PNAS

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

Marianne C. Reddan,[1,2] Sydney B. Garcia,[1,3] Golijeh Golarai,[1] Jennifer Eberhardt,[1,4] & Jamil Zaki[1]

[1]Department of Psychology, Stanford University, Stanford, CA, [2]Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, The Bronx, NY, [3]Department of Psychology, Princeton University, Princeton, NJ, [4]Graduate School of Business, Stanford University, Stanford, CA

* Jennifer Eberhardt, Marianne C. Reddan
Email: jleberhardt@stanford.edu, mcredden@stanford.edu

This PDF file includes:

Fig. S1. Enrollment Diagram. 5,296 people completed the screener survey emailed by *Bovitz Inc.*, and 1,834 eligible participants were invited to participate. To be eligible, a person had to (a) report being available for the full study (which included longitudinal follow ups not reported in the current paper), (b) not have seen any of the three films, (c) not be formerly incarcerated, and (d) have access to a computer and WIFI. Participants were automatically randomized to the intervention or to one of two control conditions at their time of enrollment. Participants were blinded to their assignment. Of those participants, 896 registered for the study on our website. 856 went on to visit one. Of those participants, 763 watched the film and returned for visit two. Following our pre-registered plan, we excluded participants who failed the post-movie comprehension quiz (Table S4), leaving 749 participants total. Some participants experienced technical errors on the website that caused them to lose visit one or visit two empathic accuracy task data (videos did not load or got stuck). In addition, some participants failed the label manipulation checks. Therefore, N = 709 participants were included in the empathic accuracy and compassion analyses if they had any useable empathic accuracy data from visit one or visit two, while 749 participants were included in the survey analyses if they had survey data from visit one or visit two. Whenever possible, we used all the data available and input NaNs for missing trials. Analyses that focused only on 'formerly incarcerated' labeled empathic accuracy trials only had N = 696 useable subjects, which is the smallest N for any analysis included in this study. Our pre-registered minimum targeted recruitment total was 600 participants, and we were able to surpass this amount. We pre-registered an enrollment range of 600 – 780 participants because we factored attrition into our power analysis. Data were analyzed only during a pre-registered checkpoint (at N = 232 participants which was ~30% of the pre-registered sample to be collected), and then after data collection for the entire longitudinal study was completed. At the time of publication of this paper, we have only analyzed visit one and visit two data for the complete data set. The impact of attrition by demographic variables of interest was negligible (see Tables S2 and S3).

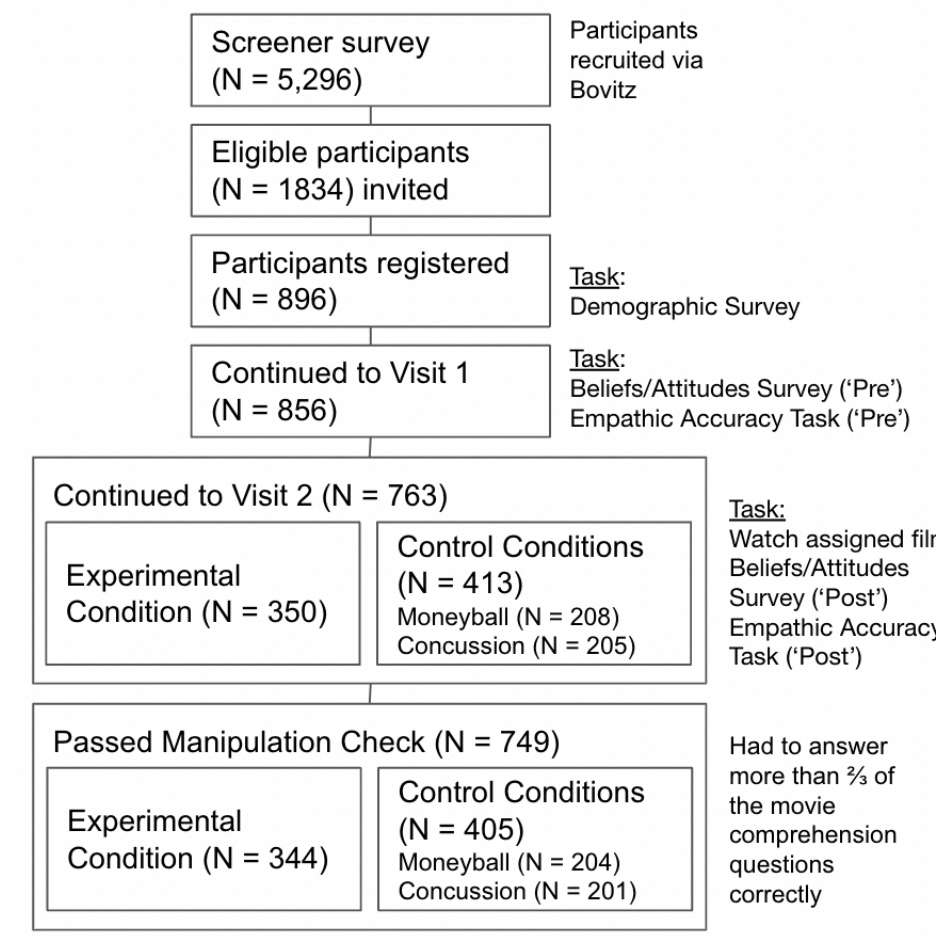
Screener survey
(N = 5,296)
Participants recruited via Bovitz
Eligible participants
(N = 1834) invited
Participants registered
(N = 896)
Task:
Demographic Survey
Continued to Visit 1
(N = 856)
Task:
Beliefs/Attitudes Survey ('Pre')
Empathic Accuracy Task ('Pre')
Continued to Visit 2 (N = 763)
Experimental Condition (N = 350)
Control Conditions
(N = 413)
Moneyball (N = 208)
Concussion (N = 205)
Task:
Watch assigned filr
Beliefs/Attitudes Survey ('Post')
Empathic Accuracy Task ('Post')
Passed Manipulation Check (N = 749)
Experimental Condition (N = 344)
Control Conditions
(N = 405)
Moneyball (N = 204)
Concussion (N = 201)
Had to answer more than ⅔ of the movie comprehension questions correctly

Fig. S2. Paradigm Schematic. A. Timeline of experimental procedures. During visit one, participants registered for the website and completed a baseline ("pre") assessment of empathic ability and compassion for incarcerated people (empathic accuracy task) and their support for criminal justice reform (beliefs & attitudes survey). On visit two participants watched one of 3 films: (*Just Mercy* (Intervention condition), *Concussion* (Control condition 1), or *Moneyball* (Control condition 2). Then participants completed the "post" empathic accuracy task (8 new video trials) and the beliefs & attitudes survey again. Participants had to complete visits one and two within a 3-day period. At the end of visit two, they completed a demographic survey, **B. Example of a trial in the empathic accuracy task.** During the empathic accuracy task participants watch videos of real formerly incarcerated people telling true emotional stories about their lives. These formerly incarcerated people are current students at various institutions. Participants rate how they think the storyteller is feeling, moment-by-moment, on a scale that ranges from 0 (negative) to 100 (positive). Participants also indicate how much compassion they felt for each storyteller in each video, on a scale that ranges from 0 (not at all) to 100 (very much). Though each storyteller is both formerly incarcerated and a student, participants are told that half of the storytellers are college students and half of the storytellers are formerly incarcerated. This label appears on screen for the first five seconds before the video appears, then remains on the duration of the video. Storytellers are White or Black men over the age of 18 who have fully consented to have their stories used in this experiment. In each instance of this task, half of the storytellers are Black and half are White. Race and label (formerly incarcerated or student) are fully counterbalanced for each instance of this task. The beliefs & attitudes survey is a multidimensional survey assessing participants' political and ideological beliefs, preferences for policies centering police and prison reform, their empathic traits, and other aspects of their personality (for a complete list of questions see Table S14 or https://osf.io/pqu7d). **C. Frequency word cloud of content descriptions of story stimuli.** A research assistant provided written summaries of the content of each video stimulus during the time of stimulus collection and quality control. To demonstrate the general content and tone of the videos used In this study, we generated a description frequency word cloud of those descriptions. The videos have not been transcribed. This figure is for descriptive purposes only.

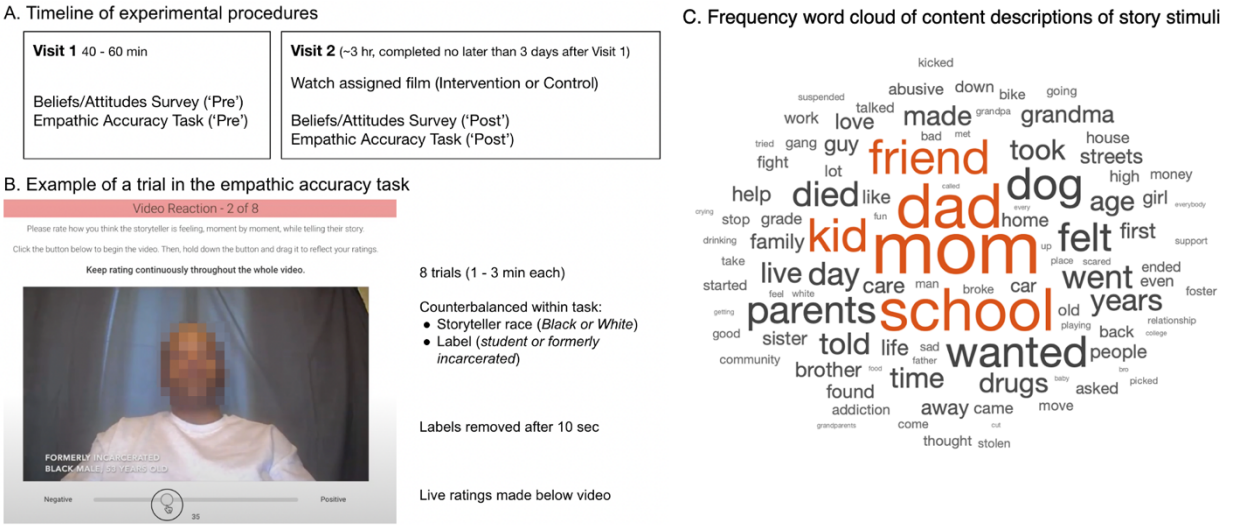

Fig. S3. Visualization of participant demographics and geographic spread. A. Geographic spread of online participants. Most of our participants live in California, Texas, and Florida. **B. Frequency word cloud of participant occupations.** Participants free responded with their current occupation. These responses were cleaned by making them lowercase, fixing typos, and consolidating acronyms and other equivalent terms. Larger font indicates greater frequency. Most participants were either retired, homemakers, not currently employed, students, or on disability. **C. Political ideology of participants**. Participants self-identified their political ideology on a drop-down menu. Most participants were Liberal Democrats (N = 383) followed by Conservative Republicans (N = 218). Participants who identified as Socialist, Anarchist, or Communist were grouped into one category called “Leftist” (N = 61). 80 participants identified as Libertarian, while five participants identified as “Far-Right.” Two participants did not respond. The length of the bar on the bidirectional political spectrum graphic indicates the number of participants in that category.

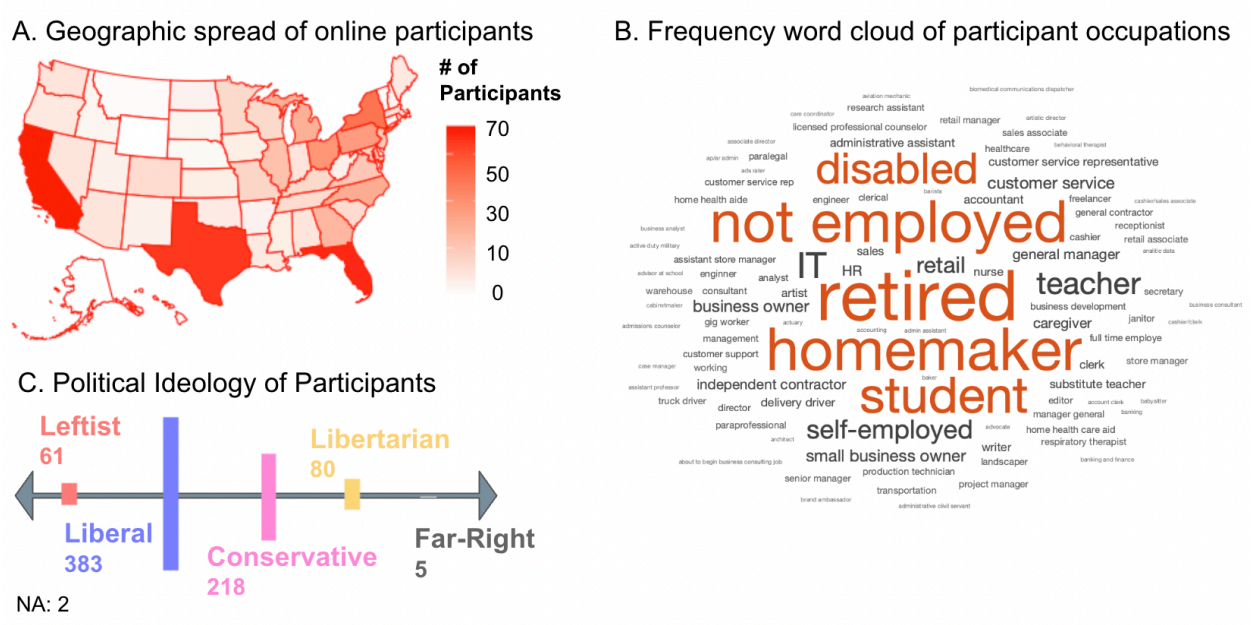

Fig. S4. Distribution of trial-by-trial compassion ratings in empathic accuracy task. All participant trial-by-trial compassion ratings from the empathic accuracy task are plotted in a frequency histogram to assess their distribution. The compassion ratings are skewed to the right indicating that participants (N = 709) report high levels of compassion overall.

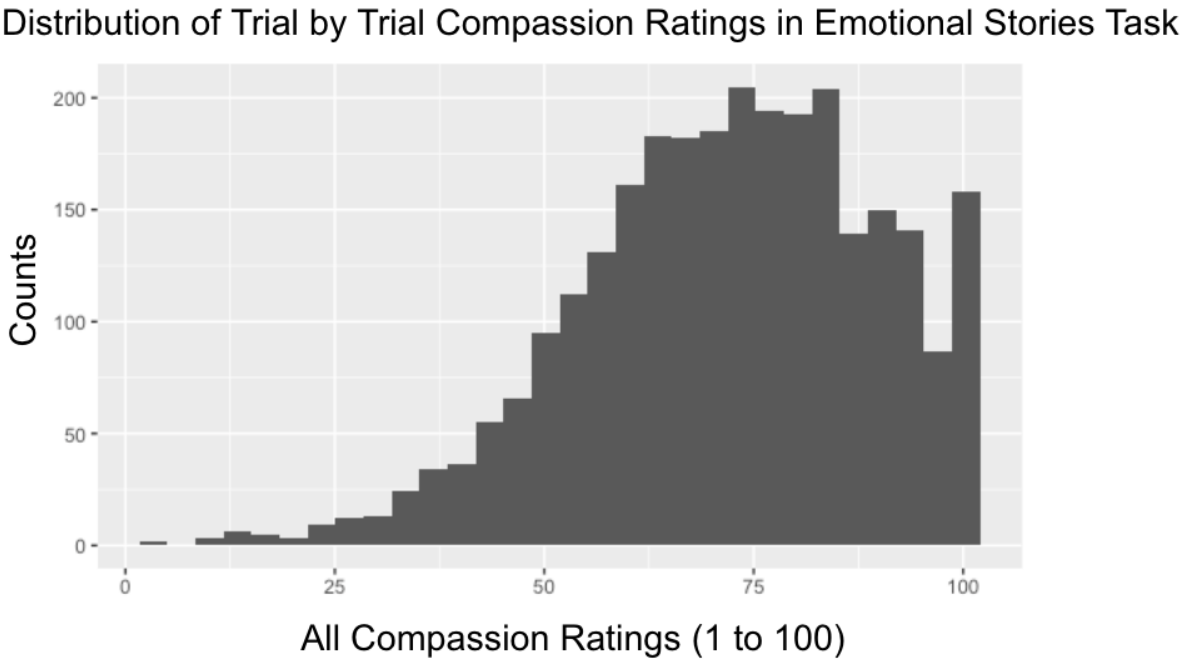

Fig. S5. Changes in empathic accuracy and compassion A. Empathic accuracy. Participants selectively increase their empathic accuracy for formerly incarcerated people (here, indicated by decreased inference error measured via RMSE) after watching the intervention film *Just Mercy* (between subjects $t(694) = 4.01$, $p < 0.001$, Cohen's $d = 0.31$; visit two student vs incarcerated labels *Just Mercy* group within subjects $t(316) = 2.83$, $p = 0.005$, $d = 0.16$; visit one vs. visit two incarcerated labels *Just Mercy* group within subjects $t(318) = 4.10$, $p < 0.001$, $d = 0.25$). Individual participants are plotted as connecting dots to visualize within-participants change. This figure unpacks the results displayed in Figure 1A; however, note that the RMSE values are *not* inverted here so that higher values indicate greater empathic error. See Table S5 for the statistics from the complete model. **B. Self-reported feelings of compassion.** Self-reported feelings of compassion toward individual storytellers were positive, overall, across all participants; however, feelings of compassion decreased slightly post-intervention and post-control, independent of both the storyteller's label ('formerly incarcerated' or 'college student') and the participant's condition assignment (intervention or control; see Table S7 for the corresponding model). Within subject post doc *t*-tests support this (*Just Mercy* group student-label visit one vs. visit two: $t(316) = 2.40$, $p = 0.02$, $d = 0.16$; incarcerated label visit one vs visit two: $t(318) = 2.37$, $p = 0.02$, $d = 0\ .14$; *Control* group student-label visit one vs. visit two: $t(372) = 4.22$, $p < 0.001$, $d = 0.23$); incarcerated-label visit one vs. visit two: $t(364) = 4.36$, $p < 0.001$, $d = 0.23$). This figure unpacks the results displayed in Figure 1B. 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.

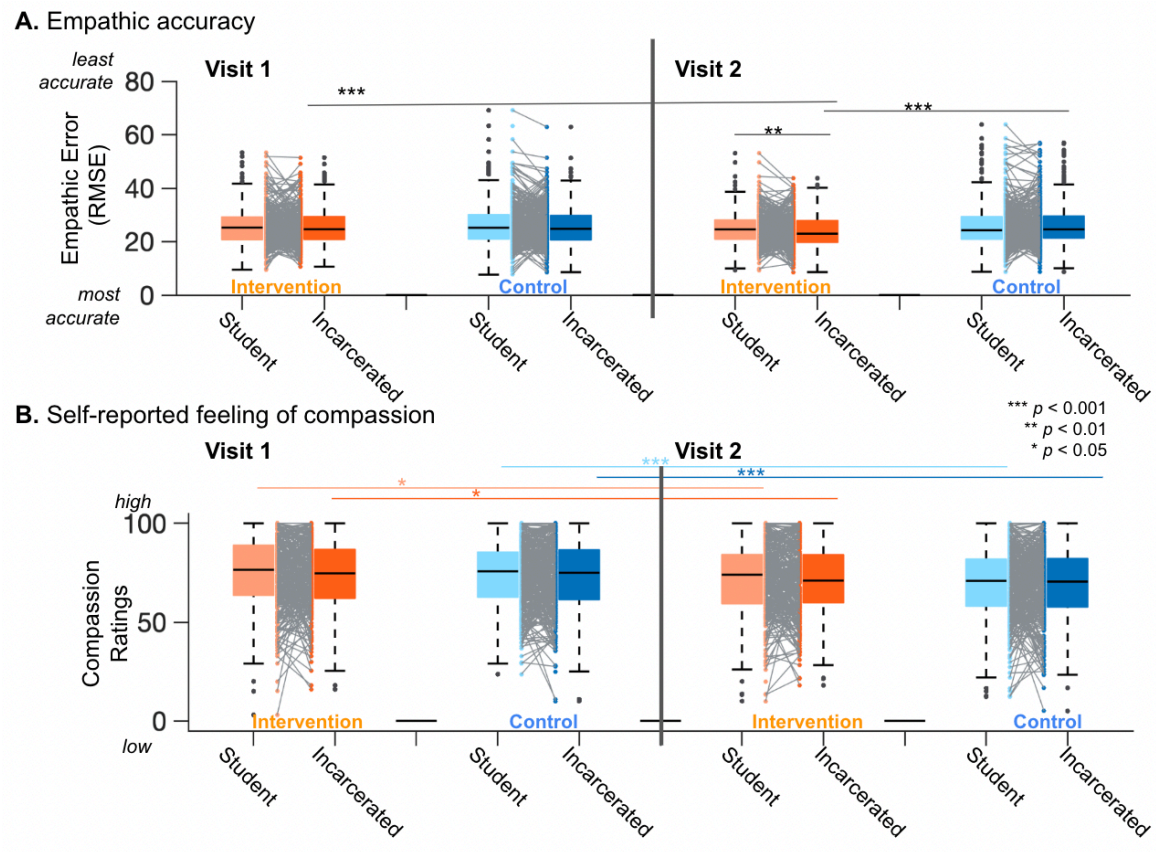

Fig. S6. Exploration of relationships among intervention, empathic accuracy, and support for reform on visit two. An exploratory mediation analysis (1,000 bootstrap samples) was conducted to test whether the film intervention influenced participant attitudes and beliefs by way of their empathic accuracy (measured via RMSE). The predictor (x) in this analysis was condition (intervention vs control). The outcome (y) was support for criminal justice reform on visit two. The mediator was empathic accuracy (M) for formerly incarcerated storytellers on visit two. We found evidence that the effect of film condition was mediated by participant's empathic accuracy towards incarcerated people. That is, the strength of the effect of the intervention on support for reform increased when controlling for error in the empathic accuracy task (see Table S12 for full model).

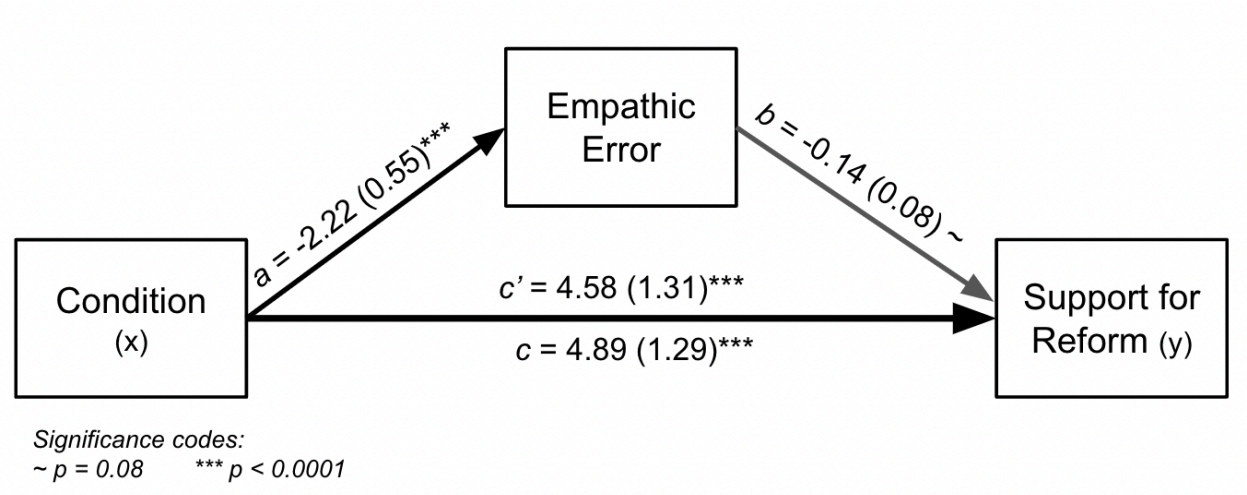

Table S1. Participant Demographics of Complete Sample. Demographic summary of participants (N = 749) who participated in visits one and two and passed the movie quiz (Table S4).

Visit	**Age**	**Gender**	**Race/Ethnicity**	**Education** (highest degree)	**Perceived SES** (*low* 1 - 100 *high*)	**Political Ideology**
Intervention Condition (viewed: *Just Mercy*)						
Visits 1 - 2 N = 344	45.15 SD 15.12	199 Women 134 Men 10 Non-binary 1 NA	3 American Indian/Alaskan Native 23 Asian 47 Black or African American 212 White 25 Hispanic or Latino 33 More than One Race 1 NA	14 Professional degree or higher 30 MA 109 BA 60 Assoc. 115 GED 15 None or Other 1 NA	5.31 SD 1.87	104 Conservative 171 Liberal 33 Other Left-wing 35 Other Right-wing 1 NA
Control Condition (viewed: *Moneyball* or *Concussion*)						
Visits 1 - 2 N = 405	43.53 SD 14.86	212 Women 182 Men 10 Non-binary 1 NA	1 American Indian/Alaskan Native 24 Asian 73 Black or African American 237 White 23 Hispanic or Latino 46 More than One Race 1 NA	27 Professional degree or higher 45 MA 102 BA 79 Assoc. 139 GED 12 None or Other 1 NA	5.30 SD 1.87	113 Conservative 213 Liberal 28 Other Left-wing 50 Other Right-wing 1 NA

Table S2. Participant Demographics at Enrollment. Demographic summary of participants (N = 896) who registered for the study. 16.40% of registered participants dropped out before visit 2 and were, therefore, not included in the main analysis. The general enrollment proportions across stratified demographic variables remains the same from registration to the final sample. See Supplementary Table 3 for significance tests.

Visit	Age	Gender	Race/Ethnicity	Education	Perceived SES	Political Ideology
Intervention Condition (viewed: *Just Mercy*)						
Visits 1 – 2 N = 388	45.75 SD 15.76	217 Women 150 Men 10 Non-binary 1 NA	3 American Indian/Alaskan Native 26 Asian 56 Black or African American 228 White 26 Hispanic or Latino 38 More than One Race 11 NA	17 Professional degree or higher 33 MA 120 BA 65 Assoc. 125 GED 17 None or Other 11 NA	6.78 SD 2.37	119 Conservative 182 Liberal 37 Other Left-wing 39 Other Right-wing 11 NA
Control Condition (viewed: *Moneyball* or *Concussion*)						
Visits 1 – 2 N = 471	44.68 SD 15.20	243 Women 205 Men 10 Non-binary 1 NA	1 American Indian/Alaskan Native 25 Asian 81 Black or African American 270 White 30 Hispanic or Latino 51 More than One Race 13 NA	28 Professional degree or higher 47 MA 117 BA 85 Assoc. 167 GED 14 None or Other 13 NA	6.87 SD 2.39	129 Conservative 240 Liberal 33 Other Left-wing 56 Other Right-wing 13 NA
NA – Never assigned to a condition due to early drop out						
N = 37	Unknown (incomplete data)	1 Woman 1 Man 35 NA	1 Asian 1 More than One Race 35 NA	1 MA 1 None or Other 35 NA	Unknown (incomplete data)	2 Liberal 35 NA

Table S3. Analysis of attrition rates by demographic variables of interest. We tested whether the proportion of drop out from registration to finishing visit two was significantly different between the intervention and control groups within demographic categories of interest. Attrition rates are reported as percent change. The only significant difference in attrition rates between groups was for participants self-identified as Hispanic or Latinx ($\chi 2 = 4.32$, $p = 0.04$). More Hispanic participants dropped out in the control group than the intervention group. This effect was small, and the resulting number of Hispanic participants was balanced. Abbreviation *n.s.* indicates no significant difference.

		Intervention				**Control**				
		Registered	Final	Dropped Out	Attrition Rate	Registered	Final	Dropped Out	Attrition Rate	**Sig.**
Gender	Woman	217	199	18	-8.29	243	212	31	-12.76	n.s.
	Man	150	134	16	-10.67	205	182	23	-11.22	n.s.
	Non-binary	10	10	0	0.00	10	10	0	0.00	n.s.
Race	Black	56	47	9	-16.07	81	73	8	-9.88	n.s.
	White	228	212	16	-7.02	270	237	33	-12.22	n.s.
	Hispanic	26	25	1	-3.85	30	23	7	-23.33	$\chi 2 = 4.32$, $p = 0.04$
	Asian	26	23	3	-11.54	25	24	1	-4.00	n.s.
	American Indian or Alaskan Native	3	3	0	0.00	1	1	0	0.00	n.s.
	More than One Race	38	33	5	-13.16	51	46	5	-9.80	n.s.
Political Ideology	Liberal	182	171	11	-6.04	240	213	27	-11.25	n.s.
	Conservative	119	104	15	-12.61	129	113	16	-12.40	n.s.
	Other Left	37	33	4	-10.81	33	28	5	-15.15	n.s.
	Other Right	39	35	4	-10.26	56	50	6	-10.71	n.s.

Table S4. Film comprehension questions for manipulation check. All participants completed a quiz after watching their assigned film to ensure they attended to the film. These questions were designed to be easy to answer if you had watched the movies, but not something that could be easily 'googled.'

Question	Answer
Just Mercy	
What was Walter McMillian doing when the police arrested him?	Driving
What law school did Bryan Stevenson attend?	Harvard
What crime was Walter McMillan accused of committing?	Murder
What illness was fellow death row inmate Herbert Richardson diagnosed with?	PTSD
What news program was McMillan's case featured on?	60 Minutes
What happened to McMillan at the end of the movie?	Released from prison
Moneyball	
Who do the A's lose to in the opening scene?	Yankees
Why are the A's losing players?	Not enough money
What school was recruiting Billy Beane to play baseball?	Stanford
What instrument does Billy Beane's daughter play?	guitar
How many consecutive games did the A's win?	20
Concussion	
What schools did Dr. Omalu attend?	All of the Above
What did Dr. Omalu throw out after each autopsy that his boss asked him to reuse?	Knives
What is the name of the first NFL player Dr. Omalu performs an autopsy on?	Mike Webster
What's the name of the brain disorder that Dr. Omalu thinks is killing NFL players?	CTE
When did the NFL host its Concussion Summit?	After NY Times article
What job did Dr. Omalu get offered at the end of the movie?	US Surgeon General

Table S5. Film intervention enhances participant's empathic accuracy when listening to the stories of formerly incarcerated men. We hypothesized that watching the intervention film, *Just Mercy*, would selectively increase one's ability to infer the emotions of formerly incarcerated storytellers; therefore, we used an LME model to test for a time (pre vs post-film) by condition (intervention vs control) by label ('formerly incarcerated' vs 'student') interaction on empathic accuracy (RMSE) in the empathic accuracy task. Participants were input as random effects. Time, condition, and label were input as fixed effects. This analysis revealed main effects of time and condition on empathic accuracy, and importantly, a significant time by condition by label interaction. Missing trial data were treated with NAs. Overall, these results indicate that empathic accuracy for formerly incarcerated people is influenced by the intervention film (see Figure 1A and Figure S5A).

Formula: RMSE ~ Time * Condition * Label + (1 | ID)
REML criterion at convergence: 18579.2

Random effects:

Groups	Name	Variance	Std. Dev.
ID	(Intercept)	21.46	4.633
Residual		33.27	5.768

Number of observations: 2788, *groups:* ID (N = 709)

Fixed effects:

	Estimate	Std. Error	*df*	*t* value	*p* val
(Intercept)	25.55	0.21	693.31	123.89	< 2e-16 ***
Time	0.46	0.16	2066.60	2.99	0.003 **
Condition	-0.69	0.29	693.31	-2.37	0.018 *
Label	0.24	0.16	2069.90	1.58	0.115
Time*Condition	0.55	0.22	2066.60	2.49	0.013 *
Time*Label	-0.08	0.22	2068.38	-0.37	0.714
Condition*Label	0.27	0.22	2069.90	1.23	0.220
Time* Condition*Label	-0.65	0.31	2068.38	-2.08	0.038 *

---

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table S6. The film intervention increases empathic accuracy towards formerly incarcerated people above and beyond participant social and political identities. The LME model from Table S5 was expanded to control for participant race, gender, socioeconomic status, and political ideology. Of the 709 participants, 74 did not give a response on the SES ladder so only 635 participants are included in this analysis. This analysis was run to test the robustness of the time by condition by label interaction on empathic accuracy, above and beyond relevant demographic variables. There were 10 levels of SES (via MacArthur Ladder), 3 levels of gender (male, female, & other), 6 levels of race (Asian, Black, Hispanic, Native American, White, or More than One Race), and 4 levels of political ideology (liberal, conservation, 'other left,' and 'other right'). These demographic variables were collected via voluntary self-report before participants began the experiment. The time by condition by label interaction held, marginally (p = 0.05), as did the main effects of time and condition. This marginal interaction was in the expected direction and was, additionally, a pre-registered analysis. There were no linear main effects of SES, Gender, Race, or Political Ideology.

Formula: RMSE ~ Time * Condition * Label + SES + gender3 + race6 + politics4 + (1 | ID)
REML criterion at convergence: 16556.3

Random effects:

Groups	Name	Variance	Std. Dev.
ID	(Intercept)	18.47	4.297
Residual		33.38	5.778

Number of observations: 2500, *groups:* ID (N = 635)

Fixed effects:

	Estimate	Std. Error	df	*t* value	P val
(Intercept)	26.93082	0.74974	604.54646	35.920	< 2e-16 ***
Time	0.48063	0.16394	1857.49228	2.932	0.00341 **
Condition	-0.73076	0.29662	606.03542	-2.464	0.01403 *
Label	0.29456	0.16405	1860.94358	1.796	0.07272 .
SES.L	1.60292	1.70989	611.95873	0.937	0.34890
SES.Q	3.44998	1.64334	612.04807	2.099	0.03619 *
SES.C	0.56413	1.46235	611.90706	0.386	0.69980
SES^4	0.82351	1.36126	611.85194	0.605	0.54543
SES^5	0.09628	1.25767	610.59373	0.077	0.93900
SES^6	1.13905	1.04047	611.07174	1.095	0.27406
SES^7	0.35329	0.79139	611.05398	0.446	0.65545
SES^8	0.81321	0.59747	608.44594	1.361	0.17399
SES^9	-0.16887	0.52249	605.44722	-0.323	0.74665
Gender.L	-0.50224	0.30553	606.38672	-1.644	0.10072
Gender.Q	1.06344	1.05479	603.39424	1.008	0.31376
Race.L	0.87322	1.22926	601.21904	0.710	0.47775
Race.Q	-0.53206	0.73702	604.71366	-0.722	0.47063
Race.C	-2.92075	1.65956	600.41653	-1.760	0.07892 .
Race^4	-4.57253	1.79337	600.27519	-2.550	0.01103 *
Race^5	-2.49001	1.16334	601.28346	-2.140	0.03272 *
Politics.L	0.07246	0.52593	605.83601	0.138	0.89046
Politics.Q	0.62021	0.56609	604.81412	1.096	0.27368
Politics.C	0.89456	0.59615	603.07185	1.501	0.13399
Time*Condition	0.51293	0.23185	1857.45774	2.212	0.02706 *

Time*Label	-0.13083	0.23193	1859.44652	-0.564	0.57276
Condition*Label	0.22756	0.23201	1860.83515	0.981	0.32680
Time*Cond*Label	-0.64309	0.32800	1859.44611	-1.961	0.05007 .

---

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table S7. Film intervention has no effect on feelings of compassion for formerly incarcerated storytellers. We hypothesized that watching the intervention film, *Just Mercy,* would selectively increase one's compassion for formerly incarcerated individuals; therefore, we used an LME model to test for a time (pre vs post-film) by condition (intervention vs control) by label ('formerly incarcerated' vs 'college student') interaction on self-reported compassion for storytellers in the empathic accuracy task. Participants were input as random effects. Time, condition, and label were input as fixed effects. No interaction was found. Furthermore, there was no main effect of condition or label. There was a main effect of time, where participants' feelings of compassion decreased moderately, overall, during the second visit, independent of the film they watched. However, it is important to note that participants reported high levels of compassion, overall, and that compassion ratings were generally skewed to the right (see Figure S4). Plots of these data can be found in Figures 1B and S5B.

Formula: Compassion ~ Time * Condition * Label + (1 | ID)
REML criterion at convergence: 23194.3

Random effects:

Groups	Name	Variance	Std. Dev.
ID	(Intercept)	136.9	11.70
Residual		167.3	12.94

Number of observations: 2788, *groups:* ID (N = 709)

Fixed effects:

	Estimate	Std. Error	*df*	*t* value	*p* val
(Intercept)	72.00	0.51	707.08	142.56	< 2e-16 ***
Time	2.34	0.35	2079.36	6.71	2.44e-11 ***
Condition	0.85	0.71	707.08	1.19	0.23
Label	0.57	0.35	2082.29	1.62	0.10
Time*Condition	-0.63	0.49	2079.36	-1.29	0.20
Time*Label	-0.02	0.49	2080.94	-0.04	0.97
Condition*Label	0.19	0.49	2082.29	0.38	0.71
Time*Condition*Label	0.21	0.70	2080.94	0.31	0.76

---

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table S8. Individual item summary of beliefs and attitudes change toward the US criminal justice system. 20 unique questions were combined into the overall measure of "support for criminal justice reform" in Figure 2A. These unique questions are itemized here along with their within-subject (pre vs. post) comparisons between the intervention and control groups. Questions were presented in a randomized order for each participant on Qualtrics. The raw responses participants made were analyzed without transformation or weighting. All featured statistics survive multiple comparisons correction (FDR; $q < 0.05$). Red indicates that support for the item decreased after the intervention. Green indicates that support for the item increased after the intervention. Effect sizes larger than Cohen's $d = 0.30$ are bolded. For the intervention group, all significant changes moved in the "progressive" or "pro-reform" direction. There were 316 participants in the intervention group and 360 participants in the control group because 73 participants had missing survey data during visit one due to not completing the survey in time before it closed. Participants were not removed entirely from analysis because they had partially completed the survey which had many items beyond the criminal justice reform items.

Question *Rating Scale: 0 (Strongly Disagree) - 100 (Strongly Agree)*	**Intervention Group** (*Pre* minus *Post* change)	**Control Group** (*Pre* minus *Post* change)
The death penalty should be an option for the most serious crimes.	$t(315) = 9.03$, $p < 0.001$, CI = [8.91, 13.88], **Cohen's *d* = 0.34**	*n.s.*
Formerly incarcerated individuals should not have to face legal barriers to accessing housing	$t(315) = -4.49$, $p < 0.001$, CI = [-6.50, -2.54], Cohen's $d = -0.15$	*n.s.*
Though the United States prison system has its problems, it is ultimately a fair system that should not be changed.	*n.s.*	*n.s.*
There are too many people in prison right now in the United States.	$t(315) = -5.42$, $p < 0.001$, CI = [-7.57, -3.53], Cohen's $d = -0.22$	*n.s.*
Formerly incarcerated individuals should have the right to vote.	*n.s.*	*n.s.*
After being released from prison, people should have their criminal conviction expunged (erased in the eyes of the law).	$t(315) = -5.64$, $p < 0.001$, CI = [-7.54, -3.64], Cohen's $d = -0.20$	*n.s.*
Whether or not you have been incarcerated should not affect your ability to get a job.	$t(315) = -3.80$, $p < 0.001$, CI = [-5.68, -1.80], Cohen's $d = -0.16$	*n.s.*
We should use our tax money to fund educational programs for those in prisons.	$t(315) = -3.81$, $p < 0.001$, CI = [-5.08, -1.62], Cohen's $d = -0.15$	*n.s.*

I often feel sympathy for incarcerated people.	*t*(315) = -8.75, p < 0.001, CI = [-9.80, -6.20], **Cohen's *d* = -0.33**	*n.s.*
It is unethical for corporations to profit off the labor of incarcerated people.	*t*(315) = -3.63, p < 0.001, CI = [-6.47, -1.92], Cohen's *d* = -0.15	*n.s.*
Incarcerated people, including those convicted of committing serious crimes, are capable of transformation.	*t*(315) = -3.79, p < 0.001, CI = [-5.20, -1.64], Cohen's *d* = -0.17	*n.s.*
In general in our country these days, Black people are treated less fairly than White people by the criminal justice system.	*t*(315) = -4.70, p < 0.001, CI = [-5.67, -2.33], Cohen's *d* = -0.14	*n.s.*
In general in our country these days, low-income defendants are treated less fairly than wealthy defendants by the criminal justice system.	*n.s.*	*n.s.*
I think someone convicted of a drug crime should be given access to rehabilitation facilities, not prison time.	*n.s.*	*n.s.*
Most police can be trusted.	*t*(315) = 3.89, p < 0.001, CI = [1.76, 5.37], Cohen's *d* = 0.12	*n.s.*
The police in the United States are too militarized.	*t*(315) = -6.45, p < 0.001, CI = [-7.28, -3.88], Cohen's *d* = -0.15	t(359) = -2.34, p = 0.02, CI = [-3.30, -0.28], Cohen's *d* = -0.06
America should defund the police and reallocate that spending to mental health care and education.	*t*(315) = -2.45, p = 0.01, CI = [-4.02, -0.44], Cohen's *d* = -0.08	t(359) = -2.41, p = 0.02, CI = [-4.10, -0.41], Cohen's *d* = -0.07
The police are the best point of contact for mental health and substance use emergencies	*n.s.*	*t*(359) = -2.44, p = 0.02, CI = [-4.06, -0.44], Cohen's *d* = -0.10
The police keep all communities, regardless of race or social status, safer.	*t*(315) = 2.79, p = 0.006, CI = [0.87, 5.05], Cohen's *d* = 0.10	*t*(359) = -2.65, p = 0.008, CI = [-4.13, -0.62], Cohen's *d* = -0.09
If I was in an emergency situation, I trust that the police would help me.	*t*(315)=5.41, p < 0.001, CI = [2.46, 5.27], Cohen's *d* = 0.13	*n.s.*

**p values corrected for multiple comparisons FDR q < 0.05*

Table S9. Film intervention increases participant support for criminal justice reform. We hypothesized that watching the intervention film, *Just Mercy*, would increase participant's support of widespread criminal justice reform; therefore, we used an LME model to test for a time (pre vs post-film) by condition (intervention vs control) interaction on the combined 'support for reform' metric derived from the Beliefs & Attitudes Survey (see Table S8). Participants were input as random effects. Time and condition were input as fixed effects. This analysis revealed main effects of time and condition and, importantly, a significant time by condition interaction. Post hoc *t*-tests revealed significant pre vs post pairwise differences in the intervention condition alone, as well as a significant between group difference in visit two. Degrees of freedom (df) fluctuate due to missing survey data. 73 participants had missing survey data during visit one due to not completing the survey in time before it closed, though they had started the survey and completed the empathic accuracy task and movie watching intervention. Participants were not removed entirely from analysis because they had partially completed the survey which had many items beyond the criminal justice reform items. Overall, these results indicate that support for criminal justice reform is influence by the intervention film (see Figure 2).

Formula: Support ~ Time* Condition + (1 | ID)
REML criterion at convergence: 10486.3

Random effects:

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	270.82	16.457
Residual		13.53	3.678

Number of obs: 1425, *groups:* ID, (N = 749)

Fixed effects:

	Estimate	Std. Error	df	*t* value	*p* val
(Intercept)	61.8983	0.6116	747.0979	101.213	< 2e-16 ***
Time	-1.4783	0.1416	676.9296	-10.439	< 2e-16 ***
Condition	1.8522	0.8649	747.0979	2.142	0.0325 *
Time*Condition	-1.6862	0.2003	676.9296	-8.419	2.26e-16 ***

Post hoc *t* tests:

	df	t	CI	Cohen's d	p value
within subjects t-test					
Intervention: Pre vs Post	315	-12.01	[-4.39, -3.16]	-0.23	< 2e-16 ***
Control: Pre vs Post	359	-1.57	[-0.90, 0.10]	-0.03	0.1182
between subjects t-test					
Pre: Intervention vs Control	675	0.79	[-1.51, 3.52]	0.06	0.4321
Post: Intervention vs Control	747	3.44	[1.85, 6.76]	0.25	0.0006 **

---

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table S10. The film intervention increases support for criminal justice reform above and beyond participant social and political identities. The LME model from Table S9 was expanded to control for participant race, gender, socioeconomic status, and political ideology. This analysis was run to test the robustness of the time by condition interaction on support for reform above and beyond relevant demographic variables. There were 10 levels of SES (via MacArthur Ladder), 3 levels of gender (male, female, & other), 6 levels of race (Asian, Black, Hispanic, Native American, White, or More than one race), and 4 levels of political ideology (liberal, conservation, 'other left,' and 'other right'). These demographic variables were collected via voluntary self-report. 76 participants did not on the SES ladder and 2 did not complete the demographic questionnaire, so the sample size is reduced to N = 671 when NaNs are eliminated. The time by condition interaction held, as did the main effects of time and condition. There were no linear main effects of SES, gender, or race. However, there is a significant linear effect political ideology (see Table S11).

Formula: Support ~ Time * Condition + SES + gender3 + race6 + politics4 + (1 | ID)
REML criterion at convergence: 9446.5

Random effects:

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	178.40	13.357
Residual		13.57	3.683

Number of obs: 1342, *groups:* ID, (N = 671)

Fixed effects:

	Estimate	Std. Error	df	t value	P val
(Intercept)	67.0816	1.8326	649.9998	36.605	< 2e-16 ***
Time	-1.4911	0.1425	669.0000	-10.466	< 2e-16 ***
Condition	1.8389	0.7555	650.0000	2.434	0.01520 *
SES.L	-4.5460	4.3317	650.0000	-1.049	0.29436
SES.Q	10.2171	4.2549	650.0000	2.401	0.01662 *
SES.C	-0.6194	3.7912	650.0000	-0.163	0.87027
SES^4	0.7400	3.3626	650.0000	0.220	0.82589
SES^5	1.3991	2.9615	650.0000	0.472	0.63678
SES^6	-0.7598	2.4210	650.0000	-0.314	0.75374
SES^7	1.7194	1.8944	650.0000	0.908	0.36442
SES^8	-1.6268	1.4880	650.0000	-1.093	0.27467
SES^9	1.2840	1.3292	650.0000	0.966	0.33440
Gender.L	0.1625	0.7786	650.0000	0.209	0.83472
Gender.Q	-7.8635	2.7541	649.9999	-2.855	0.00444 **
Race.L	0.4990	2.8646	650.0000	0.174	0.86176
Race.Q	-2.0348	1.8473	650.0000	-1.101	0.27109
Race.C	3.0293	3.8177	650.0000	0.793	0.42778
Race^4	0.3598	4.1198	650.0000	0.087	0.93043
Race^5	2.4371	2.7686	649.9999	0.880	0.37903
Politics.L	8.1598	1.3421	650.0000	6.080	2.05e-09 ***
Politics.Q	-17.4913	1.4314	649.9999	-12.220	< 2e-16 ***
Politics.C	-4.0796	1.4949	650.0000	-2.729	0.00653 **
Time*Condition	-1.6965	0.2015	669.0000	-8.419	2.30e-16 ***

---

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table S11. After watching the intervention film participants equally increased their support for criminal justice reform regardless of their political ideology. A LME model modeling participant intercepts as random effects was run to test for interactions among political ideology (four levels) and time (pre v post) in the intervention group alone. No interactions were found.

Formula: Support ~ Time* politics4 + (1 | ID)
REML criterion at convergence: 4720.9

Random effects:

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	174.75	13.219
Residual		15.66	3.957

Number of obs: 658, groups: ID, 343

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	65.58	0.92	340.41	71.03	< 2e-16 ***
TimePoint.L	2.62	0.29	314.15	9.17	< 2e-16 ***
politics4.L	-16.52	2.24	340.64	-7.38	1.19e-12 ***
politics4.Q	12.70	1.85	340.41	6.88	2.95e-11 ***
politics4.C	5.90	1.35	339.76	4.38	1.61e-05 ***
TimePoint.L:politics4.L	0.14	0.70	314.32	0.20	0.85
TimePoint.L:politics4.Q	-0.07	0.57	314.15	-0.12	0.90
TimePoint.L:politics4.C	0.40	0.41	313.66	0.96	0.34

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Degrees-of-freedom method: Kenward-roger

Table S12. Results table for follow up exploratory mediation analysis.
We sought to understand how participants' behavior in the empathic accuracy task was related to their support for criminal justice reform during visit two (after the films; see Figure S6). We found marginal evidence of a partial mediation of the effect of condition by performance on the empathic accuracy task (measured via RMSE) on support for reform. Partial mediation is determined by Baron & Kenny's *4 Steps* (Baron & Kenny, 1986). First, path c' is significant, that is, condition (X) is correlated with the outcome (Y). Second, condition (X) is significantly correlated with empathic error (M). Third, empathic error (M) *marginally* affects support for reform (Y) when controlling for X (p = 0.08) in the expected direction (less empathic error, more support for reform). Finally, there was a significant and large total effect (c) of condition on support for reform through empathic error. Note, empathic error data from visit two were missing from four people in the intervention condition and nine people in the control condition due to technical errors in the online video playback, so of the N = 709 participants with useable visit one or visit two empathic accuracy task data there are N = 696 participants in this analysis.

Regressions:

	Estimate	Std.Err	z-value	*p val*	Std.lv	Std.all
Support ~						
EmpathicError (b)	-0.141	0.080	-1.766	0.077	-0.141	-0.059
Conditon (c')	4.580	1.312	3.491	0.000***	4.580	0.131
EmpathicError ~						
Condition (a)	-2.220	0.551	-4.032	0.000***	-2.220	-0.151

Variances:

	Estimate	Std.Err	z-value	*p val*	Std.lv	Std.all
.Support	296.914	14.953	19.857	0.000***	296.91	0.98
.EmpathicError	52.859	3.748	14.104	0.000***	52.86	0.98

R-Square:

	Estimate
Support	0.023
EmpathicError	0.023

Defined Parameters:

	Estimate	Std.Err	z-value	*p val*	Std.lv	Std.all
Cond_EmpathErr_y	0.312	0.192	1.628	0.104	0.312	0.009
Total (c)	4.892	1.292	3.787	0.000***	4.892	0.140

---
*Significance codes: '***'0.001, '**'0.01, '*' 0.05*

Table S13. Petition Questions. The prompts above include the exact wording of the petition participants were asked if they would sign and share. These questions appeared only at the end of visit two after participants completed the support for reform questions. The petition questions appeared in this order where they were asked to sign first and to share second.

Please read the following petition and respond to the questions below. As of 2020, an estimated 5.2 million Americans still cannot vote because of a felony conviction. Felony disenfranchisement rates vary by state, as states institute a wide range of disenfranchisement policies. Currently, only the District of Columbia, Maine, and Vermont do not restrict voting based on a felony conviction. Meanwhile, in 11 states, voting rights are indefinitely denied to some or all of the individuals who have successfully fulfilled their prison, parole, or probation sentences. Overall, the discrepancy in felony disenfranchisement policies across the United States disproportionately impacts communities of color. Over the years, some states have changed their laws and practices to expand voting access to people with felony convictions, but a federal law would help eliminate discrepancies between states. Let your representatives know that you want a federal law to end restrictive voting due to a criminal conviction. Help support voting rights for all people! Would you be willing to sign this petition to indiciate your support?
Please read the following petition and respond to the questions below. As of 2020, an estimated 5.2 million Americans still cannot vote because of a felony conviction. Felony disenfranchisement rates vary by state, as states institute a wide range of disenfranchisement policies. Currently, only the District of Columbia, Maine, and Vermont do not restrict voting based on a felony conviction. Meanwhile, in 11 states, voting rights are indefinitely denied to some or all of the individuals who have successfully fulfilled their prison, parole, or probation sentences. Overall, the discrepancy in felony disenfranchisement policies across the United States disproportionately impacts communities of color. Over the years, some states have changed their laws and practices to expand voting access to people with felony convictions, but a federal law would help eliminate discrepancies between states. Let your representatives know that you want a federal law to end restrictive voting due to a criminal conviction. Help support voting rights for all people! Would you share this petition with friends and family?

Table S14. All questions on the beliefs and attitudes survey at timepoint 1. The questions above appeared in a random order to all participants at several timepoints throughout the experiment. In this manuscript, only the questions related to prison and policing were analyzed, however these questions were administered across this larger survey that contained components of many other questionnaires.

Question	Profile Dimension	Source
There is a person in my life with whom I can share my joys and sorrows.	Perceived Social Support	Adapted from Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet & Farley, 1988)
I can count on my friends or family when things go wrong.	Perceived Social Support	Adapted from Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet & Farley, 1988)
There is a person in my life who cares about my feelings.	Perceived Social Support	Adapted from Multidimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet & Farley, 1988)
I daydream and fantasize, with some regularity, about things that might happen to me.	Social Imagination / Fantasy	IRI - Fantasy
I often have concerned feelings for people less fortunate than me.	Empathic Concern	IRI - Empathic Concern
I really get involved with the feelings of the characters in a novel, TV show, or film.	Social Imagination / Fantasy	IRI - Fantasy
In emergency situations, I feel apprehensive and ill-at-ease.	Personal Distress	IRI - Personal Distress
I am usually objective when I watch a movie or play, and I don't often get completely caught up in it.*	Social Imagination / Fantasy	IRI - Fantasy
I try to look at everybody's side of a disagreement before I make a decision.	Perspective Taking	IRI - Perspective Taking
When I see someone being taken advantage of, I feel kind of protective towards them.	Empathic Concern	IRI - Empathic Concern
I sometimes feel helpless when I am in the middle of a very emotional situation.	Personal Distress	IRI - Personal Distress
When I see someone get hurt, I tend to remain calm.*	Personal Distress	IRI - Personal Distress
If I'm sure I'm right about something, I don't waste much time listening to other people's arguments.*	Perspective Taking	IRI - Perspective Taking
I would describe myself as a pretty soft-hearted person.	Empathic Concern	IRI - Empathic Concern
Before criticizing somebody, I try to imagine how I would feel if I were in their place.	Perspective Taking	IRI - Perspective Taking
I often comfort someone I know after they experience a hardship.	Prosociality	Baumsteiger & Siegel (2018) Measuring Prosociality: The Development of a Prosocial Behavioral Intentions Scale

I would mentor a younger or less experienced person	Prosociality	Baumsteiger & Siegel (2018) Measuring Prosociality: The Development of a Prosocial Behavioral Intentions Scale
I would assist a stranger with a small task like carrying groceries.	Prosociality	Baumsteiger & Siegel (2018) Measuring Prosociality: The Development of a Prosocial Behavioral Intentions Scale
I would help care for a sick friend or relative.	Prosociality	Baumsteiger & Siegel (2018) Measuring Prosociality: The Development of a Prosocial Behavioral Intentions Scale
I accept others even when they do things I think are wrong.	Altruism	
People should be willing to help others who are less fortunate.	Altruism	
Those in need have to learn to take care of themselves and not depend on others.	Altruism	
I am more economically privileged than most people in the United States.	Subjective SES	Adapted from The MacArthur Network on SES and Health
I have a very high standing in my workplace or community.	Subjective SES	Adapted from The MacArthur Network on SES and Health
If I got sick I would be able to access quality care.	Subjective SES	Adapted from The MacArthur Network on SES and Health
I do not worry about being able to make my rent or mortgage payments.	Subjective SES	Adapted from The MacArthur Network on SES and Health
If I get into financial trouble I can rely on a friend or family member to help me out.	Subjective SES	Adapted from The MacArthur Network on SES and Health
I often hire someone to do household tasks like cleaning, grocery shopping, or cooking (take out).	Subjective SES	Adapted from The MacArthur Network on SES and Health
I think people perform better when there is competition.	Political Ideology - Economic Scale (X)	
Unions are important for workers to protect their interests.	Political Ideology - Economic Scale (X)	
Income inequality is a major problem today.	Political Ideology - Economic Scale (X)	
Rich people should pay higher taxes.	Political Ideology - Economic Scale (X)	
Billionaires in America earned their wealth through hard work and effort.	Political Ideology - Economic Scale (X)	
Corporations should have more regulations placed on them.	Political Ideology - Economic Scale (X)	
All employees in a workplace should have a say in how things are run at their job.	Political Ideology - Economic Scale (X)	
Health care is a human right and should be free for everyone.	Political Ideology -	

	Economic Scale (X)	
Education, including higher education, is a human right and should be free and available for everyone.	Political Ideology - Economic Scale (X)	
Welfare makes people lazy.	Political Ideology - Economic Scale (X)	Attitudes Toward Poverty
Homelessness results primarily from an unfair and unaffordable housing market, and is not the fault of those who are homeless.	Political Ideology - Economic Scale (X)	
I would like to live in a society that prioritizes the welfare of people over the profits of corporations.	Political Ideology - Economic Scale (X)	
People are ultimately divided more by class than by race.	Political Ideology - Economic Scale (X)	adapted from political compass
Every person should contribute what they are able to society, and society should provide whatever it is that people need.	Political Ideology - Economic Scale (X)	from The Political Compass
Those with the ability to pay should have access to higher standards of medical care.	Political Ideology - Economic Scale (X)	from The Political Compass
In general, I think society is fair.	General System Justification	adapted from General System Justification
In general, I think the American system operates as it should.	General System Justification	adapted from General System Justification
The United States is the best country in the world to live in.	General System Justification	adapted from General System Justification
Everyone has a fair shot at wealth and happiness.	General System Justification	adapted from General System Justification
Society is set up so that people usually get what they deserve.	General System Justification	adapted from General System Justification
If people work hard, they almost always get what they want.	Economic Sys Justification	adapted from Economic System Justification
It is virtually impossible to eliminate poverty.	Economic Sys Justification	adapted from Economic System Justification
Most people who don't get ahead in our society should not blame the system; they have only themselves to blame.	Economic Sys Justification	adapted from Economic System Justification
Equal distribution of resources is a possibility for our society.	Economic Sys Justification	adapted from Economic System Justification
There are no inherent differences between rich and poor; it is purely a matter of circumstances into which you are born.	Economic Sys Justification	adapted from Economic System Justification

Some groups of people are simply inferior to other groups.	Social Dominance	adapted from Social Dominance Orientation (SDO-7; Ho et al., 2015)
I do not trust the mainstream media.	Trust	
American politicians care about my interests and well-being.	Trust	
People usually use fair procedures in dealing with others.	Trust	Belief in a Just World
I feel that people generally earn the rewards and punishments that they get in this world.	Trust	Belief in a Just World
I usually receive the outcomes that I deserve.	Trust	Belief in a Just World
I am generally subjected to processes that are fair.	Trust	Belief in a Just World
Most people can be trusted.	Trust	
Your intelligence is something about you that you can't really change.	Beliefs - Capacity for change	Growth Mindset Scale
Everyone, no matter who they are, can significantly change their status in society.	Beliefs - Social Mobility	Growth Mindset Scale
People can do things differently, but the important parts of who they are can't really be changed.	Beliefs - Capacity for change	"Kind of Person" Implicit Theory Scale
People can always substantially change the kind of person they are.	Beliefs - Capacity for change	"Kind of Person" Implicit Theory Scale
I'd always support my country, whether it was right or wrong.	Political Ideology - State Power (Y)	from The Political Compass
All authority should be questioned.	Political Ideology - State Power (Y)	from The Political Compass
Possessing illicit drugs for personal use should not be a criminal offense.	Political Ideology - State Power (Y)	from The Political Compass
The death penalty should be an option for the most serious crimes.	Beliefs - US Prison System	from The Political Compass
Formerly incarcerated individuals should not have to face legal barriers to accessing housing	Beliefs - US Prison System	
Though the United States prison system has its problems, it is ultimately a fair system that should not be changed.	Beliefs - US Prison System	
There are too many people in prison right now in the United States.	Beliefs - US Prison System	
Formerly incarcerated individuals should have the right to vote.	Beliefs - US Prison System	
After being released from prison, people should have their criminal conviction expunged (erased in the eyes of the law).	Beliefs - US Prison System	
Whether or not you have been incarcerated should not affect your ability to get a job.	Beliefs - US Prison System	
We should use our tax money to fund educational programs for those in prisons.	Beliefs - US Prison System	

I often feel sympathy for incarcerated people.	Beliefs - US Prison System	
It is unethical for corporations to profit off the labor of incarcerated people.	Beliefs - US Prison System	
Incarcerated people, including those convicted of committing serious crimes, are capable of transformation.	Beliefs - US Prison System	
In general in our country these days, Black people are treated less fairly than White people by the criminal justice system.	Beliefs - US Prison System	
In general in our country these days, low-income defendants are treated less fairly than wealthy defendants by the criminal justice system.	Beliefs - US Prison System	
I think someone convicted of a drug crime should be given access to rehabilitation facilities, not prison time.	Beliefs - US Prison System	
If I was in an emergency situation, I trust that a stranger would help me.	Perceived Social Support / Trust	
If I was in an emergency situation, I trust that a stranger of a different race would help me.	Perceived Social Support / Race	
Most police can be trusted.	Beliefs - US Policing	
The police in the United States are too militarized.	Beliefs - US Policing	
America should defund the police and reallocate that spending to mental health care and education.	Beliefs - US Policing	
The police are the best point of contact for mental health and substance use emergencies	Beliefs - US Policing	
The police keep all communities, regardless of race or social status, safer.	Beliefs - US Policing	
If I was in an emergency situation, I trust that the police would help me.	Beliefs - US Policing	
I am extroverted, that is, I am outgoing, energetic, and friendly.	Personality - Big 5	
I am agreeable, warm, and considerate.	Personality - Big 5	
I am conscientious, responsible, dependable, and goal-oriented.	Personality - Big 5	
I am moody, tense, and lack self-confidence.	Personality - Big 5	
I am open to new ideas and experiences.	Personality - Big 5	
Corporations should have more regulations placed on them.	Status Anxiety	
I find it difficult to allow myself to depend on others.	Avoidant attachment	
I am not sure that I can always depend on others to be there when I need them.	Anxious attachment	
I am nervous when anyone gets too close to me emotionally.	Avoidant attachment	
I often worry that my partner does not really love me.	Anxious attachment	
Think of this ladder as representing where people stand in the Unites States	Subjective SES	

At the top of the ladder are the people who are the best off - those who have the most money, the most education, and the most respected jobs. At the bottom are the people who are the worst off - who have the least money, least education, and the least respected jobs 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.		
Please drag the slider below to indicate how warm or cold you feel toward Black people	Racial Bias	
Please drag the slider below to indicate how warm or cold you feel toward White people	Racial Bias	
In the text boxes below, please enter your estimates of what percent of all people in the United States are the following races:	Perception of Prison Racial Disparities	
In the text boxes below, please enter your estimates of what percent of all people in prison in the United States are the following races: Note: We define the US prison population as incarcerated people sentenced to more than 1 year in state or federal prison.	Perception of Prison Racial Disparities	

Table S15. Testing for interactions between effects of interest and storyteller race on empathic accuracy. Models describing empathic accuracy and compassion in the manuscript contain averages of participant RMSE and compassion across available trails. Therefore, effects for "storyteller" and "student" labels in these models cannot be teased out by race of the storyteller because race was counterbalanced within label. Therefore, we went back and pulled the trial by trial data so that we could look for interactions with storyteller race. For this analysis, we kept as much data as possible and did not exclude for pre-registered checks on the intervention (i.e., we did not remove trials where the participant got the label check wrong, we did not remove participants who failed the quiz (Table S4), and we did not remove participants who had insufficient number of trials within labels or visits). This is important—there are 40 different video stimuli randomized across all visits, so we need a large N to be powered enough to look at these effects. For this LME analysis we had 11,560 observations, N = 820 participants, and 40 unique stimuli. We modeled individual participants as random intercepts. First, we reestablish there is a significant 3-way interaction for empathic accuracy (Time by Label by Condition: t(1.07 E+04) = 2.32, p = 0.02, Estimate = 60.53, STE = 26.15) in this larger, noisier sample. Also consistent with the main findings, there is still no 3-way interaction for compassion (Time by Label by Condition: t(10784.02) = -0.75, p = 0.46). Next, we tested if storyteller race (2 levels: Black or White) interacted with Time, Label, and Condition on empathic accuracy. It did not (t(1.076E+04) = 0.413, p = 0.68), and the 3-way interaction of interest remained significant when controlling from storyteller race (t(1.076E+04) = 2.31, p = 0.02, Estimate = 59.82, STE = 25.90) There is, however, a main effect of storyteller race where participants make more empathic errors overall when the storyteller is black (t(1.076E+04) = -14.44, p < 2e-16). Likewise, there is a main effect of storyteller race on compassion (t(10770.42) = -15.30, p < 2e-16, Estimate = -4.14, STE = 0.27), where participants report greater levels of compassion for black storytellers overall. There is no Storyteller Race by Time by Label by Condition interaction on compassion.

Formula: RMSE ~ visit * cond * SLabel * stimRace + (1 | userID)
Data: race

REML criterion at convergence: 86680.5

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.1908	-0.6698	-0.0510	0.6201	4.7744

Random effects:

Groups	Name	Variance	Std.Dev.
userID	(Intercept)	21.35	4.621
Residual		95.64	9.780

Number of obs: 11560, groups: userID, 820

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.771e+01	1.879e-01	7.850e+02	147.459	<2e-16 ***
visit.L	-4.137e-01	1.303e-01	1.088e+04	-3.176	0.0015 **
cond.L	-5.492e-01	2.657e-01	7.850e+02	-2.067	0.0391 *
SLabel.L	3.022e-01	1.295e-01	1.077e+04	2.333	0.0197 *
stimRace.L	-1.869e+00	1.295e-01	1.076e+04	-14.436	<2e-16 ***
visit.L: cond.L	-2.890e-01	1.842e-01	1.088e+04	-1.569	0.1168
visit.L: SLabel.L	4.503e-02	1.832e-01	1.076e+04	0.246	0.8058
cond.L: SLabel.L	3.396e-01	1.832e-01	1.077e+04	1.854	0.0638 .
visit.L:stimRace.L	9.799e-02	1.831e-01	1.076e+04	0.535	0.5926

cond.L:stimRace.L	-1.374e-01	1.831e-01	1.076e+04	-0.750	0.4531
SLabel.L:stimRace.L	1.374e-01	1.831e-01	1.076e+04	0.750	0.4530
visit.L: cond.L: SLabel.L	5.982e-01	2.590e-01	1.076e+04	2.309	0.0209 *
visit.L: cond.L:stimRace.L	-3.024e-01	2.590e-01	1.076e+04	-1.168	0.2429
visit.L: SLabel.L:stimRace.L	-7.514e-02	2.590e-01	1.076e+04	-0.290	0.7717
cond.L: SLabel.L:stimRace.L	-8.764e-02	2.590e-01	1.076e+04	-0.338	0.7351
visit.L: cond.L: SLabel.L:stimRace.L	1.511e-01	3.663e-01	1.076e+04	0.413	0.6800

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table S16. Testing for interactions between effects of interest and participant-storyteller racial congruency on empathic accuracy. Finally, we tested for an effect of racial similarity between the storyteller and the participant. Again, we found no interaction with the DV and our 3-way interaction of interest held again controlling for this variable (*t*(1.06 E+04) = .0.90, p = 0.37); however, there was a significant main effect of racial congruence (*t*(1.15 E+04) = -7.14, p = 1.01e-12, Estimate = -1.04, STE = 1.45); where participants were more accurate when the storyteller was the same race as them. Likewise, there is a main effect of racial congruence on compassion (*t*(11367.50) = -8.25, p < 2e-16, Estimate = -2.52, STE = 0.31), where participants report greater levels of compassion for storytellers of a different race. There is no Storyteller Race by Time by Label by Condition interaction on compassion. See Supplementary Table 14 for more details.

Linear mixed model fit by REML. t-tests use Satterthwaite's method
Formula: RMSE ~ visit * movie2 * studentYes * raceMatch + (1 | userID)
Data: race

REML criterion at convergence: 86832.1

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.3005	-0.6770	-0.0414	0.6240	4.8638

Random effects:

Groups	Name	Variance	Std.Dev.
userID	(Intercept)	21.19	4.604
Residual		97.03	9.850

Number of obs: 11560, groups: userID, 820

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.752e+01	1.893e-01	8.092e+02	145.40	< 2e-16 ***
visit.L	-4.129e-01	1.347e-01	1.085e+04	-3.07	0.002 **
movie2.L	-5.912e-01	2.677e-01	8.092e+02	-2.21	0.03 *
studentYes.L	2.532e-01	1.341e-01	1.077e+04	1.89	0.06 .
raceMatch.L	-1.035e+00	1.451e-01	1.151e+04	-7.14.	1.01e-12 ***
visit.L:movie2.L	-2.594e-01	1.905e-01	1.085e+04	-1.36	0.17
visit.L:studentYes.L	7.397e-02	1.897e-01	1.077e+04	0.39	0.70
movie2.L:studentYes.L	4.135e-01	1.897e-01	1.077e+04	2.18	0.03 *
visit.L:raceMatch.L	-4.328e-02	1.901e-01	1.081e+04	-0.23	0.82
movie2.L:raceMatch.L	-2.414e-01	2.051e-01	1.151e+04	-1.18	0.24
studentYes.L:raceMatch.L	-2.884e-01	1.897e-01	1.076e+04	-1.52	0.13
visit.L:movie2.L:studentYes.L	6.595e-01	2.683e-01	1.077e+04	2.46	0.01 *
visit.L:movie2.L:raceMatch.L	2.163e-01	2.688e-01	1.081e+04	0.81	0.42
visit.L:studentYes.L:raceMatch.L	1.166e-01	2.682e-01	1.076e+04	0.44	0.66
movie2.L:studentYes.L:raceMatch.L 4	.264e-01	2.682e-01	1.076e+04	1.59	0.11
visit.L:movie2.L:studentYes.L:raceMatch.L	3.409e-01	3.792e-01	1.076e+04	0.90	0.37

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1