portugal; RO = Romania; RU = Russia; SE = Sweden; SR = Serbia; TR = Turkey; TW = Taiwan; UK = United Kingdom; ZA = South Africa.

**Figure S4**

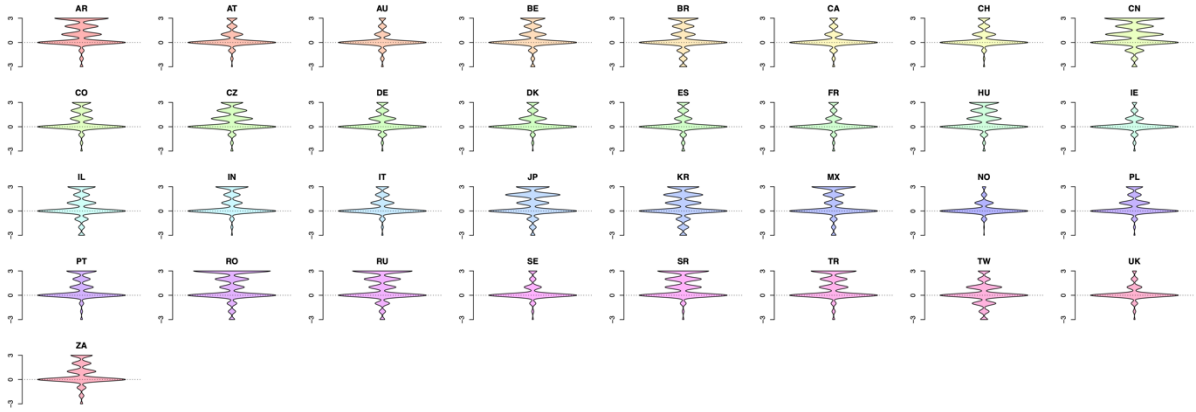
*The Distribution of Explicit Race Attitudes Across 33 Countries*




*Note.* Higher scores correspond to more bias in favor of the privileged over the stigmatized group. AR = Argentina; AT = Austria; AU = Australia; BE = Belgium; BR = Brazil; CA = Canada; CH = Switzerland; CN =China; CO = Colombia; CZ = Czech Republic; DE = Germany; DK = Denmark; ES = Spain; FR = France; HU = Hungary; IE = Ireland; IL = Israel; IN = India; IT = Italy; JP = Japan; KR = South Korea; MX = Mexico; NO = Norway; PL = Poland; PT = Portugal; RO = Romania; RU = Russia; SE = Sweden; SR = Serbia; TR = Turkey; TW = Taiwan; UK = United Kingdom; ZA = South Africa

**Figure S5**

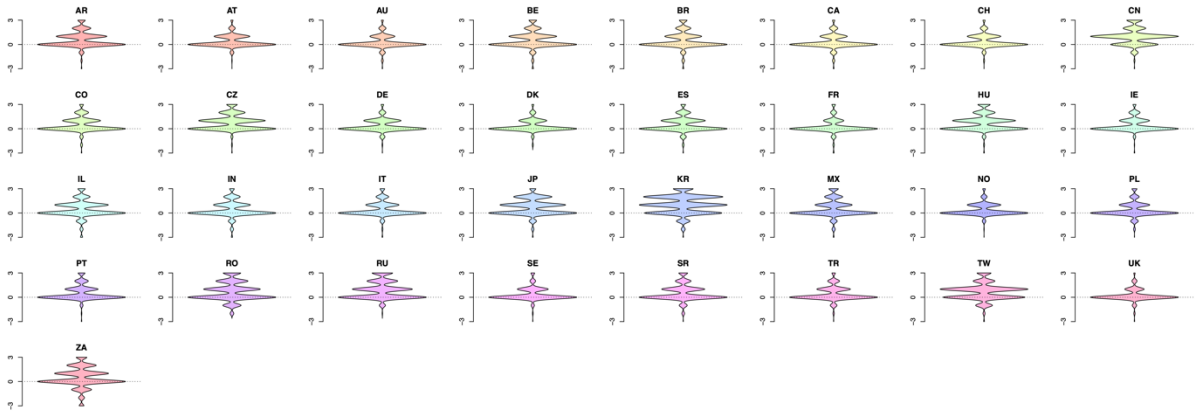
*The Distribution of Explicit Sexuality Attitudes Across 33 Countries*




*Note.* Higher scores correspond to more bias in favor of the privileged over the stigmatized group. AR = Argentina; AT = Austria; AU = Australia; BE = Belgium; BR = Brazil; CA = Canada; CH = Switzerland; CN =China; CO = Colombia; CZ = Czech Republic; DE = Germany; DK = Denmark; ES = Spain; FR = France; HU = Hungary; IE = Ireland; IL = Israel; IN = India; IT = Italy; JP = Japan; KR = South Korea; MX = Mexico; NO = Norway; PL = Poland; PT = Portugal; RO = Romania; RU = Russia; SE = Sweden; SR = Serbia; TR = Turkey; TW = Taiwan; UK = United Kingdom; ZA = South Africa

**Figure S6**

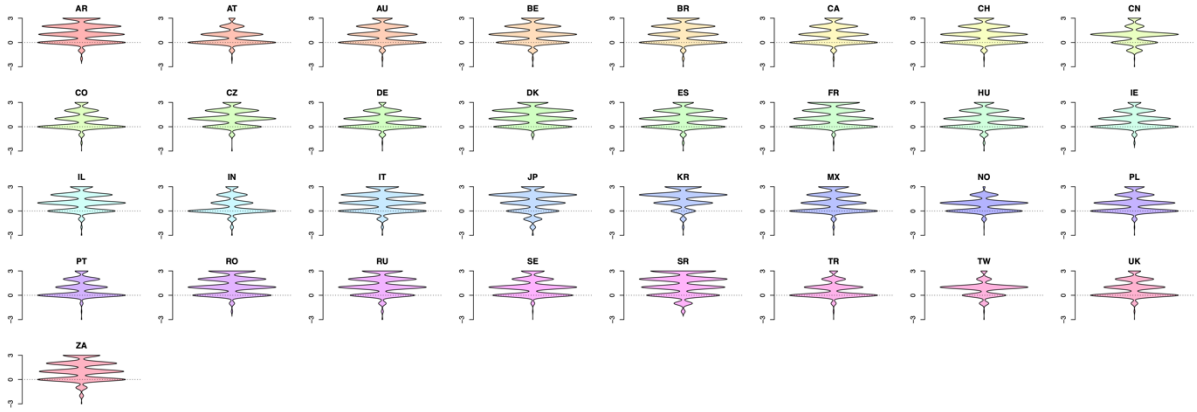
*The Distribution of Explicit Skin Tone Attitudes Across 33 Countries*




*Note.* Higher scores correspond to more bias in favor of the privileged over the stigmatized group. AR = Argentina; AT = Austria; AU = Australia; BE = Belgium; BR = Brazil; CA = Canada; CH = Switzerland; CN =China; CO = Colombia; CZ = Czech Republic; DE = Germany; DK = Denmark; ES = Spain; FR = France; HU = Hungary; IE = Ireland; IL = Israel; IN = India; IT = Italy; JP = Japan; KR = South Korea; MX = Mexico; NO = Norway; PL = Poland; PT = Portugal; RO = Romania; RU = Russia; SE = Sweden; SR = Serbia; TR = Turkey; TW = Taiwan; UK = United Kingdom; ZA = South Africa

**Figure S7**

*The Distribution of Explicit Body Weight Attitudes Across 33 Countries*




*Note.* Higher scores correspond to more bias in favor of the privileged over the stigmatized group. AR = Argentina; AT = Austria; AU = Australia; BE = Belgium; BR = Brazil; CA = Canada; CH = Switzerland; CN =China; CO = Colombia; CZ = Czech Republic; DE = Germany; DK = Denmark; ES = Spain; FR = France; HU = Hungary; IE = Ireland; IL = Israel; IN = India; IT = Italy; JP = Japan; KR = South Korea; MX = Mexico; NO = Norway; PL = Poland; PT = Portugal; RO = Romania; RU = Russia; SE = Sweden; SR = Serbia; TR = Turkey; TW = Taiwan; UK = United Kingdom; ZA = South Africa

## Additional Details on Sample Composition IV: Country-Level Starting and Ending Values

**Table S1**

*Country-Level Starting (2009) and Ending (2019) Values of Explicit Attitudes (Standardized; Standard Deviation Units)*

| Country | Age | | Race | | Sexuality | | Skin Tone | | Body Weight | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 |
| Austria | 0.00 | 0.37 | 0.45 | 0.39 | 0.48 | 0.30 | 0.37 | 0.37 | 0.49 | 0.46 |
| Argentina | 0.76 | 0.56 | 0.78 | 0.51 | 1.10 | — | 0.61 | 0.43 | — | 0.63 |
| Colombia | — | 0.34 | — | 0.47 | — | 0.43 | — | 0.46 | — | 0.51 |
| Czech Republic | 0.75 | 0.70 | 0.89 | 0.92 | 0.67 | 0.47 | 0.94 | 0.63 | 0.76 | 0.71 |
| Denmark | — | 0.61 | — | 0.45 | — | 0.21 | — | 0.38 | — | 0.86 |
| South Africa | 0.21 | 0.16 | 0.47 | 0.22 | 0.56 | 0.28 | 0.75 | 0.35 | 0.77 | 0.45 |
| Portugal | 0.76 | 0.63 | 0.76 | 0.18 | 0.91 | 0.20 | 0.66 | 0.37 | 0.52 | 0.42 |
| Russia | 0.78 | 0.76 | 1.34 | 0.76 | 0.84 | 0.74 | 1.11 | 0.69 | 0.96 | 0.85 |
| Turkey | 0.76 | 0.63 | 0.22 | 0.25 | 1.20 | 0.72 | 0.36 | 0.37 | 0.63 | 0.40 |
| Norway | 0.43 | 0.26 | 0.79 | 0.72 | 0.33 | 0.16 | 0.48 | 0.26 | 0.60 | 0.45 |
| India | 0.57 | 0.32 | 0.39 | 0.25 | 1.21 | 0.57 | 0.71 | 0.35 | 0.71 | 0.35 |
| Israel | 0.79 | 0.44 | 0.91 | 0.48 | 0.52 | 0.45 | 0.71 | 0.32 | 0.88 | 0.65 |
| Switzer-land | 0.57 | 0.42 | 0.59 | 0.32 | 0.58 | 0.26 | 0.34 | 0.21 | 0.71 | 0.61 |
| Serbia | 0.66 | 0.88 | 0.44 | 1.03 | 0.59 | 0.58 | 0.35 | 0.73 | 0.75 | 1.03 |
| Romania | 1.08 | 0.54 | — | 0.98 | 1.10 | 0.51 | 0.74 | 0.70 | 0.81 | 0.59 |
| Ireland | 0.05 | 0.43 | 0.83 | 0.44 | 0.43 | 0.16 | 0.75 | 0.34 | — | 0.48 |
| Taiwan | — | 0.50 | — | 0.58 | — | 0.32 | — | — | — | — |
| Poland | — | 0.20 | — | 0.48 | — | 0.25 | — | 0.39 | — | 0.58 |
| Mexico | 0.73 | 0.25 | — | 0.70 | 1.02 | 0.47 | 0.98 | 0.44 | 1.03 | 0.75 |
| Japan | 0.64 | 0.71 | 0.80 | 0.59 | 0.77 | 0.58 | 0.7 | 0.73 | 0.82 | 0.84 |
| Hungary | 0.57 | 0.62 | 0.91 | 0.57 | 0.87 | 0.47 | 0.75 | 0.54 | 0.59 | 0.76 |
| Brazil | 0.75 | 0.54 | 0.64 | 0.19 | 0.96 | 0.23 | 0.61 | 0.38 | 0.90 | 0.70 |
| Germany | 0.44 | 0.35 | 0.41 | 0.35 | 0.50 | 0.27 | 0.44 | 0.37 | 0.55 | 0.46 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| France | 0.51 | 0.43 | 0.28 | 0.30 | 0.40 | 0.21 | 0.23 | 0.24 | 0.99 | 0.81 |
| Belgium | 0.63 | 0.24 | 0.82 | 0.42 | 0.70 | 0.16 | 0.67 | 0.72 | 0.72 | 0.81 |
| Australia | 0.57 | 0.30 | 0.42 | 0.21 | 0.43 | 0.12 | 0.34 | 0.25 | 0.76 | 0.62 |
| Sweden | 0.36 | 0.37 | 0.61 | 0.33 | 0.43 | 0.10 | 0.56 | 0.28 | 0.84 | 0.65 |
| Spain | 0.89 | 0.41 | 0.88 | 0.55 | 0.39 | 0.17 | 0.66 | 0.42 | 0.76 | 0.67 |
| Italy | 0.60 | 0.49 | 0.52 | 0.55 | 0.67 | 0.27 | 0.43 | 0.44 | 0.88 | 0.74 |
| China | 0.63 | 0.64 | 0.95 | 0.83 | 0.76 | 0.58 | 0.93 | 0.96 | 0.56 | 0.56 |
| Canada | 0.45 | 0.31 | 0.39 | 0.12 | 0.39 | 0.20 | 0.31 | 0.23 | 0.79 | 0.60 |
| Korea | 1.17 | 1.05 | 1.29 | 0.80 | 0.21 | 0.69 | 1.26 | 0.98 | 1.26 | 0.99 |
| United Kingdom | 0.29 | 0.18 | 0.43 | 0.15 | 0.44 | 0.12 | 0.36 | 0.16 | 0.71 | 0.51 |

**Table S2**

*Country-Level Starting (2009) and Ending (2019) Values of Implicit Attitudes (Standardized; Standard Deviation Units)*

| Country | Age | | Race | | Sexuality | | Skin Tone | | Body Weight | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 | 2009 | 2019 |
| Austria | 1.80 | 0.94 | 0.84 | 0.60 | 0.44 | 0.25 | 0.70 | 0.76 | 0.82 | 0.98 |
| Argentina | 0.81 | 0.92 | 0.71 | 0.64 | 0.95 | — | 0.74 | 0.83 | — | 0.66 |
| Colombia | — | 0.72 | — | 0.60 | — | 0.50 | — | 0.80 | — | 0.66 |
| Czech Republic | 0.86 | 1.00 | 0.88 | 0.89 | 0.41 | 0.31 | 0.81 | 1.11 | 1.15 | 0.90 |
| Denmark | — | 0.92 | — | 0.62 | — | 0.08 | — | 0.68 | — | 0.81 |
| South Africa | 0.90 | 0.81 | 0.65 | 0.56 | 0.84 | 0.74 | 1.11 | 0.83 | 0.98 | 0.85 |
| Portugal | 0.82 | 0.81 | 0.86 | 0.76 | 0.64 | 0.37 | 0.79 | 0.77 | 0.66 | 0.74 |
| Russia | 0.83 | 0.87 | 0.78 | 0.73 | 0.60 | 0.51 | 0.72 | 0.92 | 0.51 | 0.57 |
| Turkey | 0.86 | 0.90 | 0.78 | 0.69 | 0.47 | 0.24 | 0.73 | 0.98 | 0.72 | 0.66 |
| Norway | 0.81 | 0.89 | 0.55 | 0.59 | 0.47 | 0.25 | 0.54 | 0.74 | 0.69 | 0.87 |
| India | 0.62 | 0.69 | 0.74 | 0.74 | 0.92 | 0.63 | 0.58 | 0.82 | 0.29 | 0.26 |
| Israel | 0.86 | 0.74 | 0.83 | 0.73 | 0.52 | 0.56 | 0.71 | 0.87 | 0.55 | 0.57 |
| Switzer-land | 0.83 | 0.95 | 0.49 | 0.60 | 0.49 | 0.15 | 0.51 | 0.70 | 0.89 | 0.90 |
| Serbia | 0.68 | 0.75 | 0.33 | 0.75 | 0.66 | 0.68 | 0.46 | 0.98 | 0.29 | 0.39 |
| Romania | 0.84 | 0.78 | — | 0.81 | 0.94 | 0.52 | 0.95 | 0.87 | 0.47 | 0.53 |
| Ireland | 0.69 | 0.78 | -0.01 | 0.66 | 1.01 | 0.42 | 1.14 | 0.83 | — | 0.75 |
| Taiwan | — | 0.91 | — | 0.84 | — | -0.17 | — | — | — | — |
| Poland | — | 0.79 | — | 0.75 | — | 0.45 | — | 0.75 | — | 0.76 |
| Mexico | 0.79 | 0.74 | — | 0.66 | 0.82 | 0.52 | 0.73 | 0.78 | 0.71 | 0.68 |
| Japan | 0.75 | 0.92 | 0.97 | 0.93 | 0.21 | 0.00 | 0.88 | 1.21 | 0.59 | 0.59 |
| Hungary | 0.86 | 0.93 | 0.89 | 0.80 | 0.72 | 0.43 | 0.74 | 0.96 | 0.72 | 0.78 |
| Brazil | 0.75 | 0.72 | 0.74 | 0.58 | 0.58 | 0.15 | 0.73 | 0.63 | 0.55 | 0.58 |
| Germany | 0.90 | 1.01 | 0.66 | 0.68 | 0.41 | 0.18 | 0.64 | 0.76 | 0.83 | 0.95 |
| France | 0.89 | 0.94 | 0.62 | 0.70 | 0.39 | 0.15 | 0.62 | 0.73 | 0.85 | 0.80 |
| Belgium | 1.01 | 0.77 | 0.68 | 0.66 | 0.71 | 0.25 | 0.72 | 0.59 | 0.76 | 0.76 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Australia | 0.72 | 0.78 | 0.66 | 0.69 | 0.59 | 0.35 | 0.68 | 0.80 | 0.72 | 0.79 |
| Sweden | 0.77 | 0.87 | 0.55 | 0.50 | 0.36 | 0.10 | 0.57 | 0.67 | 0.85 | 0.87 |
| Spain | 0.83 | 0.86 | 0.74 | 0.75 | 0.42 | 0.30 | 0.66 | 0.80 | 0.75 | 0.75 |
| Italy | 0.85 | 0.91 | 0.80 | 0.91 | 0.68 | 0.43 | 0.75 | 0.84 | 0.75 | 0.81 |
| China | 0.79 | 0.83 | 0.81 | 0.88 | 0.21 | 0.06 | 0.78 | 0.99 | 0.14 | 0.30 |
| Canada | 0.76 | 0.78 | 0.72 | 0.63 | 0.54 | 0.37 | 0.62 | 0.76 | 0.75 | 0.76 |
| Korea | 0.82 | 0.86 | 0.73 | 0.71 | 0.37 | 0.58 | 0.70 | 0.79 | 0.46 | 0.54 |
| United Kingdom | 0.76 | 0.81 | 0.64 | 0.60 | 0.60 | 0.40 | 0.57 | 0.75 | 0.67 | 0.69 |

## Additional Details on Sample Composition V: Supplemental Studies 1A–1B of Self-Selection into Tests

In addition to selecting themselves into participating in any study on Project Implicit in the first place, once they have arrived on the website, participants also have an opportunity to select themselves into particular tests. In response to reviewer feedback, we conducted two studies, which we report as Supplemental Studies 1A–1B here, in which we probed whether allowing participants to choose a test from a list — like they do on the Project Implicit demonstration website and the Project Implicit international websites — results in systematically different estimates of implicit and explicit attitudes relative to a condition in which participants are randomly assigned to take a test.

### Method

#### *Open Science Practices*

We report all measures, manipulations, and exclusions for both studies. The studies were not preregistered. Raw data files, including trial-level IAT data, analysis scripts, and materials are available for download from the Open Science Framework (OSF; https://osf.io/bfqgu/).

#### *Participants and Design*

Participants were 2,058 adult volunteers in Supplemental Study 1A and 2,072 adult volunteers in Supplemental Study 1B. Participants were recruited from the Project Implicit educational website (http://implicit.harvard.edu/). In line with standard practice (Greenwald et al., 2003), we excluded participants from subsequent analyses if they (*a*) did not complete the Implicit Association Test (IAT; Greenwald et al., 1998), which served as the main dependent measure (*n*s = 93 and 87), or (*b*) had response latencies of 300 ms or lower on at least 10% of IAT trials (*n*s = 42 and 22), indicating inattention. These exclusions left 1,920 participants in the final

sample of Supplemental Study 1A and 1,963 participants in the final sample of Supplemental Study 1B.

1,288 participants identified as female, 561 participants as male, and 34 participants as other genders in Supplemental Study 1A; 1,375 participants identified as female, 480 participants as male, and 43 participants as other genders in Supplemental Study 1B. Most participants (72.21% in Supplemental Study 1A and 68.72% in Supplemental Study 1B) were citizens of the United States, with 86 unique countries represented in Supplemental Study 1A and 99 in Supplemental Study 1B. Mean participant age was 32 years (*SD* = 14 years) in Supplemental Study 1A and 34 years (*SD* = 14 years) in Supplemental Study 1B. In Supplemental Study 1A, 1,081participants identified as White, 238 participants as Asian, 237 participants as Hispanic, 180 participants as Black, 110 participants as multiracial, 40 participants as Middle Eastern, 11 participants as Native American, and eight participants as Pacific Islander; in Supplemental Study 1B, 1,047 participants identified as White, 261 participants as Hispanic, 247 participants as Black, 195 participants as Asian, 132 participants as multiracial, 27 participants as Middle Eastern, 24 participants as Native American, and nine participants as Pacific Islander. Further demographic details are available in the open data.

In both studies, participants were randomly assigned to a choice condition or a random condition. In the choice condition (*n*s = 964 and 973), participants were allowed to select one of five IATs to complete, as they would on the Project Implicit demonstration or international website; in the random condition (*n*s = 956 and 990), participants were randomly assigned to one of the same five IATs.

***Materials***

The images used to represent each category on the IAT were borrowed for use from the Project Implicit demonstration and international websites and are available in the open materials.

***Procedure and Measures***

The procedure followed the procedure used on the Project Implicit demonstration and international websites. Specifically, participants completed measures of implicit attitudes and explicit attitudes (in randomized order). Additionally, in Supplemental Study 1B, participants completed study-specific demographic items and two exploratory scales.

**Implicit Attitudes.** Implicit social group attitudes were measured using versions of the Implicit Association Test (IAT; Greenwald et al., 1998). Each participant completed one of five IATs (age, race, sexual orientation, skin tone, or body weight). In the choice condition, participants selected one of the five IATs; in the random condition, participants were randomly assigned to one of the five IATs.

The category labels and category stimuli were identical to the ones used on the Project Implicit demonstration and international websites. The attribute labels were *Good* and *Bad*. The attribute stimuli for the former attribute included *Love*, *Cheer*, *Friend*, *Pleasure*, *Adore*, *Cheerful*, *Friendship*, *Joyful*, *Smiling*, *Cherish*, *Excellent*, *Glad*, *Joyous*, *Spectacular*, *Appealing*, *Delight*, *Excitement*, *Laughing*, *Attractive*, *Delightful*, *Fabulous*, *Glorious*, *Pleasing*, *Beautiful*, *Fantastic*, *Happy*, *Lovely*, *Terrific*, *Celebrate*, *Enjoy*, *Magnificent*, and *Triumph*, and for the latter attribute included *Abuse*, *Grief*, *Poison*, *Sadness*, *Pain*, *Despise*, *Failure*, *Nasty*, *Angry*, *Detest*, *Horrible*, *Negative*, *Ugly*, *Dirty*, *Gross*, *Evil*, *Rotten*, *Annoy*, *Disaster*, *Horrific*, *Scorn*, *Awful*, *Disgust*, *Hate*, *Humiliate*, *Selfish*, *Tragic*, *Bothersome*, *Hatred*, *Hurtful*, *Sick*, and *Yucky*. For each participant, eight attribute stimuli each were randomly selected from this larger set.

The IAT consisted of seven blocks: (*a*) category practice (categories depending on specific IAT, e.g., old people vs. young people; 20 trials); (*b*) attribute practice (pleasant vs. unpleasant words; 20 trials); (*c*) first critical block (categories depending on specific IAT, e.g., old people/good vs. young people/bad; 20 trials); (*d*) second critical block (categories depending on specific IAT, e.g., old people/good vs. young people/bad; 40 trials); (*e*) reverse category practice (categories depending on specific IAT, e.g., young people vs. old people; 28 trials); (*f*) first reverse critical block (categories depending on specific IAT and first and second critical blocks, e.g., young people/good vs. old people/bad; 20 trials); and (*g*) second reverse critical block (categories depending on specific IAT and first and second critical blocks, e.g., young people/good vs. old people/bad; 40 trials). The placement of category and attribute labels to the left and right side of the screen and the order of the two critical blocks was randomized. Participants used the "E" and "I" keys for categorization and were required to correct inaccurate responses before proceeding.

IAT D scores were calculated using the improved scoring algorithm (Greenwald et al., 2003) such that positive D scores correspond to a preference for the dominant over the stigmatized social group.

**Explicit Attitudes.** Participants responded to a 7-point relative preference item as well as two 10-point feeling thermometer items (one for each target). The order of the three measures was randomized with the constraint that the two feeling thermometer items always followed each other. The two feeling thermometer items were used to calculate a feeling thermometer difference score. Both explicit attitude items were scored such that positive scores correspond to a preference for the dominant over the stigmatized social group, thus paralleling the IAT D score.

**Time-1 Explicit Attitudes (Supplemental Study 1B).** In Supplemental Study 1B, participants additionally completed 7-point relative preference items for all five targets prior to completing the rest of the study (and, in the choice condition, before selecting a test to complete). These items were not analyzed for this project but are available in the open data.

**Exploratory Measures (Supplemental Study 1B).** In Supplemental Study 1B, following the explicit attitude items, participants completed the 10-item Motivation to Respond Without Prejudice Scale (Plant & Devine, 1998), adapted to the particular stigmatized group from the IAT. For example, participants who completed the age IAT completed the Motivation to Respond Without Prejudice Scale with respect to elderly people. Items were presented in randomized order, and participants indicated their degree of agreement or disagreement with each item on a 7-point scale.

Additionally, participants completed an 8-item scale measuring meta-perceptions of how motivated the average person is to express or conceal positive or negative feelings toward each target group. Items were presented in randomized order, and the seven response options ranged from “not at all motivated” to “extremely motivated.” These exploratory scales were not analyzed for this project but are available in the open data.

**Test-Specific Demographic Items (Supplemental Study 1B).** In Supplemental Study 1B, participants additionally completed test-specific demographic items for some tests. These items were administrated at the end of the study, following the exploratory items. Specifically, participants who completed the sexuality IAT indicated their sexual orientation (the options included heterosexual or straight, lesbian or gay, bisexual, queer, and other); participants who completed the skin tone IAT indicated their skin tone using two items, including a 10-point visual scale ranging from very light-skinned to very dark-skinned and a 5-point scale asking about

the participant's skin tone relative to others in their racial/ethnic group, with response options ranging from "much lighter" to "much darker"; and participants who completed the body weight IAT indicated their subjective weight on a 7-point scale with response options ranging from "very underweight" to "very overweight" as well as their weight in pounds and their height in inches, which allowed us to calculate their Body Mass Index (BMI). These variables were used to probe whether any differences observed across the choice and random conditions could be accounted for by systematic task-relevant demographic differences across the two conditions.

### *Analytic Strategy*

Statistical analyses were performed in the R statistical computing environment (version 4.4.1).

## Results

### *Supplemental Study 1A*

First, we tested whether the distribution of participants across tests differed by condition. As expected, in the random condition, participants completed each test with roughly equal probability, including age (21%), body weight (21%), race (20%), sexuality (19%), and skin tone (19%). These proportions did not differ from chance, $\chi^2(4) = 1.96$, $p = .743$, פ = 0.02. In contrast, in the choice condition, participants were most likely to select themselves into the race test (33%), followed by age (23%), body weight (20%), sexuality (14%), and skin tone (10%). These proportions differed from chance, $\chi^2(4) = 158.91$, $p < .001$, פ = 0.20. The two conditions significantly differed from each other, $\chi^2(4) = 67.01$, $p < .001$, V = 0.18.

Most relevant, the mean levels of implicit attitudes and the two explicit attitude measures are shown in Table S3. Overall, we observed a significant main effect of choice condition (choice vs. random), $F(1, 1910) = 12.55$, $p = .001$, partial $\eta^2 = .003$, and a significant main effect

of task (age, body weight, race, sexuality, and skin tone), $F(4, 1910) = 25.01$, $p < .001$, partial $\eta^2 = .05$, on implicit attitudes. The interaction was not significant, $F(4, 1910) = 1.84$, $p = .120$, partial $\eta^2 = .004$.

The main effect of task indicates that implicit attitudes differed across the attitude targets, which is less relevant to the present investigation. Critically, the main effect of choice condition suggests that, overall, participants were more likely to show bias when they were randomly assigned to a task than when they selected themselves into it. Although the interaction was not significant, we probed which tests primarily drove this result, and the effect of choice condition was individually significant only for the sexuality IAT.

Turning to the feeling thermometer difference score, we observed a significant Choice × Task interaction, $F(4, 1873) = 7.37$, $p < .001$, partial $\eta^2 = .02$. Follow-up analyses revealed significant differences for the sexuality task, paralleling those observed on the IAT, and the opposite effect for the body weight task such that participants in the choice condition expressed more anti-fat bias than participants in the random condition.

Finally, we also observed a significant Choice × Task interaction effect on the relative preference item, $F(4, 1855) = 6.39$, $p < .001$, partial $\eta^2 = .01$. Follow-up analyses revealed significant differences for the sexuality and body weight task, paralleling those observed for the relative preference item. Additionally, the difference was also significant for the skin tone task such that participants in the choice condition expressed less anti-dark-skinned bias than participants in the random condition.

***Supplemental Study 1B***

Supplemental Study 1B gave us the opportunity to probe the replicability of these findings and additionally involved the collection of demographic variables. These variables, in turn,

enabled us to investigate whether any robust differences across the choice and random conditions could be explained by task-relevant demographic differences.

The mean levels of implicit attitudes and the two explicit attitude measures are shown in Table S4. Unlike in Supplemental Study 1A, we observed a significant Choice × Task interaction on implicit attitudes, $F(4, 1953) = 4.51$, $p = .001$, partial $\eta^2 = .01$. Follow-up analyses revealed significant differences only for the sexuality task such that participants in the random condition showed higher levels of anti-gay bias than participants in the choice condition. We obtained parallel results on the feeling thermometer, $F(4, 1906) = 2.87$, $p = .022$, partial $\eta^2 = .01$, and relative preference items, $F(4, 1895) = 1.24$, $p = .293$, partial $\eta^2 = .003$, although the interaction for the latter was not statistically significant. The other condition differences observed in Supplemental Study 1A did not replicate.

Finally, we turned to investigating whether the significant condition difference uniquely observed for the sexuality task may have been due to demographic differences across the two conditions. Indeed, we observed a considerably larger percentage of non-straight participants in the choice condition (37%) than in the random condition (15%). This difference was statistically significant, $\chi^2(1) = 23.04$, $p < .001$, $\varphi = 0.25$.

Importantly, once the indicator variable for sexual orientation was added to the model also containing a main effect for condition (choice vs. random), the effect of the former on implicit attitudes was significant, $\beta = 0.92$, $t(379) = 10.66$, $p < .001$, whereas the effect of the latter was reduced to nonsignificance, $\beta = 0.14$, $t(379) = 1.83$, $p = .069$. We obtained a similar result for the feeling thermometer difference score, with a significant effect of sexual orientation, $\beta = 1.43$, $t(377) = 9.19$, $p < .001$, but no effect of condition, $\beta = 0.04$, $t(379) = 0.29$, $p = .775$. Finally, the same pattern also emerged for the relative preference item, with a significant effect of sexual

orientation, $\beta = 1.22$, $t(376) = 9.61$, $p < .001$, but no effect of condition, $\beta = 0.01$, $t(376) = 0.06$, $p = .951$.

**Discussion**

In Supplemental Studies 1A–1B we tested whether selecting oneself into an attitude target (as opposed to being randomly assigned to it) modulates the preferences that participants reveal on the Implicit Association Test as well as those that they express on two sets of explicit attitude items. The only systematic difference that consistently emerged in both studies was a condition effect on both implicit and explicit sexuality attitudes such that participants randomly assigned to the sexuality task showed more implicit and explicit bias than those who chose to complete the same tests. Critically, this condition difference was fully accounted for by a difference in sexual orientation across the two conditions: The proportion of non-straight participants more than doubled in the choice condition, and once this difference was accounted for, the condition effect was reduced to nonsignificance.

Together, these data suggest that the effect of self-selection into tasks on the Project Implicit websites is limited or, in the case of the sexuality test, can be accounted for, and empirically modeled using, demographic variables (e.g., more non-straight people self-select into the sexuality test). Critically for the present project, we investigated the possible effects of participant sexual orientation on explicit and implicit attitude trajectories and found large baseline differences but highly similar trajectories over time (see Table S11 and Figure S14 in the sections below). Thus, even if the number of non-straight respondents coming to the website changes over time, the overall trends toward less bias are expected to be robust to this aspect of sample composition. In sum, self-selection into task does not appear to be a major confound endangering the interpretation of the present results about patterns of change over time.

**Table S3**

*Means and Standard Deviations (in Parentheses) by Condition (Choice vs. Random) and Target (Age, Body Weight, Race, Sexuality and Skin Tone) for the Implicit Association Test, the Feeling Thermometer Difference Score, and the Relative Preference Item (Supplemental Study 1A)*

| | Implicit Association Test | | | Feeling Thermometer | | | Relative Preference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Choice | Random | Difference | Choice | Random | Difference | Choice | Random | Difference |
| Age | 0.44 (0.36) | 0.47 (0.39) | *n.s.* | -0.04 (1.80) | -0.28 (1.94) | *n.s.* | 0.43 (1.02) | 0.39 (1.16) | *n.s.* |
| Body Weight | 0.51 (0.36) | 0.51 (0.39) | *n.s.* | 0.85 (2.44) | 0.28 (1.79) | ** | 0.86 (1.12) | 0.50 (0.87) | *** |
| Race | 0.26 (0.43) | 0.32 (0.40) | *n.s.* | -0.28 (1.77) | -0.06 (1.79) | *n.s.* | 0.04 (0.95) | 0.05 (0.87) | *n.s.* |
| Sexuality | 0.23 (0.50) | 0.38 (0.45) | ** | -0.75 (3.50) | 0.31 (2.93) | *** | 0.30 (1.38) | 0.64 (1.21) | ** |
| Skin Tone | 0.49 (0.39) | 0.50 (0.38) | *n.s.* | -0.24 (1.52) | 0.14 (1.15) | *n.s.* | -0.12 (0.93) | 0.21 (0.80) | * |

*Note.* Positive values indicate a preference for the dominant over the stigmatized group. *n.s.* = not significant; * = $p < .05$; ** = $p < .01$; *** = $p < .001$, for the comparison between the choice and random conditions.

**Table S4**

*Means and Standard Deviations (in Parentheses) by Condition (Choice vs. Random) and Target (Age, Body Weight, Race, Sexuality and Skin Tone) for the Implicit Association Test, the Feeling Thermometer Difference Score, and the Relative Preference Item (Supplemental Study 1B)*

| | Implicit Association Test | | | Feeling Thermometer | | | Relative Preference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Choice | Random | Difference | Choice | Random | Difference | Choice | Random | Difference |
| Age | 0.40 (0.39) | 0.44 (0.37) | *n.s.* | -0.33 (1.89) | -0.23 (1.64) | *n.s.* | 0.20 (1.01) | 0.33 (1.06) | *n.s.* |
| Body Weight | 0.47 (0.39) | 0.47 (0.40) | *n.s.* | 0.46 (2.08) | 0.10 (2.01) | *n.s.* | 0.40 (1.04) | 0.36 (0.88) | *n.s.* |
| Race | 0.30 (0.44) | 0.27 (0.42) | *n.s.* | -0.38 (1.92) | -0.58 (2.08) | *n.s.* | -0.10 (0.99) | 0.01 (1.03) | *n.s.* |
| Sexuality | 0.17 (0.47) | 0.35 (0.46) | *** | -0.55 (3.39) | 0.04 (2.69) | ** | 0.11 (1.32) | 0.38 (1.06) | ** |
| Skin Tone | 0.48 (0.37) | 0.45 (0.41) | *n.s.* | 0.00 (0.88) | 0.13 (1.22) | *n.s.* | 0.00 (0.48) | 0.07 (0.59) | *n.s.* |

*Note.* Positive values indicate a preference for the dominant over the stigmatized group. *n.s.* = not significant; * = $p < .05$; ** = $p < .01$; *** = $p < .001$, for the comparison between the choice and random conditions.

## Robustness Checks I: Unweighted Models

In the main text, we discuss the limitations stemming from the use of unweighted models in much detail. As such, we do not interpret these models in the main text but include them in the present document for the sake of completeness (see Table S5 and Figure S8 below).

**Table S5**

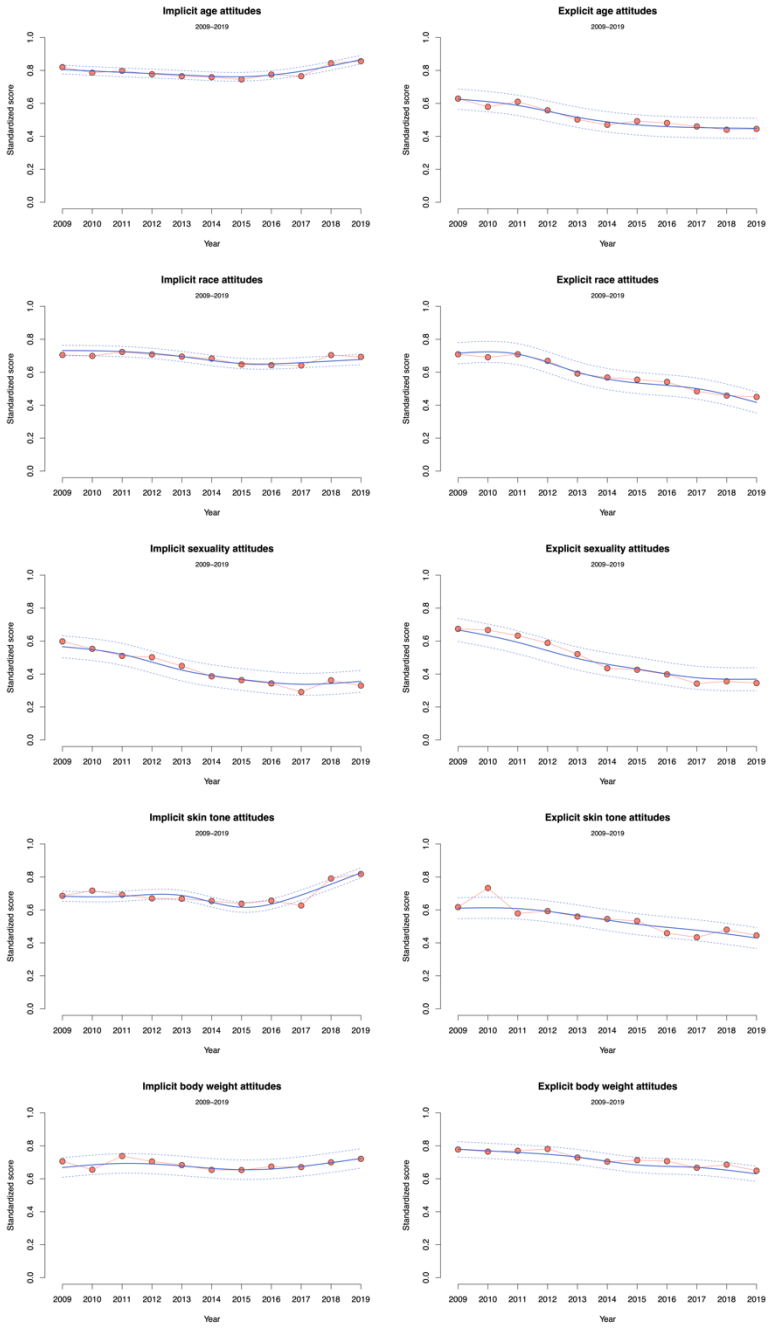
*Summary Statistics of Overall Trends by Task, from Fitted Unweighted Model Values*

| Task | Explicit attitudes | | | Implicit attitudes | | |
|---|---|---|---|---|---|---|
| | Start | End | % change | Start | End | % change |
| Age | 0.63 | 0.45 | -28.34 | 0.81 | 0.87 | 7.45 |
| Race | 0.72 | 0.42 | -41.78 | 0.73 | 0.68 | -7.39 |
| Sexuality | 0.67 | 0.37 | -44.75 | 0.57 | 0.36 | -37.61 |
| Skin tone | 0.61 | 0.43 | -29.67 | 0.68 | 0.83 | 20.86 |
| Body weight | 0.78 | 0.63 | -19.18 | 0.67 | 0.72 | 8.22 |

*Note.* All values are computed from standardized scores. The start and end values represent the first and last of 100 predicted values between 2009 and 2019, derived from the aggregate INLA models. Percent change values indicate the amount of change between the predicted start and end values.

**Figure S8**

*Overall Trends of Explicit and Implicit Attitudes across 33 Countries (Unweighted Models)*




*Note.* The solid blue lines indicate the fitted trajectories from the aggregate INLA model; the dashed blue lines indicate the 95% highest density intervals (HDIs). The red dots and red line show the raw annual means.

## Robustness Checks II: Refitting Models with Country Population as Model Weights

In the main text we fit all models using weights derived from the number of respondents within each year in each country. The inclusion of these weights ensures that each observation (i.e., the mean in each country-year) contributes to the final trend in proportion to how many respondents that observation is derived from, which helps increase the precision of the final estimate. For instance, the observation for the age task in Argentina in 2009 was based on data from 246 respondents; in 2019, it was based on data from 191 respondents. Thus, the weight that Argentina contributes to the final mean in 2009 will be slightly greater than the weight that Argentina contributes to the final mean in 2019. One limitation of this approach is that it overweights the contributions of those Western English-speaking countries that have large amounts of data in the Project Implicit International dataset (e.g., Canada, United Kingdom) rather than those countries that are actually the most populous around the world.

Here we take an alternative modeling approach: We instead weight the data according to actual (real-world) country populations, specifically how large that population is in relation to the other countries included in the dataset. For example, as of the most recent census, Argentina has a population of 46,556,542. The full sample across all 33 countries encompassed countries with a population of 4,343,542,431. Thus, the population of Argentina was about 1% of the total population; as such, it contributes to the final estimate with a weight of 0.01. We use the same real population weights for every year.

In line with the models reported in the main text, results from the reweighted models (see Table S6 and Figure S9) showed that (*a*) all explicit attitudes have decreased toward neutrality, with the fastest change in explicit sexuality attitudes; and (*b*) for implicit attitudes, sexuality attitudes decreased and skin tone attitudes increased in bias, while implicit age and race attitudes

were stable. However, unlike the primary models reported in the main text, which revealed stability, the models using population weights suggested that implicit body weight attitudes have also increased in bias over time.

Additionally, it is notable that these models had much larger error intervals than the primary models and did not always align with the observed raw data (see Figure S9). For instance, the raw data for explicit sexuality and skin tone attitudes were substantially lower than the fitted trend, suggesting that larger-population countries (such as China and) have stronger anti-gay or anti-dark-skinned attitudes than the countries that are overrepresented in the data (such as the UK and Canada). The larger error intervals are presumably also explained by the fact that this approach heavily overweights countries such as China and India, which have very large real-world populations (32.69% and 32.74%, respectively) but contribute only relatively modest numbers of participants to the dataset (2.63% and 1.02%, respectively). As such, we believe that the results of these models should be treated with caution.

**Table S6**

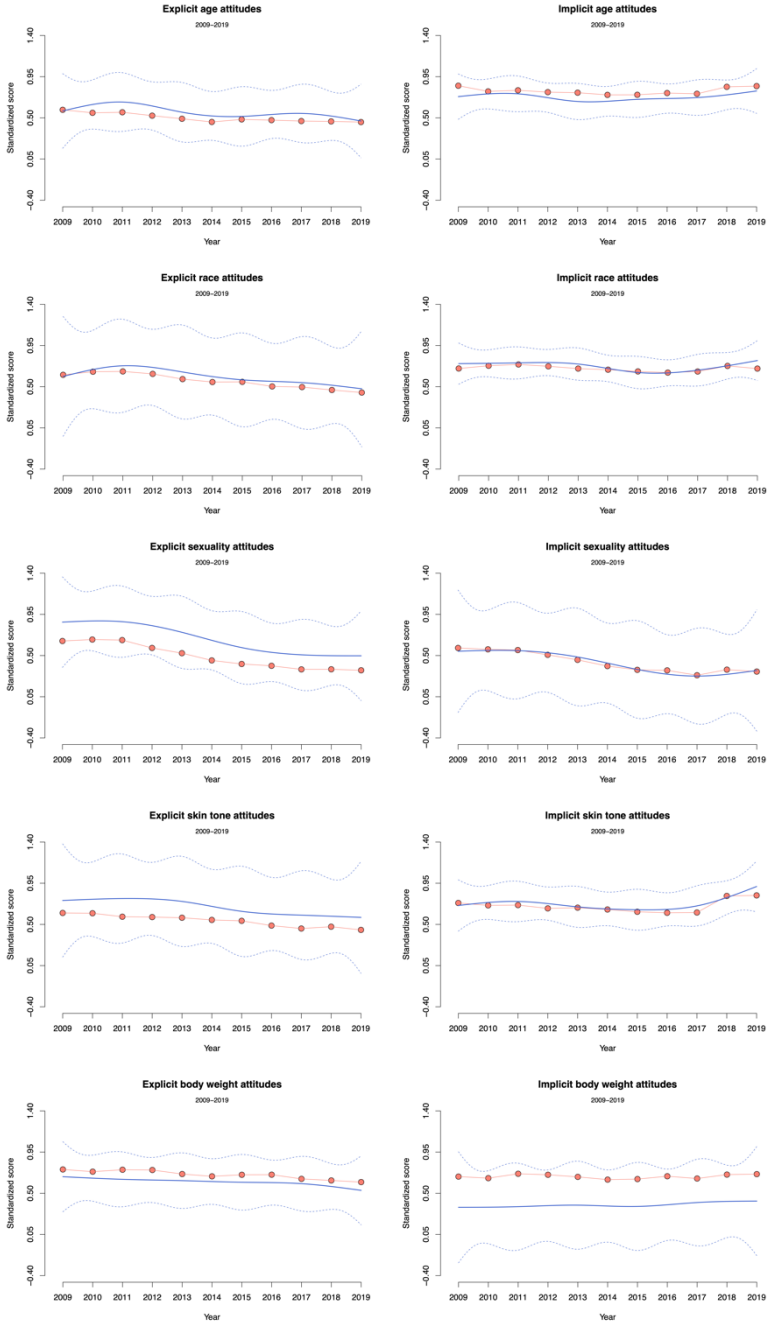
*Summary Statistics of Overall Trends by Task, from Fitted Model Values Using Population Weights*

| Task | Explicit attitudes | | | Implicit attitudes | | |
|---|---|---|---|---|---|---|
| | Start | End | % change | Start | End | % change |
| Age | 0.59 | 0.48 | -19.13 | 0.73 | 0.80 | 8.41 |
| Race | 0.67 | 0.52 | -22.36 | 0.76 | 0.79 | 4.45 |
| Sexuality | 0.89 | 0.51 | -42.69 | 0.55 | 0.34 | -38.21 |
| Skin tone | 0.78 | 0.59 | -24.42 | 0.71 | 0.91 | 29.49 |
| Body weight | 0.69 | 0.54 | -21.94 | 0.35 | 0.42 | 19.23 |

*Note.* All values are computed from standardized scores weighted by the proportional country population (i.e., the proportion of the full sample that comes from a given country). The start and end values represent the first and last of 100 predicted values between 2009 and 2019, derived from the aggregate INLA models. Percent change values indicate the amount of change between the predicted start and end values.

**Figure S9**

*Overall Trends of Explicit and Implicit Attitudes across 33 Countries Using Population Weights*




*Note.* The solid blue lines indicate the fitted trajectories from the aggregate INLA model; the dashed blue lines indicate the 95% highest density intervals (HDIs). The red dots and red line show the raw annual means, using weights for sample change.

## Robustness Checks III: Refitting Models to Citizens Only

In the main text we fit all models to respondents reporting residence within one of the 33 countries from the Project Implicit International websites. This is a relatively permissive decision given that recent residents of a country may not have sufficiently acculturated to that country to be representative of that country's culture. As such, here we refit the main models reported in the paper to those respondents reporting citizenship within those same countries to evaluate whether there are substantial differences in the trends of residents versus citizens. The results remained nearly identical to the models fit for country residents (see Table S7 and Figure S10): As in the main models, explicit attitudes decreased in bias for all five tasks, implicit attitudes decreased in bias only for the sexuality task and increased in bias for the skin tone task (with no appreciable change on the remaining tasks). The magnitudes of the mean and percentage decreases were also nearly identical across the two samples.

**Table S7**

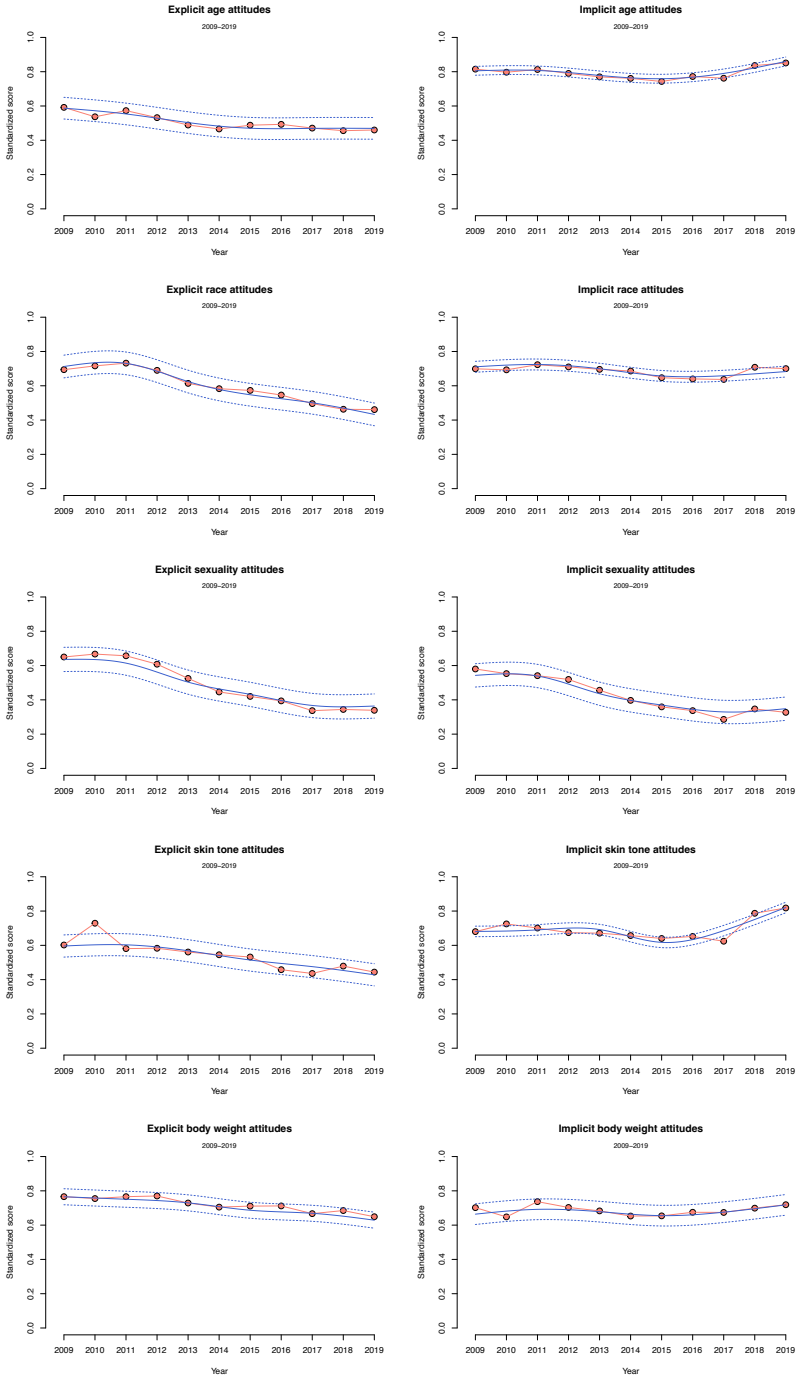
*Sample Characteristics for Citizens Only and Summary Statistics of Overall Trends by Task, from Fitted Model Estimates*

| Task | Sample size | | | Explicit attitudes | | | Implicit attitudes | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N$ | $N_{min}$ | $N_{max}$ | Start | End | % change | Start | End | % change |
| Age | 188,013 | 466 | 32,705 | 0.59 | 0.47 | -20.06 | 0.81 | 0.86 | 6.78 |
| Race | 507,032 | 597 | 238,755 | 0.71 | 0.43 | -39.23 | 0.71 | 0.68 | -3.90 |
| Sexuality | 307,586 | 755 | 38,323 | 0.64 | 0.36 | -42.76 | 0.54 | 0.35 | -35.80 |
| Skin tone | 176,488 | 461 | 27,705 | 0.60 | 0.43 | -28.17 | 0.68 | 0.82 | 20.49 |
| Body weight | 205,409 | 603 | 32,972 | 0.77 | 0.63 | -17.85 | 0.66 | 0.72 | 8.15 |

*Note.* All models are computed from standardized scores weighted by sample composition as well as by country sample size. The start and end values represent the first and last of 100 predicted values between 2009 and 2019, derived from the aggregate INLA models. Percent change values indicate the amount of change between the predicted start and end values. $N$ = total sample size; $N_{min}$ = smallest country-specific sample size; $N_{max}$ = largest country-specific sample size.

**Figure S10**

*Overall Trends of Explicit and Implicit Attitudes across 33 Countries (Citizens Only)*




*Note*. The solid blue lines indicate the fitted trajectories from the aggregate INLA model; the dashed blue lines indicate the 95% highest density intervals (HDIs). The red dots and red line show the raw annual means, using weights for sample change.

## Modeling by Sample Demographics I: Comparisons by Respondent Education (College-Educated vs. Non-College-Educated Respondents) for All Tasks

Given that Project Implicit data tends to overrepresent highly educated respondents relative to true population proportions, there is a possibility that the trends reported in the main paper were due exclusively (or at least primarily) to the liberalizing attitudes of highly educated demographic groups. To address this concern, we fit separate INLA models for respondents with and without the equivalent of college education in their countries. We used the same specifications as in the primary models, with inverse weights for country frequencies, and weighted yearly means according to demographic representation by gender, age, and ideology.

Across all tasks and for both explicit and implicit attitudes, we found similar trends regardless of respondent education level (see Table S8 and Figure S11). Specifically, for (*a*) explicit attitudes, we found parallel trends with virtually identical decreases across all tasks, with the exception of explicit skin tone attitudes where non-college-educated respondents changed descriptively faster than college-educated respondents. By contrast, for implicit attitudes we found that (*b*) only implicit sexuality attitudes showed clear decreases, with similar drops regardless of education; (*c*) implicit age, body weight, and race attitudes hardly changed at all, regardless of education; and (*d*) implicit skin tone attitudes showed curvilinear patterns, with recent increases since 2015, but especially among college-educated respondents.

These small differences notwithstanding, most respondents, regardless of education, have changed in the same direction and at similar rates across all attitudes, with correlations between temporal trajectories ranging from .79 to .99. As such, the conclusions regarding attitudinal trends over time are highly robust to respondent education.

**Table S8**

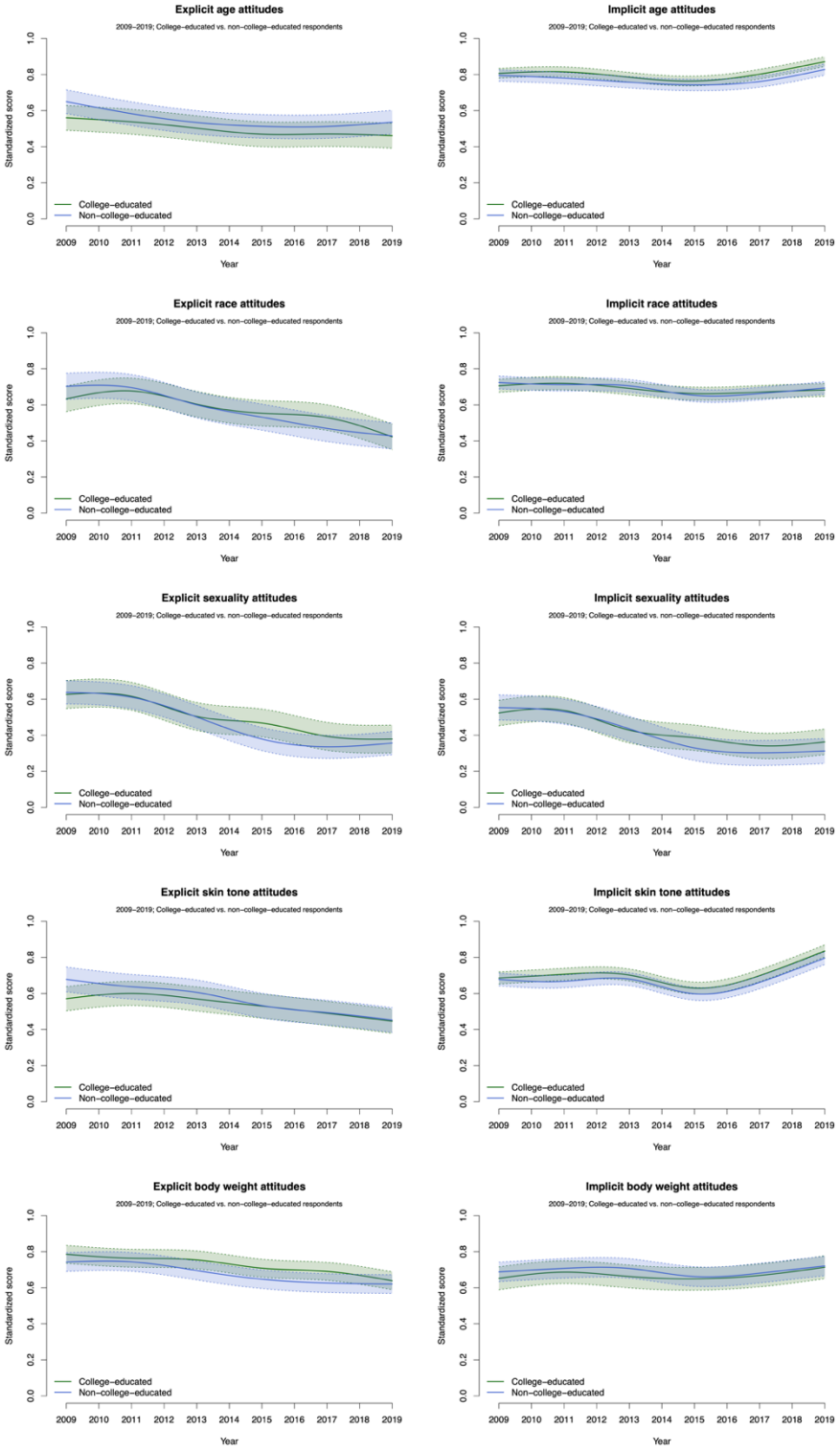
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Educational Subgroups across all Explicit and Implicit Attitudes (2009–2019)*

| Task | Participant Education | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Start | End | % change | Correlation | Start | End | % change | Correlation |
| Age | College | 0.58 | 0.43 | -26.04 | ρ = .81 | 0.84 | 0.88 | 4.56 | ρ = .93 |
| | No College | 0.62 | 0.49 | -21.35 | | 0.84 | 0.80 | -5.15 | |
| Race | College | 0.63 | 0.45 | -28.39 | ρ = .97 | 0.66 | 0.69 | 3.97 | ρ = .94 |
| | No College | 0.70 | 0.48 | -31.28 | | 0.75 | 0.70 | -6.47 | |
| Sexuality | College | 0.67 | 0.34 | -49.49 | ρ = .95 | 0.56 | 0.33 | -40.45 | ρ = .97 |
| | No College | 0.67 | 0.32 | -52.39 | | 0.62 | 0.32 | -48.97 | |
| Skin tone | College | 0.60 | 0.51 | -14.13 | ρ = .94 | 0.72 | 0.86 | 19.99 | ρ = .94 |
| | No College | 0.75 | 0.45 | -39.41 | | 0.74 | 0.77 | 5.00 | |
| Body weight | College | 0.76 | 0.78 | 2.17 | ρ = .99 | 0.66 | 0.66 | 0.19 | ρ = .79 |
| | No College | 0.77 | 0.77 | 0.75 | | 0.66 | 0.67 | 1.21 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S11**

*Comparisons of College-Educated and Non-College-Educated Respondents on Explicit Attitudes (Left Panes) and Implicit Attitudes (Right Panes)*



*Note.* College-educated respondents are shown in green and non-college-educated respondents in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Modeling by Sample Demographics II: Comparisons by Respondent Age (Younger vs. Older) for the Age Task

An exhaustive survey of all demographic differences in international patterns of attitude change is beyond the scope of the present work, which focuses on aggregate trends across tasks. As such, we focus on the specific demographic contrasts most relevant to each individual attitude task, starting with respondent age for the age task. Of note, due to the Project Implicit sample being overall quite young, even our older age group was only those above 35 (roughly corresponding to birth years of Millennials and older). This coarse age group was necessary to ensure adequate sample sizes across years and countries.

As shown in Table S9 and Figure S12, both age groups have decreased in bias consistently and at parallel rates on explicit age attitudes. This finding reinforces the conclusion that explicit attitudes change in response to domain-general pressures shared across tasks and demographic groups.

On implicit attitudes, relatively older respondents showed some recent increases in anti-old/pro-young implicit attitudes, whereas the relatively younger respondents showed no notable change. Given that even most of the older respondents may be conceptualizing the old target group as older than themselves (such as people above 70 or 80 years), the trends observed in this group may be indicative of outgroup, rather than ingroup, negativity. Perhaps, as these >35-year-olds look to economic decisions (e.g., purchasing a home, starting a family, progressing in their careers), they have become particularly frustrated with the older generations' "getting in the way" of their economic success. Indeed, these sentiments are reflected in the growing generational wars, at least in the United States (Francioli et al., 2023).

These relatively minor differences notwithstanding, we find that the patterns of change in both explicit and implicit attitudes were highly similar across younger and older respondents, with the respective time series correlated with each other at $\rho = .78$ for the former and $\rho = .87$ for the latter. As such, the present results reinforce the generalizability of the patterns reported in the main text across age groups.

**Table S9**

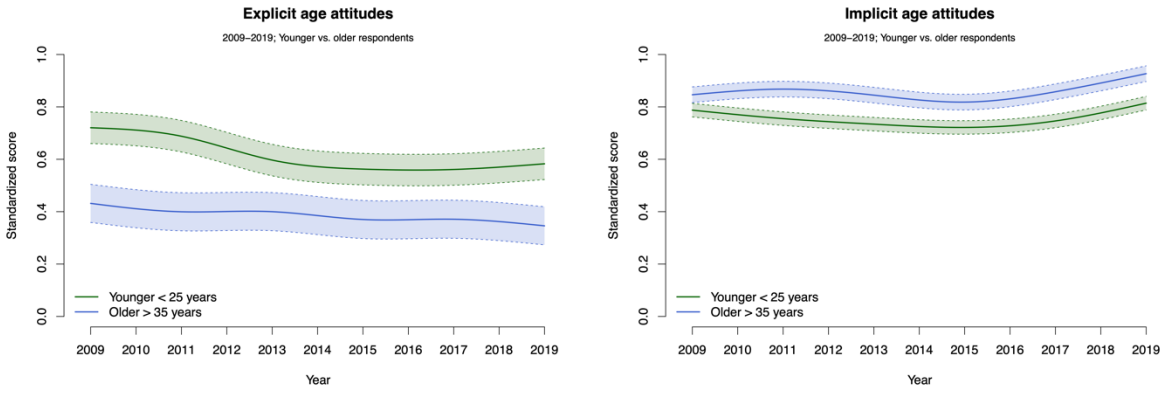
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Self-Reported Age Subgroups for Explicit and Implicit Age Attitudes (2009–2019)*

| Participant Age | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|
| | Start | End | % change | Correlation | Start | End | % change | Correlation |
| Younger (<25) | 0.70 | 0.56 | -19.43 | ρ = .78 | 0.80 | 0.79 | -0.96 | ρ = .87 |
| Older (>35) | 0.48 | 0.37 | -22.71 | | 0.80 | 0.91 | 14.18 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S12**

*Comparisons of Younger and Older Respondents on Explicit Age Attitudes (Left Pane) and Implicit Age Attitudes (Right Pane)*




*Note.* Younger respondents (below 25 years) are shown in green and older respondents (above 35 years) in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Modeling by Sample Demographics III: Comparisons by Respondent Race (White vs. Non-White) for the Race Task

In our previous work on long-term attitude trajectories in the United States (Charlesworth & Banaji, 2021), we found an effect of race such that White respondents changed faster than non-White respondents. This difference was, at least in part, explained by the fact that non-White participants already held strong and stable pro-Black/anti-White implicit and explicit attitudes.

Here, by contrast, we observed that non-White respondents (collapsing Black and other non-White racial identities to ensure adequate sample sizes given that we had only 8,722 Black participants in the data) did not already have pro-Black/anti-White attitudes (see Table S10 and Figure S13). Critically, on explicit attitudes, both White and non-White respondents decreased in bias in parallel, with similar raw change and exceedingly highly correlated trajectories ($\rho > .99$). However, because of the large baseline differences reflecting considerably stronger baseline bias among White respondents, the percent change was roughly half the rate among White relative to non-White respondents. These similar patterns of explicit attitude change across most respondents and most countries reinforce the conclusion that the explicit attitude changes are likely driven by domain-general processes that generalize across racial groups.

The two groups of participants exhibited highly correlated trajectories on implicit attitudes ($\rho = .76$), with some subtle differences. Specifically, on the one hand, White respondents exhibited curvilinear patterns similar to the overall patterns observed in implicit skin tone attitudes, with increases in bias both overall and especially since 2015. By contrast, non-White participants' implicit race attitudes decreased in bias over time, although to a moderate degree and with a potential uptick similar to that observed among White participants since 2017.

These results are notable in suggesting that White respondents in the United States may be relatively unique in their rapid drops in implicit pro-White/anti-Black attitudes, perhaps because the Black–White racial contrast is of central concern and discussion in the United States but less so in other countries around the world. Most important, White and non-White participants exhibited virtually identical trajectories on explicit race attitudes and highly similar (although subtly different) trajectories on implicit race attitudes. As such, these results attest to the robustness of the findings reported in the main paper.

**Table S10**

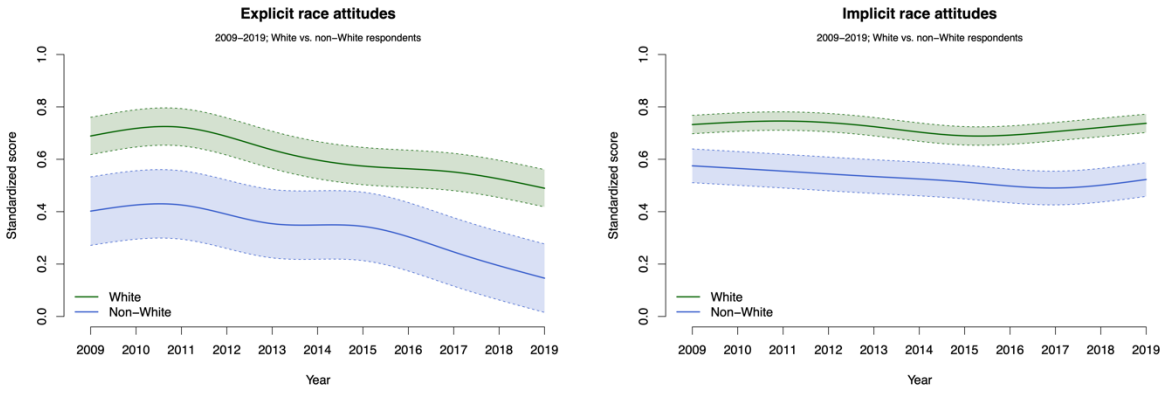
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Racial Subgroups for Explicit and Implicit Race Attitudes (2009–2019)*

| Participant | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|
| Race | Start | End | % change | Correlation | Start | End | % change | Correlation |
| White | 0.64 | 0.51 | -20.70 | $\rho > .99$ | 0.65 | 0.75 | 14.90 | $\rho = .76$ |
| Non-White | 0.37 | 0.16 | -57.32 | | 0.58 | 0.48 | -16.45 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S13**

*Comparisons of White and Non-White Respondents on Explicit Race Attitudes (Left Pane) and Implicit Race Attitudes (Right Pane)*




*Note.* White respondents are shown in green and non-White respondents (including both Black and other racial identities, such as East Asian, South Asian, and Aboriginal) in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Modeling by Sample Demographics IV: Comparisons by Respondent Sexual Orientation (Straight vs. Non-Straight) for the Sexuality Task

By far the largest baseline (intercept) differences across demographic groups were observed for straight versus non-straight (lesbian, gay, and bisexual) respondents on explicit and implicit sexuality attitudes (see Table S11 and Figure S14).

First, we found that non-straight respondents showed clear pro-gay/anti-straight bias on both explicit and implicit attitudes, and that these ingroup preferences tended to increase across the 11 years investigated. This is notable because it suggests that there is no inherent limit on ingroup preferences among stigmatized groups (e.g., non-straight respondents), and perhaps that the broader acceptance of sexual minorities internationally enables greater (albeit still weak) ingroup preference even among the stigmatized.

Second, and equally important, we also found notable drops in bias among straight respondents, mirroring the overall trends seen across countries. As such, the strong overall drop in implicit and explicit anti-gay/pro-straight attitudes is not merely due to differences in sample composition over time. Rather, even when we restrict analyses only to straight respondents, we find drops of 39% in bias on explicit attitudes and 30% on implicit attitudes.

Overall, these results suggest that, despite strong baseline differences, straight and non-straight respondents' attitudes have shifted in tandem, with temporal trajectories correlated with each other at .75 for explicit attitudes and .95 for implicit attitudes. As such, the patterns reported in the main paper cannot be explained by the overrepresentation of non-straight respondents on the sexuality task (see Supplemental Studies 1A–1B reported above), thus underscoring the robustness of our conclusions.

**Table S11**

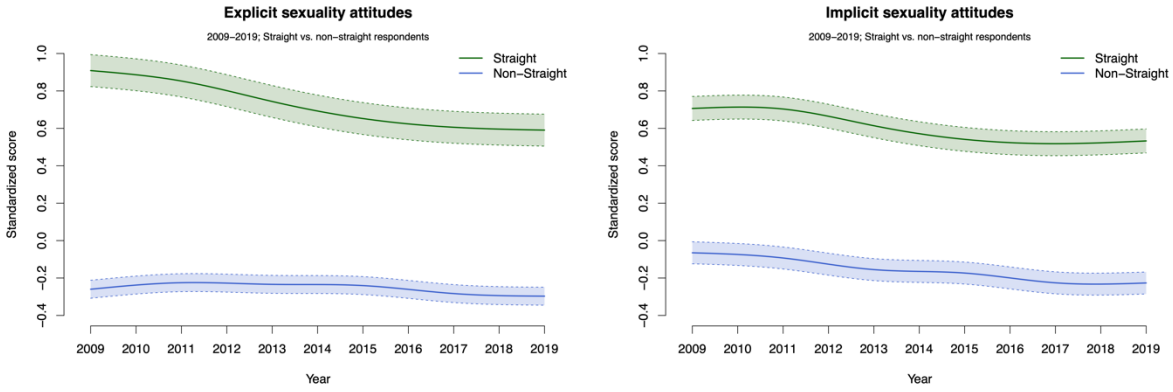
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Self-Reported Sexual Orientation Subgroups for Explicit and Implicit Sexuality Attitudes (2009–2019)*

| Participant Sexual Orientation | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|
| | Start | End | % change | Correlation | Start | End | % change | Correlation |
| Straight | 0.94 | 0.57 | -39.46 | ρ = .75 | 0.74 | 0.52 | -30.62 | ρ = .95 |
| Non-Straight | -0.23 | -0.28 | -22.80 | | -0.02 | -0.22 | -1126.44 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S14**

*Comparisons of Straight and Non-Straight Respondents on Explicit Sexuality Attitudes (Left Pane) and Implicit Sexuality Attitudes (Right Pane)*




*Note.* Straight respondents are shown in green and non-straight respondents (including gay, lesbian, bisexual and other non-straight identities) in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Modeling by Sample Demographics V: Comparisons by Respondent Skin Tone (Light-Skinned vs. Dark-Skinned) for the Skin Tone Task

Next, for the skin tone task, we turn to differences in attitude trends by respondents' self-reported skin tone, comparing those who report relatively lighter versus darker skin (collapsing those who report their skin tone as much, moderately, and slightly lighter or darker than others in their country).

The results (see Table S12 and Figure S15) showed baseline (intercept) differences in bias, with light-skinned respondents showing stronger pro-light-skin/anti-dark-skin preferences than dark-skinned respondents. However, critically, we observed no differences in the descriptive patterns or model-predicted trends over time. Specifically, both light-skinned and dark-skinned respondents exhibited clear drops in bias on explicit attitudes and the curvilinear patterns reported in the main text on implicit attitudes.

Intriguingly, this result suggests that the concerns driving the patterns in implicit attitudes — such as perceived threats of immigration — are affecting respondents similarly regardless of their own skin tone. That is, perceived threats of immigration may activate increased attention to subtle skin tone distinctions and may even make dark-skinned individuals more likely to make negative attributions toward their own group.

Future research is now poised to directly investigate these questions across countries with different immigration histories. For now, the key takeaway is that respondents' self-reported skin tone did not moderate the patterns of change in explicit or implicit skin tone attitudes, with highly similar trajectories on both the former ($\rho = .83$) and the latter ($\rho = .91$).

**Table S12**

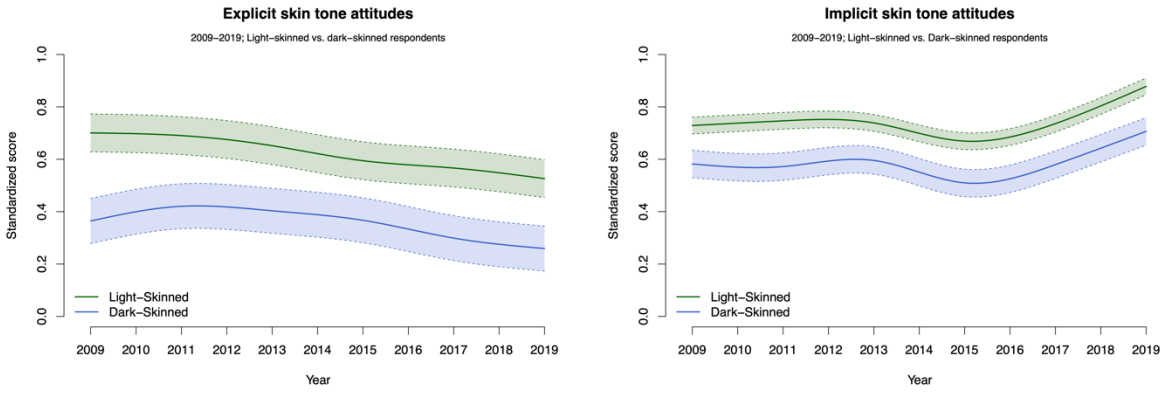
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Self-Reported Skin Tone Subgroups for Explicit and Implicit Skin Tone Attitudes (2009–2019)*

| Participant Skin Tone | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|
| | Start | End | % change | Correlation | Start | End | % change | Correlation |
| Light-Skinned | 0.72 | 0.53 | -26.40 | ρ = .83 | 0.76 | 0.87 | 14.56 | ρ = .91 |
| Dark-Skinned | 0.38 | 0.23 | -38.45 | | 0.57 | 0.69 | 20.38 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S15**

*Comparisons of Light-Skinned and Dark-Skinned Respondents on Explicit Skin Tone Attitudes (Left Pane) and Implicit Skin Tone Attitudes (Right Pane)*




*Note.* Light-skinned respondents are shown in green and dark-skinned respondents in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Modeling by Sample Demographics VI: Comparisons by Respondent Body Weight (Heavier vs. Thinner) for the Body Weight Task

Finally, turning to the body weight task, we compared the trends among respondents who self-reported as being relatively heavier or thinner (collapsing very, moderately, and slightly heavier or thinner to ensure adequate sample sizes across countries and time). Again, we found meaningful and expected baseline differences, with thinner respondents showing stronger pro-thin/anti-fat explicit and implicit attitudes across time. However, critically, despite these baseline differences, we also saw parallel trends across time, with both thinner and heavier groups showing slightly decreasing explicit body weight attitudes and relatively stable implicit body weight attitudes (see Table S13 and Figure S16).

Note that for explicit body weight attitudes, the raw data for the final year (2019) suggested an increase (see Table S13); however, this last mean was perhaps an anomaly due to relatively smaller sample sizes. In fact, the fitted values (see Figure S16) tend to show consistent decreases over time. The raw data for the penultimate year (2018) show an ending value of $M$ = 0.77 among thinner respondents, reflecting a drop of 17%, and an ending value of $M$ = 0.53 among heavier respondents, reflecting a drop of 10%. Critically, both explicit and implicit attitude trajectories were correlated with each other at $\rho = .93$, suggesting highly similar temporal trends among lighter and heavier respondents.

Thus, we interpret these data as showing consistent patterns with the overall results, regardless of respondents' own body weight, underscoring the robustness of the conclusions reported in the main paper.

**Table S13**

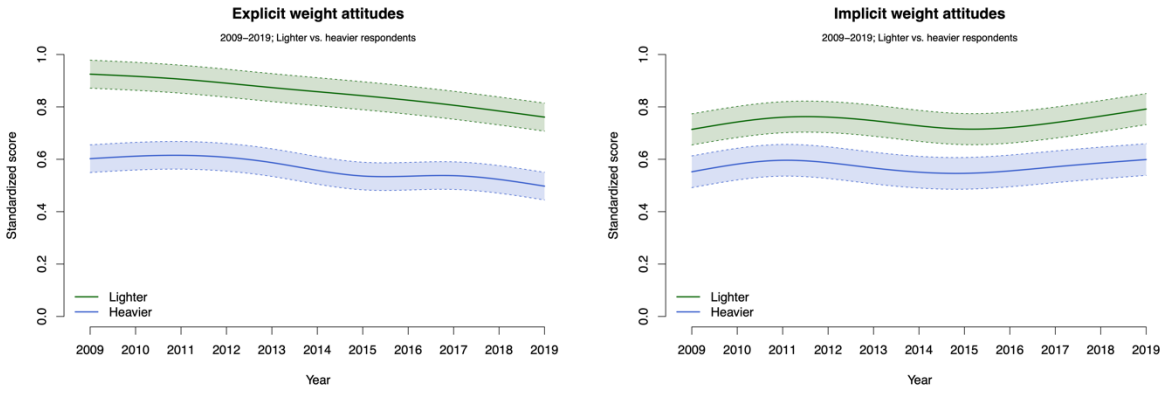
*Comparisons of Raw Starting and Ending Values, Percent Change, and Correlation of Fitted Trajectories by Self-Reported Body Weight Subgroups for Explicit and Implicit Body Weight Attitudes (2009–2019)*

| Participant | Explicit attitudes | | | | Implicit attitudes | | | |
|---|---|---|---|---|---|---|---|---|
| Body Weight | Start | End | % change | Correlation | Start | End | % change | Correlation |
| Thinner | 0.93 | 0.93 | -0.15 | ρ = .93 | 0.71 | 0.74 | 3.46 | ρ = .93 |
| Heavier | 0.59 | 0.58 | -2.02 | | 0.56 | 0.57 | 1.58 | |

*Note.* All values represent standardized scores. The start and end values represent the first and last of observed values between 2009 and 2019. Percent change values indicate the amount of change between the observed start and end values. The correlations are derived from model-predicted trajectories.

**Figure S16**

*Comparisons of Lighter and Heavier Respondents on Explicit Body Weight Attitudes (Left Pane) and Implicit Body Weight Attitudes (Right Pane)*




*Note.* Lighter respondents are shown in green and heavier respondents in blue, with solid lines corresponding to fitted trends and dashed lines corresponding to 95% credible intervals.

## Exploratory Analyses: Ecological Correlates of Implicit and Explicit Attitude Change

Previous research has sought to document the cultural and ecological correlates of the distribution of implicit and explicit attitudes across space within the United States (Hehman et al., 2021). Results from those studies showed that human features of the environment consistently predicted implicit and explicit attitudes. For example, the presence of more mental health providers was correlated with lower levels of bias. In contrast, ecological correlates produced inconsistent effects. For example, air pollution was correlated with explicit anti-Black and anti-Muslim attitudes but not with other attitudes. This null result for ecology is perhaps surprising given the large body of research suggesting that ecological features of the environment shape psychological outcomes, ranging from rice-based agriculture (Talhelm et al., 2014) to the hardness of the soil requiring ploughing (Alesina et al., 2013) to the propensity for disease threats (Schaller & Park, 2011) (for reviews of socio-ecological psychology, see Sng et al., 2018; Uskul & Oishi, 2020).

The preponderance of these ecological findings raises the possibility that the null result in previous research linking ecology and implicit/explicit attitudes could have been due to the limited focus only on the United States and on a single time point for both the ecological predictors and attitudes. Thanks to the Project Implicit International dataset, which contains data varying both across countries and over time, combined with new ecological databanks (Wormley et al., 2022), we now have an exciting opportunity to overcome these past limitations. Specifically, in the supplemental exploratory analyses reported here we newly consider how ecological variation across space (across 33 countries) and time (across 11 years) may map onto parallel spatial and

temporal variation in implicit and explicit attitudes. The results will provide new empirical insights into how international trends in attitude change may (or may not) be tied into broader features of cultural and ecological environments.

We note that these analyses are exploratory in that they aim to discover empirical patterns in a bottom-up manner from the available data (similar to Hehman et al., 2021). Although specific hypotheses could be tested in a top-down manner (e.g., that greater disease risk should correspond to higher levels of prejudice; Schaller & Park, 2011), here we opt for a data-driven approach for two reasons. First, the goal of the present project is not to test a particular theoretical perspective but rather to characterize (and start to understand) international trends in explicit and implicit social group attitudes. Second, discrepant findings emerging from past research (e.g., that disease risk sometimes has a null effect on prejudice; Inbar et al., 2016) also warrant a more inductive, empirically driven approach. As such, we offer theoretically driven interpretations of discovered empirical patterns with the hope that the nuanced results can be used to inform theorizing about not only whether but when ecology is linked to explicit and implicit attitude change around the world.

**Identifying and Preparing Ecological and Attitude Data**

One need only look at a recent publication in ecology to realize that there are thousands of potential ecological and population variables that could be studied as correlates of explicit and implicit attitude change, ranging from biomass and the ruggedness of the terrain to mortality and population rates. Because the aim of the current project is to provide an initial exploration of these ecological correlates, we constrain our sample space by using the nine time-stamped ecological variables from the Ecology–Culture Dataset (ECD; Wormley et al., 2022).

The ECD contains data across 220 countries for nine ecological variables (e.g., GDP and life expectancy) and 72 cultural variables (e.g., tightness–looseness, gender inequality, and personality). Notably, all cultural variables have data collected for only one timepoint and thus are not suited to understand the spatial and temporal patterning between covariates and attitudes. Given that the temporal trends are the primary focus of the present paper, we narrow our focus further to the nine variables from ECD that have at least some temporal variation (see Figure S17), although not all of these encompass the full 11 years of the available attitude data. Next, we also visualize the temporal and spatial variation in the implicit and explicit attitudes across the five key attitude tasks (see Figure S18).

Critically, when dealing with correlations between two timeseries, one must always consider the possibility of spurious correlations resulting simply from similar underlying trends. For instance, if two variables are both decreasing in their trends, then we will inevitably find a correlation between the two. However, such a correlation may very well not indicate any meaningful relationship in country-level temporal trends between the two variables. For example, if unemployment rate is decreasing worldwide and explicit attitudes are also decreasing worldwide, the correlation between the two time series will be positive. However, such a correlation need not imply that the two variables share any common variance at the country level.

As such, such analyses typically rely on detrended timeseries, which can be created, for example, by subtracting *t–1* from *t*, thereby removing shared variability in temporal patterns. Such detrended corrections help ensure that any observed correlation is due to the rises/falls in the timeseries and not merely to superficial similarity in trends. Thus, in our primary analyses, we normalize (z-score) both the ecological and attitude variables within each year. This analysis removes shared trends and considers country-level variability at each time point. As can be seen

in comparison with Figures S19 (ecological variables) and S20 (attitudes), which show the raw (unstandardized) values, this method eliminates overwhelming trends (e.g., falls in unemployment and explicit attitudes; see above) but retains spatial variation in the yearly rises/falls.

We then apply Representational Similarity Analysis (RSA; Kriegeskorte et al., 2008) across the normalized matrices of ecological variables and attitudes to measure the overlap in spatiotemporal variation. To do so, we vectorize each matrix such that each value in the vector is tagged as coming from a particular country in a particular year (e.g., Argentina in 2009) and then correlate the vectorized matrices with each other. Higher correlations indicate that the spatiotemporal variation in an ecological variable matches the spatiotemporal variation in an attitude variable.

**Ecological Correlates of Explicit Attitude Change**

The most important result emerging from the analysis of the ecological correlates of explicit attitudes is that the patterns of positive and negative correlations were highly similar to each other, regardless of the particular test (see Figure S21). For example, the detrended time series of population density was consistently positively correlated with the detrended time series of all explicit attitudes (including age, race, sexuality, skin tone and body weight; see top row), whereas the detrended time series in temperature was consistently negatively correlated with the detrended time series of all explicit attitudes (see bottom row). This result, in turn, suggests that the explicit attitude trajectories themselves tended to move in tandem with each other, thus dovetailing with the interpretation offered in the main text according to which explicit attitudes reflect shared, relatively domain-general sources of change.

Although we emphasize again that the results are exploratory and not designed to test specific ecological theories, it is helpful to offer some initial interpretation of the ecological variables that are correlated most consistently with the explicit attitude trajectories. We turn to these analyses next. Interested readers are invited to further consult Figure S21 for the results involving the remaining ecological variables (including rainfall, life expectancy, GDP, mortality, disease threat, income inequality, and unemployment). Offering a detailed interpretation of each of these variables is beyond the scope of the present project.

The most consistently positive ecological correlate of explicit attitude change was population density, such that an increase in population density within a year and country corresponded to an increase in explicit bias within that year and country. The relationship with population density was present for all explicit attitude topics, with Spearman's correlations ranging from $\rho = .20$ to $\rho = .40$ (all $ps < .001$). This result may be seen as surprising from certain theoretical perspectives. For example, according to contact theory (Allport, 1954; Pettigrew & Tropp, 2006), higher population density (and, correspondingly, more opportunities for intergroup contact) may correspond to less prejudice. (Although it should be noted that other preconditions of prejudice reduction, such as that contact should be on equal footing and condoned by authorities, may have been absent.)

In any case, this result is more straightforward to reconcile with perspectives according to which demographic changes can exacerbate conflict around limited resources (Goldstone, 2002). This interpretation is further bolstered by the fact that although the correlation between the detrended time series was consistently positive, the relationship was strongest for explicit race attitudes, $\rho = .34$, $p < .001$, and explicit skin tone attitudes, $\rho = .40$, $p < .001$, respectively. Given that the relationship is strongest for these two attitude targets, we can further speculate that the

negative effect of population density may be especially powerful if the population growth is particularly strong among stigmatized racial groups, which could activate specific threats around demographic change among members of dominant groups (Craig & Richeson, 2014a, 2014b, 2018).

On the other end of the spectrum, the most consistent negative ecological correlate of explicit attitude change was temperature, with significant negative correlations for four out of five attitude targets, ranging from $\rho = -.29$ to $\rho = -.32$ (all $ps < .001$). This result suggests that countries that experienced especially pronounced increases in temperature in any given year were most likely to exhibit changes in explicit attitudes toward less bias. The relationship for explicit sexuality attitudes was also descriptively negative ($\rho = -.08$), but did not reach significance, $p = .190$. These results are perhaps most easily explained by emerging perspectives linking seasonal and/or temperature patterns to prejudice (Hohm, O'Shea, et al., 2024; Hohm, Wormley, et al., 2024); however, the mechanism accounting for the link between these ecological variables and social group-based negativity remains to be seen.

Therefore, we also consider a third variable that could help explain the negative relationship between temperature change and explicit attitude change: industrialization. Specifically, countries exhibiting the greatest temperature rises and anomalies due to industrialization and, consequently, rising greenhouse gas emissions (e.g., Canada and the UK; Samborska, 2024) are also those countries that have been particularly liberalizing in their explicit attitudes across many groups. As such, whether temperature changes have a causal effect on explicit attitude change or, alternatively, both are caused by the same third variable remains to be seen.

More generally, these analyses are exploratory and, as such, the present conclusions should be more rigorously tested in future work, ideally using data showing greater variation

across countries and time and/or using more sophisticated analytic techniques that can account for the types of spurious correlation explained above. For now, the primary contribution from these supplemental analyses is that, regardless of the attitude target, explicit attitudes tend to exhibit similar ecological correlates, suggesting that they are all tied into similar domain-general features of the environment.

**Ecological Correlates of Implicit Attitude Change**

In contrast to explicit attitudes, which were characterized by consistent relationships with ecological variables regardless of the attitude targets, ecological correlates of implicit attitude change showed more test-specific variation. Specifically, drawing on the distinction between body-related and sociodemographic stigmas (Charlesworth, Sanjeev, et al., 2023), we observe different sets of correlates for change in implicit attitudes toward sociodemographic social categories (specifically, race, skin tone, and sexuality) vs. body-related social categories (specifically, age and body weight).

What is more, the clusters of sociodemographic and body-related implicit attitudes showed not only different but even opposite patterns of results. For example, change in population density was positively correlated with change in implicit race attitudes, $\rho = .22$, $p < .001$ and implicit skin-tone attitudes, $\rho = .14$, $p = .001$, but negatively correlated with change in implicit age attitudes, $\rho = -.19$, $p < .001$, and implicit body weight attitudes, $\rho = -.18$, $p = .001$. We also observed opposing relationships with temperature, which showed a negative relationship with implicit race attitudes, $\rho = -.25$, $p < .001$, but a positive relationship with implicit body weight attitudes, $\rho = .22$, $p < .001$.

Across these opposing relationships we note that the patterns for sociodemographic social categories tended to be similar to their explicit counterparts (e.g., positive correlations with population density but negative correlations with temperature), whereas ecological correlates of change diverged across implicit and explicit attitudes for body-related social categories. These results dovetail with findings from the social cognition literature suggesting that implicit–explicit correlations tend to be higher for sociodemographic than for body-related attitude targets (Nosek et al., 2007).

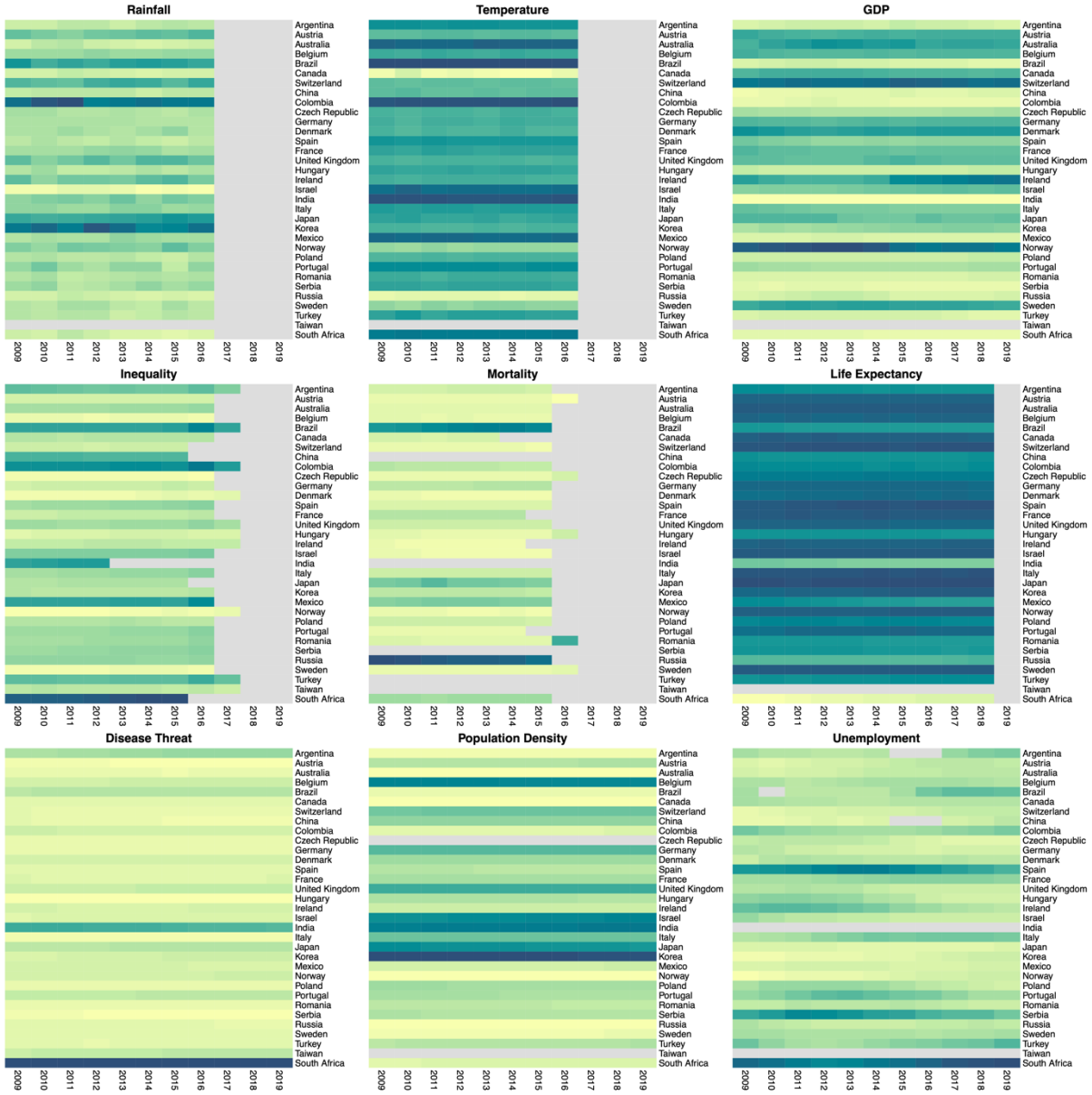
Perhaps the linkage between sociodemographic implicit attitudes and ecological factors may be driven by similar reasons as for their explicit counterparts (e.g., population density activating perceptions of demographic change and resource threat). In contrast, body-related implicit attitudes have opposing ecological correlates to most other (implicit and explicit) attitudes, which aligns with their generally opposing trajectories (i.e., they are the only ones that are largely stable over history; Charlesworth, Sanjeev, et al., 2023) . Furthermore, it is intriguing that body weight and age are the only attitude targets represented in the data whose spatiotemporal patterning is positively aligned with the spatiotemporal patterning of ecological disease threat or mortality. This result could reflect that body-related implicit attitudes are tied to physical concerns and pathogens, again aligning with past theorizing about the particular societal role of these stigmas (Park et al., 2007).

To summarize: Whereas the detrended explicit attitude trajectories showed consistent relationships with the detrended trajectories of ecological correlates, the relationship was more variable for implicit attitudes. Specifically, sociodemographic targets (race, skin tone, and sexuality) exhibited similar patterns to explicit attitudes, whereas body-related targets (age and body

weight) showed a pattern opposite of both explicit attitudes and implicit attitudes toward socio-demographic targets. Nevertheless, we emphasize once again the exploratory nature of these analyses and speculative nature of our conclusions. From our perspective, the results provide merely initial empirical suggestions that could be used as a starting point for deliberative tests of ecological hypotheses in future work.

**Figure S17**

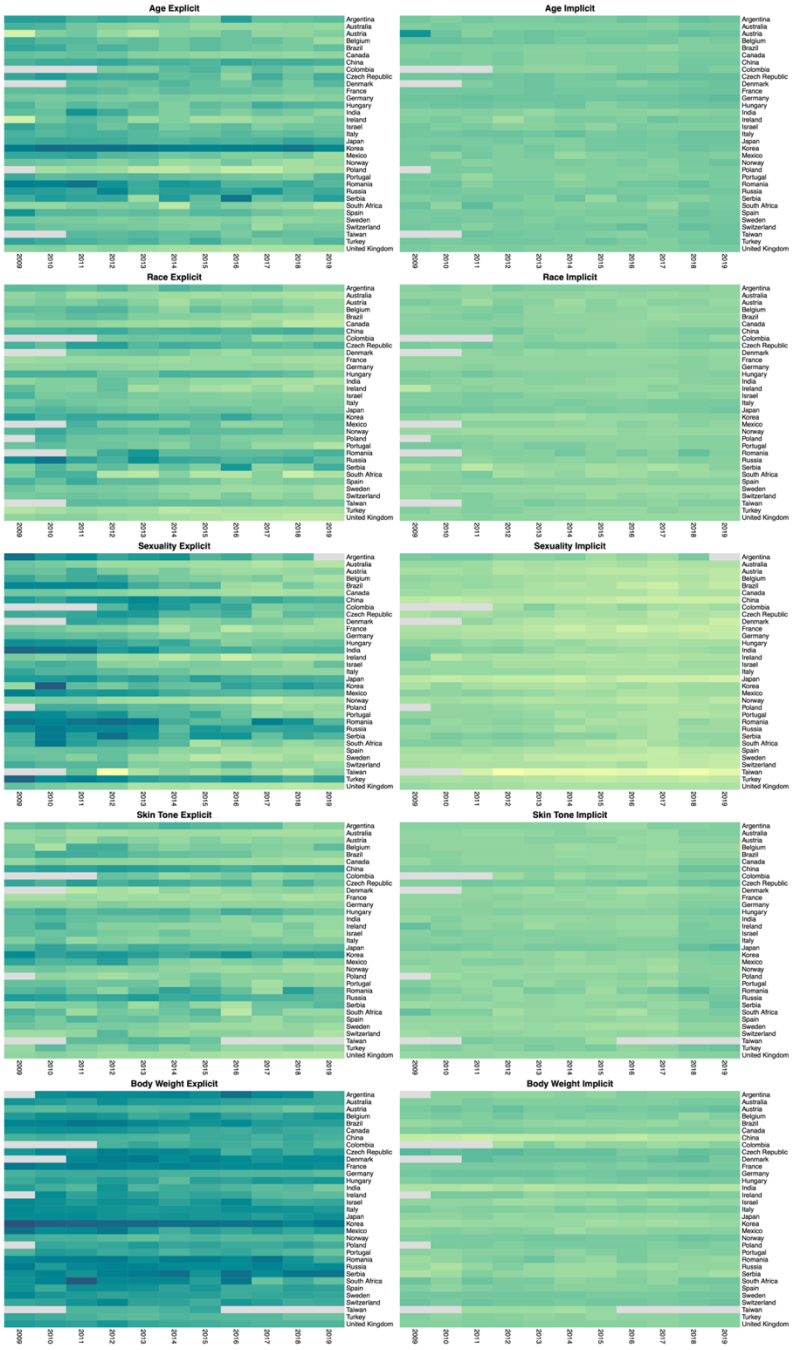
*Z-Scored (Within-Year Scaled) Variation in 9 Ecological Covariates across Time (X-Axis) And Countries (Y-Axis)*




*Note.* All values are visualized such that lighter colors indicate that the variable is (relatively) lower compared to the average variable value in that year (e.g., for the first column of the population density plot, the population density variable was normalized across all countries in 2009). For example, life expectancy was previously lower and has grown in magnitude over time across nearly all countries. By detrending the data, these spurious effects are removed.

**Figure S18**

*Z-Scored (Within-Year Scaled) Variation in Explicit and Implicit Attitudes across 5 Tasks (Columns), Time (X-Axis), And Countries (Y-Axis)*

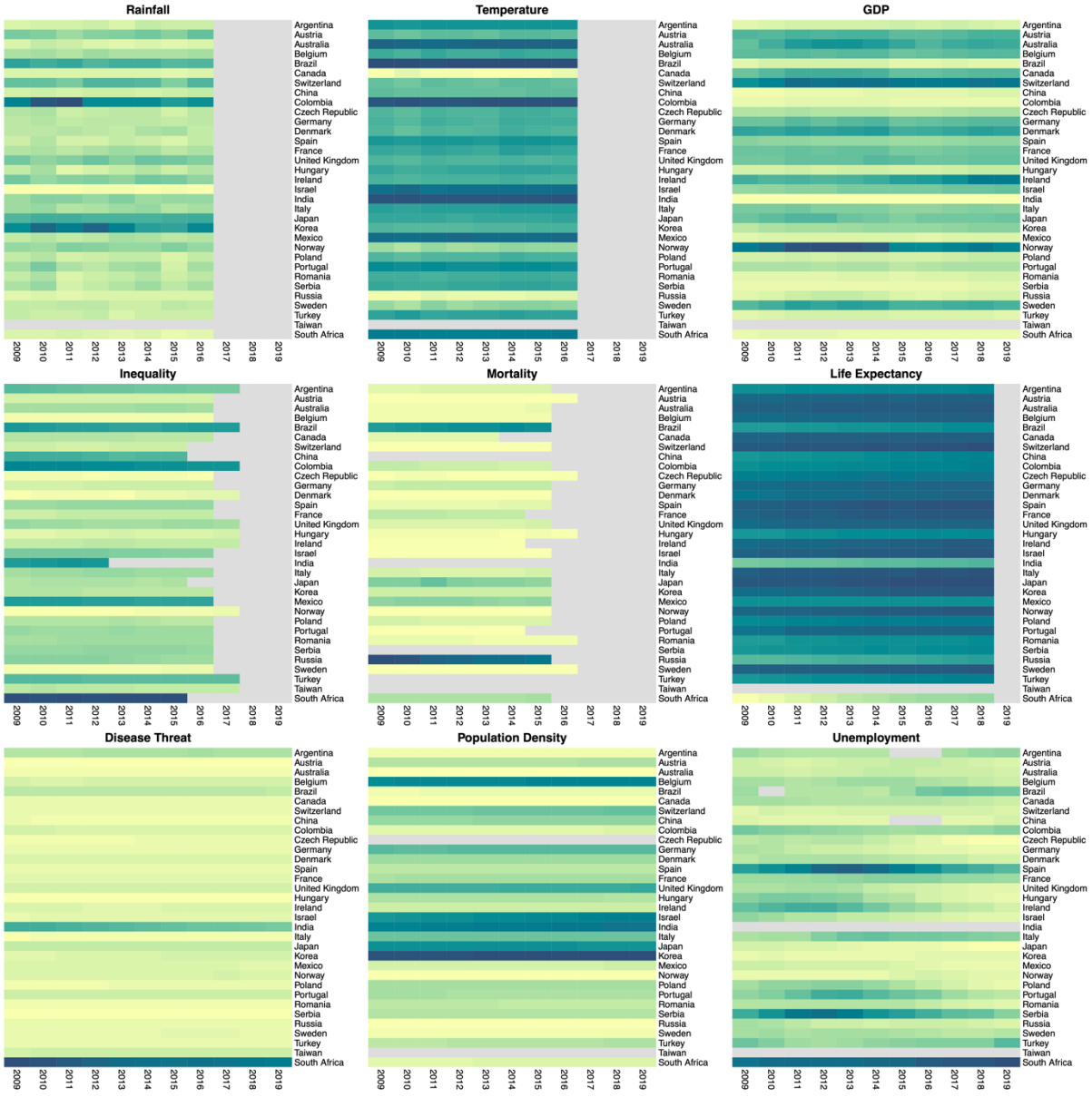



*Note.* All values are visualized such that lighter colors indicate that the attitude is (relatively) lower compared to the average attitude value for that year (across countries). For example, in

2009, implicit body weight attitudes were lower in Serbia and China than in other countries in that year and are therefore shaded with lighter colors.

**Figure S19**

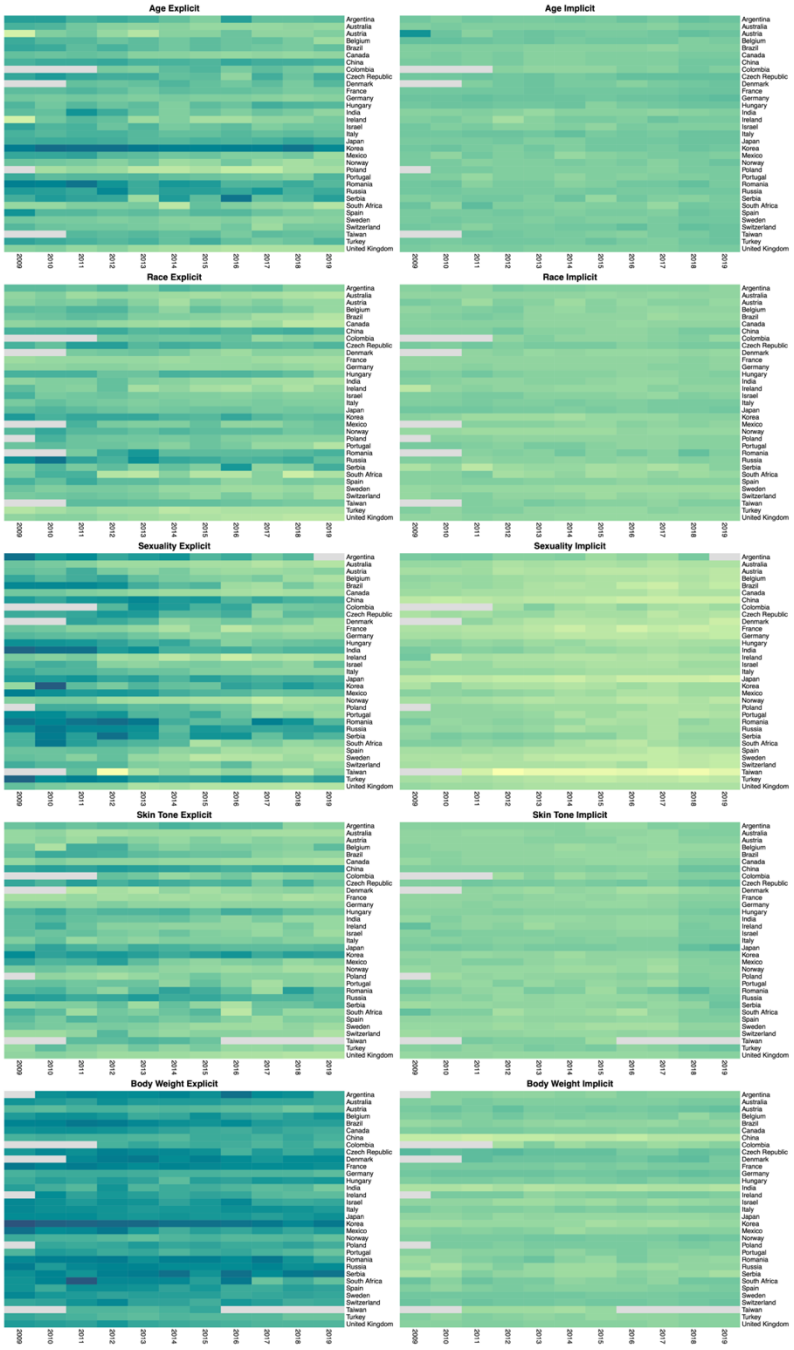
*Raw Variation in 9 Ecological Covariates across Time (X-Axis) And Countries (Y-Axis)*




*Note.* All values are visualized such that lighter colors indicate that the variable is (relatively) lower compared to the average variable value. For example, life expectancy was previously lower and has grown in magnitude over time across nearly all countries.

**Figure S20**

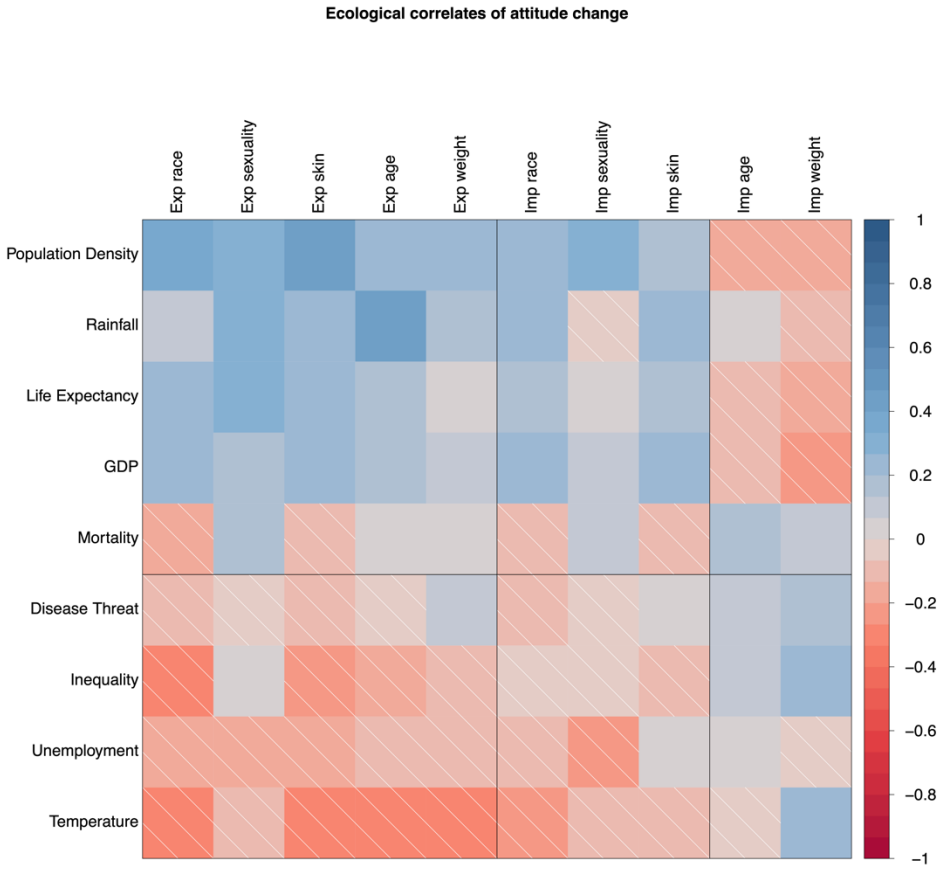
*Raw Variation in Explicit and Implicit Attitudes across 5 Tasks (Rows), Time (X-Axis), And Countries (Y-Axis)*



*Note.* All values are scaled such that lighter shades indicate that the attitude is (relatively) lower compared to the average attitude value for that task. For example, sexuality attitudes (both implicit and explicit) have consistently become lower in bias across all countries and thus have lighter colors in later years (more to the right of the plot). In contrast, implicit skin tone attitudes have exhibited unique increases in 2018 and 2019 across most countries, as indicated by relatively darker colors for those two years.

**Figure S21**

*Z-Scored (Within-Year Scaled) Representational Similarity in the Temporal and Spatial Variation of Attitudes (X-Axis) and Ecological Correlates (Y-Axis).*




*Note.* All values are scored such that blue colors indicate positive correlations (i.e., the trends are similar across the variables), while red colors indicate negative correlations (i.e., the trends are opposite across the variables). Darker colors (red or blue) indicate stronger correlations. Ecological predictors (y-axis) are ordered from those variables that have the most negative average correlation (on the bottom of the plot; i.e., temperature and unemployment), to those variables that have the most positive average correlation (on the top of the plot, i.e., population density and rainfall). Variables showing similar patterns are surrounded by black boxes.

## Supplemental References

Alesina, A., Giuliano, P. & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, *128*(2), 469–530. https://doi.org/10.1093/qje/qjt005

Allport, G. W. (1954). *The nature of prejudice*. Addison–Wesley.

Baumgartner, H. & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, *38*(2), 143–156. https://doi.org/10.1509/jmkr.38.2.143.18840

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., Navon, M., Rabinovich, Y., Lofaro, N. & Kurdi, B. (2023). The Project Implicit International Dataset: Measuring implicit and explicit social group attitudes and stereotypes across 34 countries (2009–2019). *Behavior Research Methods*, *55*(3), 1413–1440. https://doi.org/10.3758/s13428-022-01851-2

Charlesworth, T. E. S., Sanjeev, N., Hatzenbuehler, M. L. & Banaji, M. R. (2023). Identifying and predicting stereotype change in large language corpora: 72 groups, 115 years (1900–2015), and four text sources. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000354

Chen, C., Lee, S. & Stevenson, H. W. (1994). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, *6*(3), 170–175. https://doi.org/10.1111/j.1467-9280.1995.tb00327.x

Craig, M. A. & Richeson, J. A. (2014a). More diverse yet less tolerant? How the increasingly diverse racial landscape affects White Americans' racial attitudes. *Personality and Social Psychology Bulletin*, *40*(6), 750–761. https://doi.org/10.1177/0146167214524993

Craig, M. A. & Richeson, J. A. (2014b). On the precipice of a "majority-minority" America. *Psychological Science*, *25*(6), 1189–1197. https://doi.org/10.1177/0956797614527113

Craig, M. A. & Richeson, J. A. (2018). Hispanic population growth engenders conservative shift among non-Hispanic racial minorities. *Social Psychological and Personality Science*, *9*(4), 383–392. https://doi.org/10.1177/1948550617712029

Francioli, S. P., Danbold, F. & North, M. S. (2023). Millennials versus Boomers: An asymmetric pattern of realistic and symbolic threats drives intergenerational tensions in the United States. *Personality and Social Psychology Bulletin*, *50*(11), 1546–1562. https://doi.org/10.1177/01461672231164203

Goldstone, J. A. (2002). Population and security: How demographic change can lead to violent conflict. *Journal of International Affairs*, *56*(1), 3–21.

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

Hehman, E., Ofosu, E. K. & Calanchini, J. (2021). Using environmental features to maximize prediction of regional intergroup bias. *Social Psychological and Personality Science*, *12*(2), 156–164. https://doi.org/10.1177/1948550620909775

Hohm, I., O'Shea, B. A. & Schaller, M. (2024). Do moral values change with the seasons? *Proceedings of the National Academy of Sciences*, *121*(33), e2313428121. https://doi.org/10.1073/pnas.2313428121

Hohm, I., Wormley, A. S., Schaller, M. & Varnum, M. E. W. (2024). Homo temporus: Seasonal cycles as a fundamental source of variation in human psychology. *Perspectives on Psychological Science*, *19*(1), 151–172. https://doi.org/10.1177/17456916231178695

Hui, C. H. & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, *20*(3), 296–309. https://doi.org/10.1177/0022022189203004

Inbar, Y., Westgate, E. C., Pizarro, D. A. & Nosek, B. A. (2016). Can a naturally occurring pathogen threat change social attitudes? Evaluations of gay men and lesbians during the 2014 Ebola epidemic. *Social Psychological and Personality Science*, *7*(5), 420–427. https://doi.org/10.1177/1948550616639651

Johnson, T., Kulesa, P., Cho, Y. I. & Shavitt, S. (2005). The relation between culture and response styles. *Journal of Cross-Cultural Psychology*, *36*(2), 264–277. https://doi.org/10.1177/0022022104272905

Kriegeskorte, N., Mur, M. & Bandettini, P. A. (2008). Representational similarity analysis: Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4. https://doi.org/10.3389/neuro.06.004.2008

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G. & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88. https://doi.org/10.1080/10463280701489053

Park, J. H., Schaller, M. & Crandall, C. S. (2007). Pathogen-avoidance mechanisms and the stigmatization of obese people. *Evolution and Human Behavior*, *28*(6), 410–414. https://doi.org/10.1016/j.evolhumbehav.2007.05.008

Pettigrew, T. F. & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Plant, E. A. & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811–832. https://doi.org/10.1037//0022-3514.75.3.811

Samborska, V. (2024, March 18). *How much have temperatures risen in countries across the world?* OurWorldinData.Org. https://ourworldindata.org/temperature-anomaly

Schaller, M. & Park, J. H. (2011). The behavioral immune system (and why it matters). *Current Directions in Psychological Science*, *20*(2), 99–103. https://doi.org/10.1177/0963721411402596

Sng, O., Neuberg, S. L., Varnum, M. E. W. & Kenrick, D. T. (2018). The behavioral ecology of cultural psychological variation. *Psychological Review*, *125*(5), 714–743. https://doi.org/10.1037/rev0000104

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X. & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, *344*(6184), 603–608. https://doi.org/10.1126/science.1246850

Uskul, A. K. & Oishi, S. (2020). Editorial overview: What is socio-ecological psychology? *Current Opinion in Psychology*, *32*, 181–184. https://doi.org/10.1016/j.copsyc.2020.01.001

Wormley, A. S., Kwon, J. Y., Barlev, M. & Varnum, M. E. W. (2022). The Ecology–Culture Dataset: A new resource for investigating cultural variation. *Scientific Data*, *9*(1), 615. https://doi.org/10.1038/s41597-022-01738-z