PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES OPEN ACCESS

Film intervention increases empathic understanding of formerly incarcerated people and support for criminal justice reform

Marianne C. Reddan[a,b,1], Sydney B. Garcia[a,c], Golijeh Golarai[d], Jennifer L. Eberhardt[a,d,e,1], and Jamil Zaki[a,1]

Affiliations are included on p. 8.

Contributed by Jennifer L. Eberhardt; received December 27, 2023; accepted August 16, 2024; reviewed by Mina Cikara and Abigail Marsh

Nuanced portrayals of stigmatized groups in media have been shown to reduce prejudice. In an online experiment (N = 749), we tested whether a feature film depicting incarcerated peoples' experiences in the criminal justice system can increase a) empathic accuracy and compassion toward people who have been incarcerated and b) support for criminal justice reform. We measured baseline empathic accuracy via a well-validated task, where participants infer the emotions of people sharing stories about difficult life events. All storytellers were formerly incarcerated and students. However, in half the videos we labeled them as "formerly incarcerated" and in the remaining half as "college student." We then surveyed people's baseline attitudes toward criminal justice reform. Next, we assigned participants to watch one of three films. The intervention film chronicled the true stories of Black men on death row. Two docudramas of similar length served as control films. Finally, participants completed the empathic accuracy task and survey again and were given the opportunity to sign a petition. Compared to those who watched a control film, participants who watched the intervention film more accurately inferred the emotions of storytellers labeled "formerly incarcerated," and increased their support for criminal justice reform. These effects held for conservative and liberal participants alike. However, the film had no effect on feelings of compassion. Together, these results demonstrate the power of narrative interventions to not only increase empathic accuracy for members of a severely stigmatized group, but to increase support for reforms designed to improve their lives.

empathy | storytelling | narrative interventions | incarceration | criminal justice reform

Significance

Shifting people's responses to members of stigmatized groups can be difficult. Yet, after watching a narrative film about incarceration, we found that a nationally representative sample of study participants in the United States—a country that imprisons a greater share of its population than any other industrialized nation in the world—became more accurate at inferring the emotions of formerly incarcerated people. The film also boosted participants' desire to see sweeping reforms within the criminal justice system itself. These effects were observed regardless of participants' self-described political orientation. Building upon prior work demonstrating the power of storytelling, our findings show that sharing personal narratives can increase connection with highly stigmatized groups.

In 1986, Walter McMillian, a Black 45-y-old logger living in Alabama, was arrested for murder. He was innocent—he was miles away at a family gathering at the time the crime occurred. Yet, he was convicted based on false eyewitness testimony. He spent six years on death row before a lawyer, Bryan Stevenson, was able to overturn his conviction (1).

Walter's story is emblematic of a host of systemic problems within the criminal justice system in the United States. The US has the highest incarceration rate in the world–nearly 2 million Americans are in a prison or jail right now (2). Exposure to violence and isolation puts imprisoned people at risk of mental health conditions that persist well after release (3, 4). Employment prospects are limited for previously incarcerated people (5, 6), and the negative consequences of incarceration reverberate into communities for generations (7). Indeed, punitive policies and the culture of stigmatization that surrounds incarceration may contribute to the high rate of recidivism in the US (8). Meanwhile, there is substantial evidence for the positive impact of rehabilitative policies (9, 10).

In view of this evidence, many experts have called for criminal justice reforms that promote rehabilitation over retribution. However, public support for such reforms often falls short—incarcerated people in the US are typically viewed as an "extreme outgroup" and they are frequently dehumanized (11, 12). Such negative public perceptions pose a significant challenge to reform (13).

Could exposure to stories like Walter's influence the way Americans relate to people who have been incarcerated? Research demonstrates that increasing empathy toward others—including members of stigmatized groups—increases people's desire to help (14–19). One avenue to increase empathy is through immersive narratives which provide access to other people's perspectives. Narrative interventions can reduce prejudice and promote positive attitudes toward members of stigmatized groups (20–22), which can then generalize to the group as a whole (23–25). For example, a previous study found that sharing narratives about a stigmatized group through live theater (26) increased self-reported feelings of empathy toward the stigmatized groups being portrayed. Empathy interventions can, however, backfire (27). That is, they can unintentionally promote the status quo by making the target of empathy appear as less capable and in need of sympathy, rather than

Author contributions: M.C.R., G.G., J.L.E., and J.Z. designed research; M.C.R., S.B.G., J.L.E., and J.Z. performed research; M.C.R. and S.B.G. analyzed data; J.L.E. and J.Z. directed analysis; and M.C.R., S.B.G., G.G., J.L.E., and J.Z. wrote the paper.

Reviewers: M.C., Harvard University; and A.M., Georgetown University.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: mcreddan@stanford.edu, jleberhardt@stanford.edu, or jzaki@stanford.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2322819121/-/DCSupplemental.

Published October 21, 2024.

in need of material resources or power (28). Narrative interventions with lasting beneficial effects alter values and beliefs instead of merely pulling on heartstrings (see ref. 29 for a review). For example, a field experiment following the Rwandan genocide found that participants who listened to a radio soap opera that highlighted the value of reconciliation reported higher empathy for Rwandan prisoners and genocide survivors, compared to participants who listened to a control soap opera (30).

By extension, these findings suggest that nuanced narratives about incarcerated people in popular media may enhance empathy for incarcerated people in the US. However, it is unknown if self-reported feelings of empathy for stigmatized groups in prior studies reflect real changes in empathic cognition. Changes in self-reported empathy may, instead, reflect motivation to respond in a manner that is socially desirable. Indeed, actually increasing peoples' empathy toward a stigmatized outgroup may be challenging. Empathy requires cognitive effort (31–34), and individuals may be inclined to resist extending their cognitive resources to empathize with members of outgroups (35). The fact that incarcerated people are typically viewed as a dangerous outgroup may lead others to avoid exerting the cognitive effort that empathic inference and perspective-taking requires (13).

Here, we test the impact of a narrative intervention—a Hollywood docudrama—by directly measuring empathic responses to previously incarcerated individuals. We expect a professionally produced Hollywood docudrama to be effective, first, because it offers nuanced portrayals of incarcerated people, and individualized information about outgroup members has been shown to weaken stereotypes and increase the overlap between perceptions of self and others (36, 37). Second, such films are emotionally evocative and designed to captivate public attention. These characteristics provide an ideal opportunity for "narrative transportation"(38) which could increase viewers' ability to empathize with stigmatized outgroup members and could even change viewers' attitudes and behaviors. Still, scientific evidence of the impact of docudramas is lacking, as Hollywood films are rarely systematically evaluated by randomized control trials. Therefore, it remains unknown how impactful such films are in changing viewer empathy toward individual members of outgroups or their opinions and behaviors toward those outgroups.

We hypothesized that a full-length feature film about the case of Walter McMillian would increase empathy for formerly incarcerated people and support for criminal justice reform. Specifically, we predicted that the film would enhance viewers' ability to infer the emotions of people who have been incarcerated in real life—not just the actors who play them on screen—and that this enhanced empathy would coincide with greater support for policies that offer assistance to incarcerated and formerly incarcerated people.

To test this, we used a mixed longitudinal design where we measured a) participants' ability to infer the emotions of formerly incarcerated people, and b) their attitudes toward criminal justice reform both before and after watching either the intervention film or one of two equally engaging control films with "underdog" narratives that did not address incarceration. We measured empathic accuracy, a component of empathy that draws from cognitive reasoning, by using a well-validated task (39, 40). To construct this task, we collected videos of formerly incarcerated people describing emotional events in their lives. After being recorded, these storytellers provided moment-by-moment ratings of how they felt as they relayed their story. Online participants were later asked to view these videos and rate what they thought the storyteller was feeling at each moment using the same scale, which ranged from 0 (negative emotion) to 100 (positive emotion; *SI Appendix*, Fig. S2). The correspondence between a viewer's inference and a formerly incarcerated storyteller's self-report served as a measure of "empathic accuracy," or agreement between two people about what one of them feels. Each participant performed the empathic accuracy task twice, both before and after watching the film they were assigned. To further test the specificity of our intervention, the videos featured a label that indicated the storyteller was either "formerly incarcerated" or a "college student." All storytellers, in real life, had been incarcerated and were (or had been) enrolled in a college-level educational program, so both labels were accurate. Counterbalancing labels across participants assured that the labels, and not features of storytellers themselves, accounted for any effects we observed. Following each rating task, we measured self-reported feelings of compassion, which is the motivation to improve another's well-being, for storytellers on a scale from 0 (not at all) to 100 (very much). We predicted that watching a narrative film intervention about carceral injustice would selectively improve participants' empathic accuracy and compassion for formerly incarcerated storytellers relative to students.

We recruited a nationally representative sample of the US with regard to race, gender, subjective socioeconomic status, and political ideology (*SI Appendix*, Fig. S3 and Table S1) via a panel recruitment service. Participants were randomly stratified based on the aforementioned demographic variables to the intervention and control conditions. Participants completed the entire study alone on their personal computers at home. Once they began the experiment, they had three days to complete all components across two visits (*SI Appendix*, Figs. S1 and S2). This design allows time to vary between watching the film and completing the postfilm measures, which helps to mitigate demand effects on changes in self-report measures. A total of 749 participants completed part or all of the two visits and met our preregistered standards for data quality control (i.e., confirmed attention to the assigned film).

Results

Film Intervention Improved Empathic Understanding of Formerly Incarcerated Men's Emotions. We tested whether participants who watched the intervention film (*Just Mercy*; Warner Bros. 2019; N = 327) improved in their ability to infer the emotions of formerly incarcerated men relative to participants who watched the control films (N = 382). We assessed inference accuracy by calculating the Rms error (RMSE) between the participants' ratings and the storyteller's self-ratings during the empathic accuracy task.

Consistent with our preregistered hypothesis, we found a significant three-way interaction among visit (pre- vs. postfilm), film (intervention vs. control films), and storyteller label ("formerly incarcerated" vs. "college student") on empathic inference accuracy (measured via RMSE), in a linear mixed effects (LME) model that treated participant intercepts as random effects and visit, film, and label as fixed effects [$t(2068.38) = -2.08$, $P = 0.04$, $\beta = -0.65$, STE = 0.31; *SI Appendix*, Table S5]. This interaction held—though marginally—[$t(1859.54) = -1.96$, $P = 0.05$, $\beta = -0.64$, STE = 0.33; *SI Appendix*, Table S6] when controlling for participant race, gender, socioeconomic status, and political ideology. This suggests that the effect of the intervention on empathic accuracy cannot be explained by participants' social and political identities.

Planned pairwise *t* tests confirmed that a participant's ability to infer the emotions of formerly incarcerated storytellers significantly increased after watching *Just Mercy*, both within participants who viewed *Just Mercy* (post vs. pre: paired $t(318) = 4.10$,

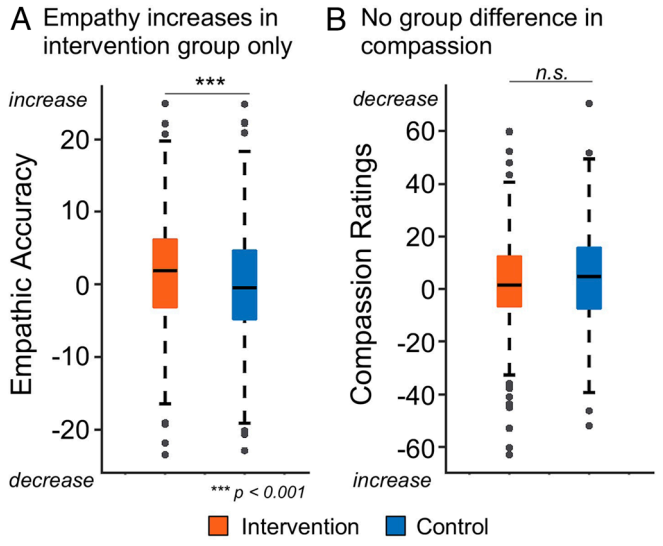

Fig. 1. Changes in empathic accuracy and compassion toward formerly incarcerated storytellers after film intervention. (*A*) Empathic accuracy increases in the intervention group only. Participants in the intervention group (N = 327), relative to the control group (N = 382), demonstrated an increase in their ability to accurately infer the feelings of formerly incarcerated people during the empathic accuracy task after watching the film ($P < 0.001$). Empathic accuracy was measured by taking the RMSE between participant inference ratings and the storyteller's self-ratings. Plotted are average (post–pre) change scores in RMSE for video trails where storytellers were labeled "formerly incarcerated." For plotting purposes, we inverted the RMSE change scores so that positive values indicate greater accuracy. (*B*) No change in compassion. There are no group differences in the change in compassion for formerly incarcerated storytellers. Plotted are average (post–pre) change scores in compassion for video trials where storytellers were labeled "formerly incarcerated." Bars indicate the two groups (intervention and control). All data are represented in box plots where the median is a black line and the upper and lower "whiskers" represent the bounds of the quartiles.

$P < 0.001$, Cohen's $d = 0.25$) and between the *Just Mercy* and control groups [$t(694) = 4.01$, $P < 0.001$, Cohen's $d = 0.31$; Fig. 1*A*]. This post-intervention effect was specific to storytellers labeled "formerly incarcerated" relative to "college student" [paired $t(316) = 2.83$, $P = 0.005$, Cohen's $d = 0.16$; *SI Appendix*, Fig. S5*A*].* Importantly, the primary three-way interaction among visit (pre- vs. postfilm), film (intervention vs. control films), and storyteller label ("formerly incarcerated" vs. "college student") on empathic accuracy holds when controlling for the race of the storyteller [$t(1076) = 2.31$, $P = 0.02$, $\beta = 0.60$, STE = 0.26] and the racial congruency between the participant and the storyteller [$t(1076) = 2.46$, $P = 0.01$, $\beta = 0.66$, STE = 0.27; *SI Appendix*, Tables S15 and S16 for details].

To summarize, after watching the film *Just Mercy*, participants demonstrated enhanced ability to infer the emotions of storytellers labeled "formerly incarcerated" relative to those labeled "college student." Participants' preexisting social and political identities had little influence on these effects.

Film Intervention had no Effect on Compassion for Formerly Incarcerated Men. We tested whether participants in the intervention condition (N = 327), relative to those in the control condition (N = 382), changed in their self-reported feelings of compassion for the formerly incarcerated men featured in the empathic accuracy task. Contrary to our expectations, we found no effect of the intervention on self-reported feelings of compassion for formerly incarcerated men. Using an LME model that treated participant intercepts as random effects, we found no interaction of visit, film, or label when predicting compassion toward storytellers [$t(2080.94) = 0.31$, $P = 0.76$], nor did we find a main effect of film or label (Fig. 1*B* and *SI Appendix*, Table S7). However, there was a significant main effect of visit, where compassion decreased, overall, during visit two [$t(2079) = -6.71$, $P < 0.001$], though self-reports of compassion were generally quite high overall (*SI Appendix*, Fig. S4).

Film Intervention Increased Support for Criminal Justice Reform. We tested whether participants who watched the intervention film (N = 344) increased their support for criminal justice reform relative to participants in the control conditions (N = 405). To measure support for criminal justice reform, we followed a preregistered plan to create an index of support (from 0 to 100) that averaged across 20 questions. The questions assessed broad support, such as "*Though the United States prison system has its problems, it is ultimately a fair system that should not be changed,*" and more specific policy preferences such as "*We should use our tax money to fund educational programs for those in prison.*" *SI Appendix*, Table S8 for individual items and their statistics.

Using a LME model with participant intercepts treated as random effects, we found a significant interaction between visit (pre vs. post) and film (intervention vs. control) on support for reform [$t(676.93) = -8.41$, $P < 0.0001$; *SI Appendix*, Table S9 for full model]. That is, after watching *Just Mercy* participants were more supportive of criminal justice reform, relative to participants who watched the control films (two-tailed independent samples $t(747) = 3.49$, $P < 0.001$, STD = 17.05, CI = [1.85, 6.76], Cohen's $d = 0.25$; Fig. 2*A*). This interaction remained significant when controlling for participant race, gender, socioeconomic status, and political ideology, indicating that the effectiveness of the intervention is not explained by participant social and political identities [$t(669) = -8.41$, $P < 0.0001$; *SI Appendix*, Table S10] despite significant individual differences in support for reform by political ideology prior to watching the intervention film [$t(650) = 6.10$, $P < 0.001$; *SI Appendix*, Table S11].

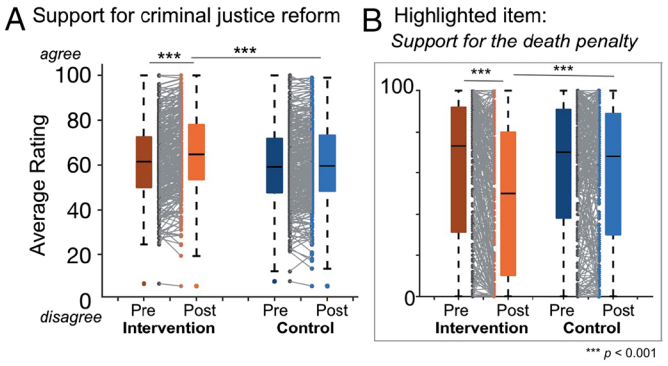

Fig. 2. Support for criminal justice reform increases after movie intervention. (*A*) Support for criminal justice reform. Twenty criminal-justice-oriented "attitudes & beliefs" questions were combined into a single metric, per person, so that high values indicated support for progressive criminal justice reform and low values indicated opposition. Participants selectively increase their support for criminal justice reform after watching the intervention film *Just Mercy* (N = 344), relative to control (N = 405; *SI Appendix*, Table S9 for statistics). (*B*) Highlighted item: Support for the death penalty. The film intervention specifically addresses the death penalty, therefore, the effect of the film on this specific individual item is highlighted. The intervention film selectively decreased support for the death penalty ($d = -0.34$; *SI Appendix*, Table S8 for statistics). Data are represented in box plots where the median is a black line and the upper and lower "whiskers" represent the bounds of the quartiles. Individual participants are plotted as connecting dots to visualize within-participant changes.

*Note that sample size may fluctuate due to incomplete trial-by-trial participant responses within conditions. We analyzed all data where possible (*Materials and Methods* for details).

To test whether the impact of *Just Mercy* varied by viewer's preintervention ideology, we disaggregated the intervention group (in Fig. 2*A*) by four levels of political ideology (Fig. 3). These levels were 1) "other" left (N = 33), which included participants that identified as socialist or more economically left than liberal, 2) liberal (N =171), 3) conservative (N = 104), and 4) "other" right (N = 35), which included participants who identified as libertarians or more economically right than conservative. Participants in the intervention condition who did not identify with a political ideology (N = 1) were excluded from this analysis. In a LME model with participant intercepts treated as random effects, we found no significant interactions between visit and political ideology in the intervention group [$t(313.66) = 0.957$, $P = 0.34$; *SI Appendix*, Table S11]. However, two-tailed paired t tests (post vs. pre) *within* each political ideology subgroup, revealed that watching *Just Mercy* significantly increased support of criminal justice reform (Other Left: $t(27) = 3.52$, $P = 0.002$, mean difference = 3.45; Liberal: $t(159) = 8.67$, $P < 0.0001$, mean difference = 4.08; Conservative: $t(94) = 7.02$, $P < 0.0001$, mean difference = 3.42; Other Right: $t(30) = 3.34$, $P = 0.002$, mean difference = 3.90). Participants with missing visit-one data were eliminated from pairwise analyses, therefore, sample size may appear lower in the pairwise tests.

Exploratory Analysis: Does a Participant's Empathic Accuracy Mediate the Film's Effect on Criminal Justice Reform? To determine the psychological processes that may mediate the impact of *Just Mercy* on viewers' support for criminal justice, we performed an exploratory mediation analysis. We tested the relationships among film condition (X), empathic accuracy toward formerly incarcerated storytellers as measured by RMSE (M), and support for criminal justice reform (Y) on visit two. Of the 709 participants who completed the empathic accuracy tasks on visits one and two, only N = 696 could be included in this analysis due to insufficient data in the "formerly incarcerated" label trials on visit two (*Materials and Methods* for details). We found evidence for a partial mediation of the effect of condition on support for reform by participants' performance on the empathic accuracy task. That is, the impact of the intervention on support for reform decreased when controlling for empathic accuracy (*SI Appendix*, Fig. S6 and Table S12).

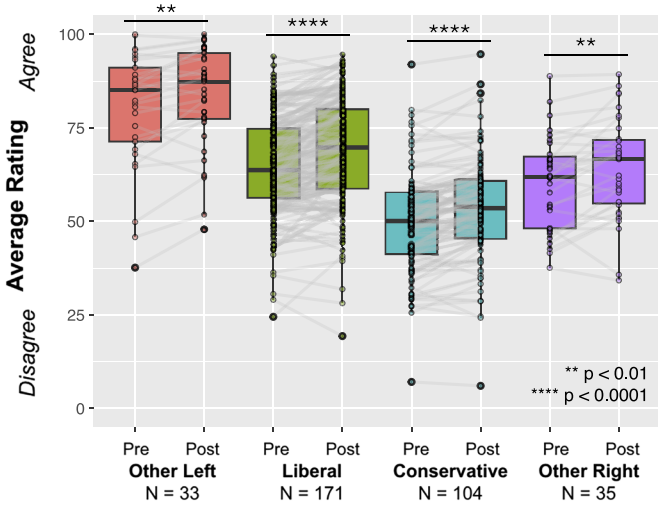

Fig. 3. Change in support for criminal justice reform by political ideology in the intervention group. Participants increased their support for criminal justice reform after the film intervention, regardless of their political ideology. However, participants with more right-wing political ideologies had a lower baseline support for reform upon entering the study, relative to participants with more left-wing political ideologies (*SI Appendix*, Tables S10 and S11). Here, "Other Left" encompasses participants who identified as socialist, communist, or anarchist; while "Other Right" encompasses participants who identified as economically libertarian or far-right. Liberal includes participants who identified as Democrats. Conservative includes participants who identified as Republicans. Data are represented in box plots where the median is a black line and the upper and lower "whiskers" represent the bounds of the quartiles. Individual participants are represented as dots and are connected to highlight within participant pre/post comparisons.

Exploratory Analysis: Does the Film Intervention Increase People's Willingness to Sign and Share a Petition to Restore Voting Rights to Formerly Incarcerated People? We also assessed the impact of the intervention film on a behavioral measure of policy support: whether participants indicated that they would sign and share a petition in support of a federal law to restore voting rights to people with criminal convictions (*SI Appendix*, Table S13). We modeled the questions after those commonly found on the popular website *change.org*. These questions were only included in visit two, a total of 749 people answered them.

We found that participants who had watched *Just Mercy* were marginally more willing to sign the petition and share it with friends and family compared to participants who had watched the control films (Fig. 4*A*). A Pearson's chi-squared test with Yates' continuity correction revealed that a marginally higher proportion of participants in the intervention condition signed the petition ($\chi2 = 3.23$, $P = 0.07$) than the control group. This is a 7.66% percent increase in signatures from the intervention group relative to the control groups. Though this result is in the pre-registered predicted direction, it was not significant. Similarly, participants in the intervention group were again only *marginally* more likely to share the petition with friends and family ($\chi2 = 3.47$, $P = 0.06$; Fig. 4*B*) than participants in the control group. This effect of the film on petition support is also in the predicted direction, though it is not significant.

Next, we compared our petition results with two snapshots of national polling data that surveyed support for voting rights for people with felony convictions inside and outside of prison (Fig. 4*C*). The proportion of people who signed our petition, in both conditions, was significantly higher than a 2019 national poll (N = 1002) conducted before the release of *Just Mercy* in 2020 (Intervention $\chi2 = 260.08$, $P < 0.001$; Control $\chi2 = 229.42$, $P < 0.001$) and a 2022 national poll (N = 1000) conducted at the same time as our data collection (Intervention $\chi2 = 99.69$, $P < 0.001$; Control $\chi2 = 73.96$, $P < 0.001$). Though willingness to sign the petition was only marginally influenced by the intervention film ($P = 0.07$), support for voting restoration on visit two of our experiment was significantly higher than the national average.

Discussion

A narrative film based on the true story of Walter McMillian—*Just Mercy*—improved viewers' empathic accuracy, that is, their ability to correctly infer the feelings of formerly incarcerated men, and increased viewers' support for criminal justice reform. Indeed, watching *Just Mercy* increased viewers' opposition to the death penalty by nearly 20%, an effect size of $d = 0.34$. This effect on policy preferences is higher than more resource-intensive interventions, such as those that involve attending live theater ($d = 0.16$ to 0.33; (26)) and "deep canvassing," a political campaign strategy where volunteers have conversations with constituents to increase voter turnout (typically resulting in a 10% increase; see ref. 41). Notably, the prosocial effects of *Just Mercy* on viewer's orientation toward incarcerated people occur regardless of viewers' initial political ideology. Together, these effects suggest a film's ability to induce cognitive empathy (i.e., increased empathic accuracy) is a

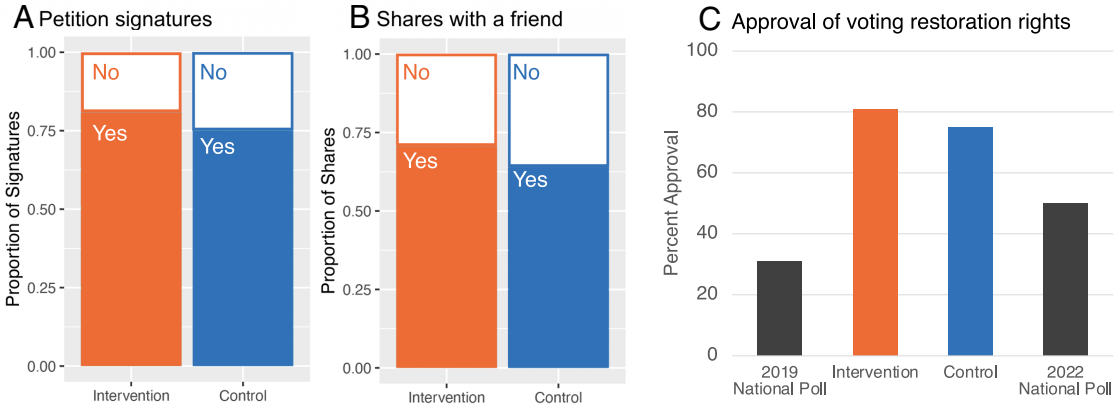

Fig. 4. Film intervention marginally motivates action to restore voting rights to formerly incarcerated people. (*A*) Petition signatures. The film intervention marginally motivated people to sign a real petition to restore voting rights to formerly incarcerated people in states that do not allow it, relative to the control conditions. Plotted are the proportions of all respondents who signed (solid color, "Yes") and who did not sign (in outline color, "No") in the intervention (N = 344) and control groups (N = 405). The difference in signatures by group is in the predicted direction but is not significant ($P = 0.07$). (*B*) Shares with a friend. The film intervention also marginally motivated people to share the petition with a friend, relative to the control condition. Plotted are the proportions of all respondents who shared (solid color, "Yes") and who did not share (in outline color, "No") in the intervention (N = 344) and control groups (N = 405). Although in the predicted direction, the difference in shares by group is also not significant ($P = 0.06$). (*C*) Approval of voting restoration rights. Here, we compare the percentage of participants who signed "Yes" in the intervention and control groups to two equivalent national polls of registered voters. The first poll was conducted before the film *Just Mercy* was released (Hill-HarrisX 2019) while the second was conducted during the same time we collected our data (Lake Research Partners 2022). Overall percentages in our study were significantly larger than both national polls ($P < 0.001$). This could be a demand effect, or it is possible that listening to stories from people labeled formerly incarcerated influences approval votes in this poll regardless of film intervention.

key factor in how it can influence an individual's support for political policies.

Contrary to our expectations, we found that self-reported compassion for formerly incarcerated people decreased for all participants at visit two. There are several factors to consider when interpreting this unexpected trend. First, compassion ratings were, on average, quite high across all participants. Even when they decreased, they were still on the very positive end of the spectrum. Therefore, ceiling effects in the compassion measure may have made it difficult to assess actual changes to this subcomponent of empathic processing. Second, given our lengthy (~4-hour) experimental design, it is possible that participants experienced compassion fatigue—a feeling of emotional exhaustion after experiencing empathy (31, 42). Yet, despite this fatigue, compassion toward targets labeled "formerly incarcerated" dropped less than it dropped for those labeled "college student" for participants in the intervention condition. This could indicate that *Just Mercy* buffers against compassion fatigue, which partially supports our hypotheses. Though we must be cautious when interpreting null findings, our results suggest that films which selectively activate cognitive components of empathy (here, empathic accuracy) may be more effective at inducing attitude change than films which primarily intend to elicit sympathy or compassion (which here, was already near ceiling).

In designing the study, we made a number of methodological choices to increase robustness. First, the large representative sample in our study ensures that our findings are generalizable to the US population. Second, assessing empathic accuracy and compassion for storytellers who have been incarcerated provided an opportunity to approximate participants' responses to people in real life (not actors on a screen) who experience real stigma. Third, by giving participants an opportunity to indicate whether they would sign and share a change.org-style petition, we mimicked real-world campaigns for criminal justice reform and were able to quantify the impact of the film intervention (*Just Mercy*) on viewers' willingness to take action for social change (also see ref. 43). Together, these considerations underscore both the sensitivity of our methods in assessing the impacts of a narrative film intervention, and the relevance of our findings to real-world attitudes toward members of a stigmatized group.

Nevertheless, our experimental design had limitations. First, we measured empathic accuracy *before* participants watched the film as well as after they watched the film. This was intentional—we needed to acquire a baseline measure of empathic accuracy. However, this may have primed participants, especially those in the *Just Mercy* condition, to pay more attention to themes of incarceration and therefore increased the film's impact (44). Second, we used the relatively neutral label of "formerly incarcerated" for the storytellers. Using more negatively charged labels, such as "former felon" or "ex-con," may result in different levels of empathy and compassion. Third, the long duration of the experiment and the heavy emotional load of the stories might have caused fatigue or influenced participant dropout and missingness in the data. Finally, though our control films were carefully matched to *Just Mercy* on multiple important perceptual and emotional characteristics (i.e., rooting for the underdog) to isolate the effect of *Just Mercy's* storyline—we do not know whether *Just Mercy* induced a substantially different mood than the control films. Mood congruency is known to influence affective but not cognitive empathy (45), therefore, the effects of mood are important to disentangle from the effects of the storyline in the film. Future research might explore how these design choices impact the magnitude of the effects that we report here.

Our findings suggest narrative elements that tap into viewers' empathy are particularly important in shifting attitudes. Future studies are needed to parse the impacts of different narrative elements in eliciting empathic responses and the extent to which those responses translate to changes in policy preferences (see ref. 46). Our methods may provide a useful approach in quantifying these effects. Teasing apart the useful elements of narratives would also provide valuable information for guiding the entertainment industry's selection of scenes in their productions. Future studies are needed to determine how long-lasting these effects are and to better understand the mechanisms underlying viewers' empathic responses. Greater clarity on the mechanism—both psychologically and neurobiologically—could significantly inform the development of future interventions.

Motivating people to seek out empathy-enhancing entertainment, specifically when the content of those narratives contradicts viewers' preexisting attitudes or beliefs, remains a major challenge. However, our results validate the power of narratives to improve human connection and promote the desire for system-level change.

Materials and Methods

Materials and procedures for this study were approved by the Stanford University Non-Medical Institutional Review Board. The experiment was preregistered prior to data collection (see: https://osf.io/pqu7d). All study participants gave informed consent to participate in the study.

Experimental Design. We ran an online experiment to test two preregistered hypotheses: that watching an intervention film with nuanced narratives about incarcerated people (*Just Mercy)*, compared to watching a control film (*Concussion* or *Moneyball)*, will 1) improve observers' compassion and ability to accurately infer the emotions of storytellers when they are labeled as "formerly incarcerated" relative to when they are labeled as "college student" and 2) increase support for a variety of criminal justice reforms. We ran additional analyses to test for a) moderation by storyteller race as well as by b) observer race, gender, socioeconomic status, and political ideology (preregistered), and c) the mediational role of empathy on support for criminal justice reform (not preregistered).

This experiment was conducted within a longitudinal randomized controlled trial which included five online "visits" spanning over the course of 3 mo. Data were collected in three waves (each involving five visits for enrolled participants). The first wave began in June 2022, the second began in July 2022, and the third began in November 2022. Here, we focus on only visit one (pre-intervention) and visit two (post-intervention). Participants completed both visits within the span of three days and were allowed to take as many breaks as they needed.

On visit one, participants completed a baseline assessment of empathic accuracy for formerly incarcerated people (8 video trials) and a survey with measures of attitudes about the criminal justice system (see *SI Appendix*, Tables S8 and S14 for individual items). Survey items were presented in a random order. On visit two, participants were randomly assigned to watch *Just Mercy (2019*; Intervention film*)*, *Concussion* (2015; Control condition 1), or *Moneyball* (2011; Control condition 2). After watching the assigned film, participants completed another empathic accuracy task (8 new video trials), and the same survey again with two additional questions about a petition. Finally, they completed a demographic survey assessing their gender, race, socioeconomic status (SES), political affiliation, and personal familiarity with the criminal justice system.

Intervention and Control Stimuli. We predicted that the storyline of the intervention film (*Just Mercy*), which is based on the true story of a lawyer (Bryan Stevenson) who worked to overturn the sentences of men on death row who were wrongly convicted–would be the primary contributor to any effects we found. We carefully selected the two control films, *Concussion* and *Moneyball,* to match with the intervention film on key features, so that the storyline would be the main difference across the films. Thus, the control films are of the same genre (DocuDrama) and broad theme (a lead male protagonist that goes up against a system) as *Just Mercy*, and all films are rated PG-13, ~2 h long, released after 2010, have an IMDB rating ~7, have a Rotten Tomatoes above 55%, and box offices above ~50 million or higher. We additionally selected one control film that had a main character that was of the same race and gender as the intervention film (a Black man; *Concussion,* Columbia Pictures, 2015) and one control film that had a main character that was not the same race (a White man; *Moneyball*, Columbia Pictures, 2011) so that we could test one alternative hypothesis: that the race of the main character of the film would affect changes in support for a protagonist fighting to change a system.

Stimulus Collection for the Empathic Accuracy Task. The stimuli for the empathic accuracy task consisted of real-life stories told by formerly incarcerated men (in 1 to 3 min). We focused on men because men make up more than 90% of the US prison population. We recruited our sample of formerly incarcerated male storytellers by working with organizations and individuals who aid formerly incarcerated people, and through snowball sampling, in which storytellers recommended other formerly incarcerated men to assist us with the study. Storytellers had to be male, over the age 18, and were (or had recently been) a college student. We also required storytellers to have access to a computer and reliable internet connection. Data collection was completed during COVID lockdowns, so we were unable to set up a video recording room in person. Those interested in assisting were asked to contact us via email. After determining eligibility, a research assistant mailed the storytellers' equipment (i.e., a webcam, ring light, gray backdrop, and gray sweatshirt). This equipment was used to standardize the video quality, lighting, backdrop, and clothing of the storytellers.

The research assistant met with each storyteller via Zoom to answer questions and ensure they knew how to use our website to record their stories. They were instructed to tell four emotional life stories, each no longer than 3 min. After each story, the storyteller watched their video and used a slider scale to continuously rate how they felt while telling that story on a scale from 0 (negative emotion) to 100 (positive emotion). Next, they completed the demographic questionnaire and the beliefs and attitudes survey. Finally, they were paid for their contribution.

At least two members of the study team watched each video and coded it for video quality, audio quality, and content (specifically, that the story told was unrelated to incarceration). We recruited storytellers until we obtained suitable videos featuring stories from 20 storytellers who were perceived as Black and 20 storytellers who were perceived as White. In total, we recruited 73 storytellers, but some identified as biracial. We ran a pilot study to determine which storytellers were perceived primarily as Black or White before finalizing their inclusion as stimuli. For this purpose, a nationally representative sample of participants (N = 51) were recruited online via Bovitz Inc. They viewed the faces of each storyteller and were asked to identify the storyteller's race. They were able to select more than one race. Stimuli from biracial participants were utilized in the present experiment only if they were categorized by more than 50% of raters as monoracial Black or White.

After selecting the 40 stimulus videos, two researchers then made time edits to some of the videos (i.e., to shorten some stories, or, for example, to cut out a section where a storyteller mentions their son's name). A professional video editor standardized the sound levels of all videos, improved the visual quality of some videos, added a gray background to one video, and edited face tattoos off one storyteller. A researcher then created two versions of each video. One version included the label "college student" and listed the person's race, sex, and age. The other version used the label "formerly incarcerated" and included the same three demographic variables (e.g., "Formerly Incarcerated," Black Man, Age 52). The labels appeared without the video for the first 5 seconds, then the video appeared. And the labels remained on the screen for the duration of the video. We do not have consent to share all of the videos publicly but summarize their content in *SI Appendix*, Fig. S2.

Recruitment and Randomization. For the main study, we recruited a nationally representative sample through the participant recruitment company Bovitz Inc. Our preregistered goal was to collect N = 780 participants (which was well over the 318 participants our preregistered power analysis indicated was needed to detect a difference in means in empathy between the treatment and control groups of 0.32 with 80% power) because we wanted to examine moderation of effects by gender, race, class, and political affiliation. We did not expect all eligible participants identified by Bovitz to want to participate in a longitudinal study, so we deliberately oversampled during the eligibility phase in order to reach our target N. 1834 individuals who completed our screener survey met inclusion criteria (i.e., had not seen any of the films, had never been incarcerated, and reported that they could participate in the full study which included longitudinal follow-ups not reported here), and all were invited to participate (see enrollment diagram in *SI Appendix*, Fig. S1). We assigned them to the between-subject treatment groups via stratified-block randomization by gender, race, SES, and political affiliation (*SI Appendix*, Table S1). SES was assessed via the MacArthur Ladder, which instructs the participant: *"Think of this ladder as representing where people stand in the United States. At the top of the ladder are the people who are best off–those who have the most money, the most education, and the most respected jobs. At the bottom are the people who are worst off–who have the least money, least education and the least respected or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the very bottom. Please indicate the rung where you think YOUR FAMILY stands at this time in your life, relative to other people in THE UNITED STATES."* There are 10 rungs, so the data are converted to values 1 to 10.

Stop-Check Analysis. Initially, 33% of participants were randomly assigned to each film (*Just Mercy*, *Concussion*, or *Moneyball*). However, following our preregistered plan, we collected data in waves so that after ~30% of our total desired sample size of participants completed visits one and two, we could pause data collection to check for statistically significant differences between the two control films (*Concussion* and *Moneyball*) on the criminal justice reform questions. The purpose of this was to maintain control validity while maximizing our power within our budget and time constraints. After the second wave of data collection, we paused data collection and found no meaningful differences between control films on the criminal justice reform questions. For example, changes in attitudes to the death penalty did not significantly differ between the two control films in a two-tailed independent samples *t* test [*Concussion* vs. *Moneyball*: $t(123.24) = -0.58$, $P = 0.56$, CI = (−8.13, 4.45)]. Therefore, the rest of data collection stratified randomization such that 50% of participants were randomly assigned to *Just Mercy*, 25% were assigned to *Concussion* and 25% to *Moneyball*.

Manipulation Check. To be included in our analyses, participants had to correctly answer two-thirds of the questions on a simple quiz about their assigned film (*SI Appendix*, Table S4). This threshold was preregistered before data collection. 98.17% of participants passed the quiz. Six participants who watched *Just Mercy*, four who watched *Concussion*, and four who watched *Moneyball* were eliminated.

Attrition, Exclusions, and Final Sample. In total, 856 participants completed at least some portion of visit one. However, after exclusions due to participant-reported website and/or computer errors, duplicate submissions, and preregistered standards for inclusion (i.e., attrition from pre- to postintervention or failing the manipulation check), our total sample was 749 participants. However, not all of 749 participants completed every portion of the two visits completely. Due to there being multiple platforms for different portions of data collection for each visit, and due to participant error and fatigue, there may be random missing data from these participants at various stages of data collection. We used all data whenever possible. Attrition did not differ by condition (*SI Appendix*, Tables S2 and S3). Demographic variables of interest were stratified across groups in the final sample, and there were no differences between the intervention group and control group on their self-reported familiarity with the criminal justice system [$t(745) = -0.76$, $P = 0.45$; Intervention mean = 60.63; Control mean = 61.91].

Criminal Justice Policy Preferences and Petition Analysis. Support for criminal justice reform was operationalized using a preregistered index averaging across participants' responses to 20 items, each of which was reported on a scale of 0 to 100. *SI Appendix*, Table S6 for individual items and their statistics. Support for signing and sharing the mock petition was also included in only the visit two survey (*SI Appendix*, Table S13 for exact wording). These surveys were hosted on Qualtrics, with the Qualtrics survey page embedded on the study website. Question order was randomized for each unique visit and for each unique participant. There was missingness in the data due to some participants experiencing technical errors and/or skipping through the website page (which hosted the Qualtrics survey) without completing it. Missingness did not differ by condition. Thus, we excluded participants with missing data in visit one survey responses (N = 73) when conducting pairwise (pre- vs. post-intervention) *t* tests (Fig. 2). We excluded participants with missing perceived SES data (N = 76) from analyses that use this to control for the effects of social class. Additionally, two people did not complete any of the demographic questionnaire, so their data were also excluded in analyses where those variables are used to control for individual differences in social identities.

We inputted all 749 participants (with NAs in place of missing data) in an analysis of the relationships among condition, time, and support for criminal justice reform. Analysis was performed using the LMER package in R (R Core Team, 2018). Condition (intervention vs. control) and time (pre- vs. postintervention) were modeled as fixed effects. Participant intercepts were modeled as random effects. LME models were fit by REML. Satterthwaite's method for estimating degrees of freedom was used for inference (*t* tests). All p-values reported throughout this paper are two-tailed.

Snapshot of National Polling Data. We compared the results of our petition to snapshots of national polling data surveying restoration of voting rights to people convicted of a felony. To find these snapshots, we searched the internet for polls asking a similar question at two timepoints 1) right before the film *Just Mercy* was released (Dec 2019), and 2) during our data collection (June 2022 to Dec 2022). We looked for polls related to voting restoration for people convicted of a felony that made their raw counts public (not just percentages). We selected a poll by Hill-HarrisX conducted April 27 to 28, 2019, on a national sample of registered voters (N = 1002; see ref. 47) and a poll by Lake Research Partners conducted July 11 to 17, 2022, also on a national sample of registered voters (N = 1000; see ref. 48). The Hill-HarrisX poll was worded as follows: "Should each of the following groups be able to vote in national elections or not?" The list included "people serving time in prison because of a felony." The Lake Research poll was worded as follows: "Would you support or oppose a law that guarantees the eligibility to vote for all citizens 18 and older, including citizens with felony convictions both inside and outside of prison?" Our wording can be found in *SI Appendix*, Table S13. Note, the national polls include voting restoration for people currently in and out of prison whereas our petition emphasizes restoration for people who completed their sentence.

Empathic Accuracy Analysis. We measured participants' ability to infer the feelings of storytellers both before and after the film intervention with an empathic accuracy task (*SI Appendix*, Fig. S2). In each task, participants viewed 8 short (1 to 3 min) videos of men (all over age 18) telling negative, emotional life stories, while continuously rating how they believed the storyteller felt while telling their story, on a scale from negative to positive emotion. These videos were presented in a randomized order and drawn from the set of 40 stimulus videos. In each session, two storytellers were Black and labeled "formerly incarcerated," two were White and labeled "formerly incarcerated," two were Black and labeled "college student," and two were White and labeled "college student." Label assignment was randomized across participants and visits. This was done so that we could assess the effect of the label "formerly incarcerated" above and beyond the race of the storyteller and the content of the story. After watching each video, participants were asked to recall the label on the video as a manipulation check to ensure participants had seen the label.

For the analysis, first, we preprocessed the rating data. All rating traces were smoothed by a moving average of 10 samples. Next, we aligned the storyteller's self-ratings (emotion ratings made by the storytellers while viewing their own video after recording it) with each observer's inference ratings in time. Participants had to have at least one usable rating trace within a single visit to be included in the complete analysis. We also preregistered that we would exclude rating traces from the analysis due to failed manipulation checks (i.e, the participant did not correctly identify the label) and random technical error. Indeed, some participants had slow WIFI connections that affected their ability to watch the videos and make ratings, which resulted in misalignment between the ground truth self-ratings and the observer ratings. We made a non-preregistered decision to exclude any observer rating trace that was more than 12 seconds longer than its ground truth rating. We chose these criteria because it is the length of social media "reels" and other short videos that are designed to induce content comprehension in the shortest period possible. Trials that fell within these criteria were cropped as appropriate and realigned to the start of the self-rating trace. Overall, a total of 2,625 (out of a total of 11,984; 21.90%) video trials were removed from the analysis due to failed manipulation checks and/or misalignment. The number of missing traces did not differ by condition. A total of 40 participants lost all trial data due to extreme misalignment or failed manipulation checks. We allowed participants into the analysis so long as they had some usable data from either visit one or visit two. A total of N = 709 participants met these criteria, but certain analyses that focus only on "formerly incarcerated" labeled trials on visit two have a total N of 696 participants that met this criterion. In analyses that account for these effects above and beyond political and social identities, there is a total of N = 671 data due to missingness in SES ladder responses and the demographic survey as explained earlier. The sample size of this analysis is still within the preregistered total range of participants (N = 600 to 780).

Using both participants' ratings and storytellers' ratings, we calculated empathic accuracy—our measure of one's ability to correctly infer the feelings and emotions of others—by taking the root mean squared error (RMSE) between the participant's inference rating trace and the storyteller's self-reported emotional rating trace. RMSE is the ideal metric for assessing accuracy in this paradigm because it is scale-dependent, and the ratings we wish to compare are made on the same scale. That is, we do not simply want to measure whether an observer

noticed a positive change in the storyteller's emotion, but that they accurately assessed the initial baseline (i.e., level 70) and any subsequent change (e.g., an increase to level 90). RMSE is always nonnegative; a value of 0 indicates a perfect fit to the "ground truth" self-ratings. Therefore, lower RMSE indicates higher accuracy. Because this is counterintuitive, we invert RMSE for plotting purposes in Fig. 1*A* so that positive values indicate greater accuracy. That is, we flipped the sign of the Post vs. Pre change in RMSE so that an increase in error would yield a negative change score and a decrease in error would yield a positive change score. However, the true RMSE values are used in all statistical tests and are denoted as "empathic error" in such tests including the mediation analysis (See *SI Appendix*, Fig. S6). Our preregistration describes the use of Pearson's correlation, but we decided to use RMSE before conducting our analyses due to its superiority as a measure. RMSE is a stricter test of accuracy than taking a Pearson's correlation between the two traces, because Pearson's correlations are invariant to the mean of the trace, that is, a participant could rate a trace entirely in the "negative" portion of the scale but change it at the same rate and magnitude as a storyteller who rated themselves entirely in the "positive" portion of the scale.

We hypothesized that watching the intervention film, *Just Mercy*, would selectively increase one's ability to infer the emotions of formerly incarcerated individuals; therefore, we used a LME model to test for a time (pre vs. post) by condition (intervention vs. control) by label ("formerly incarcerated" vs. "college student") interaction on empathic accuracy (RMSE) in the empathic accuracy task. RMSE values were averaged within participant for each label within time. Participants were input as random effects. Time, condition, and label were input as fixed effects. A second LME model that controlled for participant race, gender, SES, and political ideology was then run to test the robustness of this interaction above and beyond relevant demographic variables. LME models were fit by REML. Satterthwaite's method for estimating degrees of freedom was used for inference (*t* tests) unless otherwise specified.

Compassion Analysis. Compassion for each storyteller was measured during the empathic accuracy tasks. After participants watched and rated each storyteller's video, they reported their level of compassion for the storyteller on a scale from 0 to 100. Data were removed with the same criteria as described above in the empathic accuracy analysis. We hypothesized that watching the intervention film, *Just Mercy*, would selectively increase one's compassion for formerly incarcerated individuals; therefore, we again used an LME model to test for a time (pre vs. post) by condition (intervention vs. control) by label ("formerly incarcerated" vs. "college student") interaction on compassion in the empathic accuracy task. Raw compassion values were averaged within participant by label within time. Participants were input as random effects. Time, condition, and label were input as fixed effects.

Data, Materials, and Software Availability. Anonymized spreadsheets, code, figure files data have been deposited in NarrativeIntervention_PNAS_2023 (https://osf.io/eugjd/) (49). All study data are included in the article and/or *SI Appendix*.

Author affiliations: [a]Department of Psychology, Stanford University, Stanford, CA 94305; [b]Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, The Bronx, NY 10461; [c]Department of Psychology, Princeton University, Princeton, NJ 08544; [d]Department of Psychology, Stanford SPARQ, Stanford University, Stanford, CA 94305; and [e]Graduate School of Business, Stanford University, Stanford, CA 94305

1. B. Stevenson, *Just Mercy: A Story of Justice and Redemption* (Spiegel & Grau, 2014).
2. Prison Policy Initiative, United States profile. (2024). Available at: https://www.prisonpolicy.org/profiles/US.html. Accessed 25 July 2024.
3. J. Schnittker, M. Massoglia, C. Uggen, Out and down: Incarceration and psychiatric disorders. *J. Health Soc. Behav.* **53**, 448–464 (2012).
4. K. Turney, C. Wildeman, J. Schnittker, As fathers and felons: Explaining the effects of current and recent incarceration on major depression. *J. Health Soc. Behav.* **53**, 465–481 (2012).
5. D. Pager, The mark of a criminal record. *Am. J. Sociol.* **108**, 937–975 (2003).
6. D. Pager, B. Western, N. Sugie, Sequencing disadvantage: Barriers to employment facing young black and white men with criminal records. *Ann. Am. Acad. Pol. Soc. Sci.* **623**, 195–213 (2009).
7. J. P. Dunlea, D. Goel, L. Heiphetz, Children's socio-moral judgments and behaviors toward peers with and without incarcerated parents. *Child Dev.* **93**, e515–e530 (2022).
8. P. Smith, C. Goggin, P. Gendreau "The Effects of Prison Sentences and Intermediate Sanctions on Recidivism: General Effects and Individual Differences" (Solicitor General of Canada, 2002).
9. S. Han. M. Valdovinos Olson, R. C. Davis, Reducing recidivism through restorative justice: An evaluation of Bridges to Life in Dallas. *J. Offender Rehabil.* **60**, 444–463 (2021).
10. M. Ellison, K. Szifris, R. Horan, C. Fox, A rapid evidence assessment of the effectiveness of prison education in reducing recidivism and increasing employment. *Probat. J.* **64**, 108–128 (2017).
11. M. Vasiljevic, G. T. Viki, "Dehumanization, moral disengagement, and public attitudes to crime and punishment" in *Humanness and Dehumanization*, P. G. Bain, J. Vaes, J. P. Leyens, Ed. (Psychology Press, 2013), pp. 129–146.
12. J. M. Binnall, Respecting beasts: The dehumanizing quality of the modern prison and an unusual model for penal reform. *J. Law Policy* **17**, 31 (2008).
13. J. C. Deska, S. M. Almaraz, K. Hugenberg, Dehumanizing prisoners: Remaining sentence duration predicts the ascription of mind to prisoners. *Pers. Soc. Psychol. Bull.* **46**, 1614–1627 (2020).
14. C. D. Batson, L. L. Shaw, Evidence for altruism: Toward a pluralism of prosocial motives. *Psychol. Inq.* **2**, 107–122 (1991).
15. C. D. Batson, J. Chang, R. Orr, J. Rowland, Empathy, attitudes, and action: Can feeling for a member of a stigmatized group motivate one to help the group? *Pers. Soc. Psychol. Bull.* **28**, 1656–1666 (2002).
16. A. A. Marsh, The caring continuum: Evolved hormonal and proximal mechanisms explain prosocial and antisocial extremes. *Annu. Rev. Psychol.* **70**, 347–371 (2019).
17. J. A. Okonofua, D. Paunesku, G. M. Walton, Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5221–5226 (2016).
18. G. M. Walton *et al.*, Lifting the bar: A relationship-orienting intervention reduces recidivism among children reentering school from juvenile detention. *Psychol. Sci.* **32**, 1747–1767 (2021).
19. C. D. Batson *et al.*, Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group?. *J. Pers. Soc. Psychol.* **72**, 105–118 (1997).
20. D. Broockman, J. Kalla, Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science* **352**, 220–224 (2016).
21. E. L. Paluck, R. Porat, C. S. Clark, D. P. Green, Prejudice reduction: Progress and challenges. *Annu. Rev. Psychol.* **72**, 533–560 (2021).
22. S. Murrar, M. Brauer, Entertainment-education effectively reduces prejudice. *Group Process. Intergroup Relat.* **21**, 1053–1077 (2018).
23. A. Alrababa'h, W. Marble, S. Mousa, A. A. Siegel, Can exposure to celebrities reduce prejudice? The effect of Mohamed Salah on Islamophobic behaviors and attitudes. *Am. Polit. Sci. Rev.* **115**, 1111–1128 (2021).
24. T. F. Pettigrew, L. R. Tropp, A meta-analytic test of intergroup contact theory. *J. Pers. Soc. Psychol.* **90**, 751–783 (2006).
25. E. Schiappa, P. B. Gregg, D. E. Hewes, The parasocial contact hypothesis. *Commun. Monogr.* **72**, 92–115 (2005).
26. S. Rathje, L. Hackel, J. Zaki, Attending live theatre improves empathy, changes attitudes, and leads to pro-social behavior. *J. Exp. Soc. Psychol.* **95**, 104138 (2021).
27. J. Vorauer, "The case for and against perspective-taking" in *Advances in Experimental Social Psychology*, J. M. Olson, M. P. Zanna, Eds. (Elsevier Academic Press, 2013), pp. 59–115.
28. J. Dixon, L. R. Tropp, K. Durrheim, C. Tredoux, "Let them eat harmony": Prejudice-reduction strategies and attitudes of historically disadvantaged groups. *Curr. Dir. Psychol. Sci.* **19**, 76–80 (2010).
29. J. Zaki, M. Cikara, Addressing empathic failures. *Curr. Dir. Psychol. Sci.* **24**, 471–476 (2015).
30. E. L. Paluck, Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *J. Pers. Soc. Psychol.* **96**, 574–587 (2009).
31. J. Zaki, Empathy: A motivated account. *Psychol. Bull.* **140**, 1608–1647 (2014).
32. N. Epley, E. M. Caruso, "Perspective taking: Misstepping into others' shoes" in *Handbook of Imagination and Mental Simulation*, (Psychology Press, 2009), pp. 295–309.
33. K. J. K. Klein, S. D. Hodges, Gender differences, motivation, and empathic accuracy: When it pays to understand. *Pers. Soc. Psychol. Bull.* **27**, 720–730 (2001).
34. M. W. Kraus, S. Côté, D. Keltner, Social class, contextualism, and empathic accuracy. *Psychol. Sci.* **21**, 1716–1723 (2010).
35. M. Cikara, S. T. Fiske, Bounded empathy: Neural responses to outgroup targets' (mis)fortunes. *J. Cogn. Neurosci.* **23**, 3791–3803 (2011).
36. D. A. Redelmeier, A. Tversky, Discrepancy between medical decisions for individual patients and for groups. *N. Engl. J. Med.* **322**, 1162–1164 (1990).
37. A. D. Galinsky, G. B. Moskowitz, Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *J. Pers. Soc. Psychol.* **78**, 708–724 (2000).
38. M. C. Green, T. C. Brock, The role of transportation in the persuasiveness of public narratives. *J. Pers. Soc. Psychol.* **79**, 701–721 (2000).
39. J. Zaki, J. Weber, N. Bolger, K. Ochsner, The neural bases of empathic accuracy. *Proc. Natl. Acad. Sci.* **106**, 11382–11387 (2009).
40. W. Ickes, Empathic accuracy. *J. Pers.* **61**, 587–610 (1993).
41. R. M. Alvarez, A. Hopkins, B. Sinclair, Mobilizing pasadena Democrats: Measuring the effects of partisan campaign contacts. *J. Polit.* **72**, 31–44 (2010).
42. C. D. Cameron, L. T. Harris, B. K. Payne, The emotional cost of humanity: Anticipated exhaustion motivates dehumanization of stigmatized targets. *Soc. Psychol. Personal. Sci.* **7**, 105–112 (2016).
43. R. C. Hetey, J. L. Eberhardt, Racial disparities in incarceration increase acceptance of punitive policies. *Psychol. Sci.* **25**, 1949–1954 (2014).
44. R. Bilali, J. R. Vollhardt, Priming effects of a reconciliation radio drama on historical perspective-taking in the aftermath of mass violence in Rwanda. *J. Exp. Soc. Psychol.* **49**, 144–151 (2013).
45. H. Pan *et al.*, Mood congruency affects physiological synchrony but not empathic accuracy in a naturalistic empathy task. *Biol. Psychol.* **184**, 108720 (2023).
46. S. Keen, A theory of narrative empathy. *Narrative* **14**, 207–236 (2006).
47. The Hill, Poll: 69 percent of Americans say prisoners shouldn't be allowed to vote. (2019). Available at: https://thehill.com/hilltv/what-americas-thinking/441863-poll-69-percent-of-voters-say-prisoners-shouldnt-be-allowed-to/. Accessed 25 July 2024.
48. D. Gotoff, S. Markowitz, P. McCauley, R. Ferrentino New national poll shows majority favor guaranteed right to vote for all. https://www.sentencingproject.org/app/uploads/2022/10/Guaranteed_Right_to_Vote_Survey_Findings.pdf. Accessed 20 November 2022.
49. M. C. Reddan, S. B. Garcia, G. Golarai, J. L. Eberhardt, J. Zaki, NarrativeIntervention_PNAS_2024 Data and code to accompany manuscript publication. OSF. https://osf.io/eugjd/. Deposited 26 December 2023.