The Stability of Criminal Justice Policy Views: Evaluating the Effects of Factual Corrections and Appeals to Social Identity

Jane Esberg, Jonathan Mummolo and Sean J. Westwood*

September 4, 2020

Abstract

Recent protests have brought criminal justice to the forefront of U.S. politics. Moving preferences on policies like mandatory minimums, however, remains a central challenge to widespread reform. Across six experiments ($N > 11,000$), we show that changing criminal justice policy preferences remains difficult. A common explanation for widespread support for punitive policy is that most Americans believe crime is rising even during periods of decline. However, we find while the public is willing to accept factual corrections about crime rates, this never prompts reconsideration of policy opinions. Additional experiments deploying common persuasive designs show co-partisan elite cues have no effect, but individuals update their opinions when factual corrections are combined with forced consideration of opposition views or when pressured by in-group members. These interventions are cognitively burdensome, logistically challenging to scale and produce only small effects. Policy preferences are movable, but simple information treatments are ineffective, complicating criminal justice reform.

*Jane Esberg is a Postdoctoral Fellow at Princeton’s Empirical Studies of Conflict Project and the International Crisis Group, jesberg@princeton.edu. Jonathan Mummolo is an Assistant Professor of Politics and Public Affairs at Princeton University, jmummolo@princeton.edu. Sean Westwood is an Assistant Professor of Government at Dartmouth College, sean.j.westwood@dartmouth.edu. Funded by Stanford University, Princeton University, Dartmouth College and the National Science Foundation (NSF Grant #15-571); approved by Institutional Review Boards at Stanford University, Princeton University and Dartmouth College.
Decades of mass incarceration and continued high-profile police killings of unarmed men and women of color continue to prompt calls for systemic criminal justice reforms. However, as prior research on political persuasion has shown (Hopkins, Sides and Citrin, 2019; Green et al., 2011; Lawrence and Sides, 2014), convincing the mass public to support meaningful policy changes—let alone the wholesale revamping or abolition of institutions—can be extremely challenging. In the arena of criminal justice policy, the outlook is further complicated by widespread misperceptions on the closely related issue of crime: despite decades of overall declines in crime rates across the country, polls frequently show majorities of Americans reporting a belief that crime is rising (McCarthy, 2014). These rampant misperceptions have produced concern that support for punitive policy is an artifact of mistaken beliefs about public safety (Indermaur et al., 2012; Pfeiffer, Windzio and Kleinmann, 2005). More broadly, if voters so wildly misperceive important social conditions, theories of how public opinion affects the formation of polices meant to address those same social circumstances may break down, and disincentivize government officials from enacting policies to improve objective circumstances (Bartels, 2009; Downs, 1957; Ferejohn, 1986; Healy and Malhotra, 2010; Fiorina, 1981; Key, 1966; Lenz, 2013).

But such arguments rest on an important assumption: voters’ views on policy matters are informed by their perceptions of related social conditions. Empirical evidence for this assertion is slim. With rare exceptions (Cobb and Kuklinski, 1997; Gilens, 2001), prior work shows that correcting mistaken beliefs about seemingly relevant facts does little to alter related policy preferences. For example, Hopkins, Sides and Citrin (2019) finds that telling Americans the size of the immigrant population does not change attitudes toward migrants. Lawrence and Sides (2014) shows that giving correct, politically-relevant statistics—about U.S. racial composition, median income, educational attainment, and the unemployment and poverty rates—does not affect attitudes on affirmative action and aid to the poor. And Green et al. (2011) shows introducing high school students to a civics curriculum emphasizing
constitutional rights led to knowledge gains, but did not increase support for civil liberties. Across issue areas, citizens appear ready and willing to correct their mistaken beliefs, but hardly ever use that new knowledge to update their policy preferences.

The considerable scholarly focus on factual corrections as a means of inducing policy opinion change has failed to incorporate increasing evidence from social psychology showing citizens orient their politics around group identities [Iyengar et al., 2019; Jardina, 2019; Jefferson, 2020; Mason, 2015; Rogowsk and Sutherland, 2016; White, 2007; White, Laird and Allen, 2014]. Beyond just explaining the attitudes citizens hold, group-based identity is consistently one of the largest sources of persuasion in the social-psychological research (Asch, 1951; Wood, 2000; Chaiken, Giner-Sorolla and Chen, 1996). While facts about the status of political issues might be insufficient to prompt reconsideration of policy positions, people will change their opinions to reduce perceived distance from groups (either partisan or demographic), to meet normative expectations of the groups to which they belong (Chen, Shechter and Chaiken, 1996; Petty and Wegener, 1998; White and Laird, 2020a) and when prompted to consider the perspective of an opposing group (Kirmani and Campbell, 2009; Galinsky et al., 2008; Todd, Bodenhausen and Galinsky, 2012). Group identity is not the only possible source of opinion change, but as one of the most effective sources of persuasion in prior empirical work (Wood, 2000; Chaiken, Giner-Sorolla and Chen, 1996), it represents a promising strategy for prompting individuals to rethink policy attitudes. Providing facts may in itself be insufficient to move attitudes, but group identity may offer a means to both correct misperceptions and shift policy opinion.

In this paper, we investigate the prospects of garnering mass support for criminal justice reform by evaluating the ability of factual misperceptions and appeals to group identity to alter policy views in the salient and normatively important arena of criminal justice policy (Soss and Weaver, 2017). The promotion of public safety is a central mandate of government. But the massive rate of misperceptions of crime in the population, discussed
in detail below, makes this issue area an important test case for the study of how factual corrections influence policy views. The potential to change so many minds about the true status of crime in America affords a high ceiling for downstream effects on the share of individuals who may subsequently update closely related policy opinions. By the same token, if effectively correcting such widespread misbeliefs fails to induce changes in closely related policy attitudes, we may be less confident that such corrections would alter policy views more generally.

Drawing on six experiments involving more than 11,000 individuals, we first show consistent evidence that correcting misperceptions of crime does not influence views on criminal justice policies. Contrary to accounts in which individuals are resistant to updating factual beliefs (Kuklinski and Hurley 1994; Kuklinski et al. 1998; Nyhan and Reifler 2010, 2016), citizens readily update beliefs about crime trends when supplied with official statistics. For example, we find that relative to respondents who received no factual information on crime—only roughly 20% of whom correctly described the direction of recent homicide trends—over 70% of respondents who received corrective information correctly perceived the status of crime later in the survey. We see similar patterns when including distractor tasks to increase temporal distance between factual interventions and measurement of perceptions, or financial incentives for truthful responses (Bullock et al. 2015). However, we also show that acquiring these facts in isolation has virtually no effect on attitudes toward criminal justice policies, patterns that hold when presenting facts in a range of information environments and across multiple types of factual interventions, including information on homicide rates and mass incarceration.

To investigate whether group identity affects policy views—and the role of information in strengthening this effect—we deploy several experimental designs from social psychology intended to prompt sufficient cognition to carefully reevaluate opinion in the face of new evidence (e.g., Cacioppo and Petty 1984), to compel opinion change (e.g., Asch 1951) or to
think about the logic underpinning positions held by the out-group (e.g., Himmelroos and Christensen, 2014; Galinsky et al., 2008). Our results show that it is remarkably difficult, but not impossible, to change policy preferences with group identity-based interventions. Providing an elite cue from an extremely popular copartisan U.S. Senator—a relatively subtle intervention meant to remind individuals of how their partisan group views an issue—fails to move attitudes. However, two more cognitively burdensome studies show that opinion change is possible. First, we show that combining corrective facts about crime alongside social pressure to conform to the attitudes of people who are similar to the participant on multiple demographic traits (race, gender and partisanship) induces opinion change. Relative to those receiving no corrective facts or social cues, approximately 6% of participants changed their opinion on criminal justice policy. In addition, a written perspective taking task wherein individuals are asked to argue for the position of the group with whom they disagree prompts approximately 14% of participants to update their policy opinions relative to control. This effect, however, is more or less constant whether or not we couple the exercise with corrective facts, and the magnitude of these opinion shifts is modest. Importantly, these effects were almost exclusively limited to the single issue mentioned explicitly in our prompts, mandatory minimum sentencing. In other words, effects generally failed to spill over to closely related issues (Hopkins et al., 2015), suggesting prompting opinion change on a suite of related measures would require interventions tailored to each separate policy. Moreover, corrective information appears to have only a tenuous role in bolstering group-based appeals. Finally, we find no evidence that group-based appeals—on their own—cause people to reassess the status of crime, casting doubt on the possibility that relevant facts play an intermediary role in the formation of related policy preferences.

These findings make an important contributions to the study of policy opinion change. Recent studies have reported a heretofore puzzling empirical regularity—the lack of policy opinion updating following the provision of corrective information that is directly related
to the social conditions policies are nominally aimed at improving (Hopkins, Sides and Citrin, 2019; Nyhan et al., N.d.; Green et al., 2011, e.g.). In light of our results, this pattern is far less mysterious. Correcting mistaken beliefs about the status of relevant social conditions fails to change opinion on closely related policies because individuals simply rely on other information when forming policy views, including the views of salient groups. These findings also offer a bittersweet consolation to political observers concerned that widespread misperceptions may distort the policymaking process (Allcott and Gentzkow, 2017; Prior, 2007; Pariser, 2011). Far from a world in which misperceptions of objective measures of social conditions lead voters to embrace policy views they would otherwise shun, our results suggest such factual knowledge is largely irrelevant.

At the same time, these findings highlight the difficulty of persuading individuals to rethink their policy preferences, results which suggest substantial challenges for advocates of systemic criminal justice reform. While interventions focused on the opinions of salient groups caused opinion change, thereby revealing an important mechanism in this process, these exercises required significant cognitive expenditure that are difficult to achieve in the mass public (Cacioppo and Petty, 1984). In addition, those individuals that did update their policy views generally did not exhibit opinion shifts that were large in magnitude. Taken together, our results describe a world in which facts alone are insufficient to move opinion and group-based interventions are effective but impractical—patterns that illuminate the extreme difficulty of political persuasion.

The tenuous link between misperceptions and policy views

A central reason for concern over factual misperceptions is the presumed effect they have on policy preferences: if citizens cannot correctly evaluate the political world, they may develop policy views based on faulty information or fail to reward (punish) politicians for improving (failing to improve) conditions. Believing crime is on the rise when it is falling, for example, could lead to more restrictive criminal justice policy preferences than conditions
demand. Discussing the role of fake news in the 2016 election, some have even claimed that “The informational underpinnings of democracy have eroded,” (Madrigal, 2017).

But the empirical literature on misinformation, especially in recent years, suggests concerns over ignorance in the electorate are inflated, at least as they relate to policy attitudes. Studies show that while individuals will readily update mistaken beliefs when presented with facts, this factual knowledge rarely translates to shifts in attitudes toward related policies. For example, Hopkins, Sides and Citrin (2019) finds that providing Americans with information about the size of the immigrant population does not change attitudes toward migrants. Lawrence and Sides (2014) shows that giving correct, politically-relevant statistics — about U.S. racial composition, median income, educational attainment, and the unemployment and poverty rates — does not affect attitudes on related policies including affirmative action and aid to the poor. And Green et al. (2011) shows in a randomized controlled trial that introducing high school students to a civics curriculum emphasizing constitutional rights led subjects to have better knowledge of civil liberties, but did not increase support for civil liberties.

Focusing specifically on the area of crime, we see at best mixed evidence of the effect of information on policy attitudes. In work done concurrently with this study, Nyhan et al. (N.d.) shows corrective information on crime trends can improve the accuracy of perceptions but has little effect on candidate choice. Pfeiffer, Windzio and Kleinmann (2005) find a correlation between misperceptions of crime and support for harsher penalties, but the study provides no compelling evidence for a causal link. Another set of studies in criminology find evidence for only modest or fleeting effects. Indermaur et al. (2012) shows that providing information on sentencing options has little effect on attitudes towards sentencing, particularly over time. Singer and Cooper (2009) identify a small effect of providing crime statistics on opinion of the criminal justice system, but does not measure attitudes on specific policies. Two exceptions are noteworthy. Norris and Mullinix (2019) find that facts about wrongful
incarceration do reduce support for the directly related issue of capital punishment, but do not increase support for broader police reforms. And Gilens (2001) finds that providing information on the crime rate significantly affects support for prison spending. Studies showing stronger evidence for the effect of information on policy typically involve information treatments that discuss the relevant policy itself (Hopkins and Mummolo, 2017), relieving subjects of the cognitive burden of linking facts to related policy. For example, Howell and West (2009) shows that providing information on the amount being spent on schools and teacher pay alters support for spending on schools and teachers. Sides (2016) shows that informing individuals about who benefits from the estate tax changes support for that very tax.

In sum, there is at best mixed support for the notion that individuals update policy preferences based on the objective status of relevant social conditions. And when facts do move policy preferences, they typically center on the policy itself, not an outcome the policy is meant to address. When individuals are asked to make larger mental leaps—connecting, for example, information on the size of the immigrant population to border policy (Hopkins, Sides and Citrin, 2019)—they either fail or decline to do so. In the words of Lawrence and Sides (2014), these results raise the question: “Why did correct information matter so little?” (4). In what follows, we outline an alternative theoretical perspective on how policy preference change can occur.

If Not Facts, Then What? Group Identity and Persuasion

Normative models of political opinion are often premised on the rational consideration of factual information (e.g., Mill 1895). As recent work suggests, and as we confirm below, citizens are unlikely to update policy positions based on newly acquired facts. Given the low marginal utility for using new knowledge to update one’s policy beliefs, rational citizens may simply forgo the exercise for uses of time that pay greater dividends (Downs, 1957). But are policy opinions entirely immovable? In the second half of this paper we evaluate alternative
sources of opinion change derived from group identity models (Wood, 2000), and whether correcting widely-held misperceptions can bolster these effects.

We focus on three distinct group identity-based mechanisms to test for possible opinion change on criminal justice policy. These approaches vary in the amount of cognitive engagement (effort) required of citizens—from a follow-the-leader heuristic cue (Lenz, 2013) at the low end, to a task that requires participants to think deeply about the reasoning of those who hold an opposing view on a specific policy. The simplest group-based pressure comes from elite cues from co-partisan group members. Indeed, recent work shows that for many issues simply relying on party cues is often a particularly effective strategy for opinion formation and updating (Kim and LeVeck, 2013) and may have limited negative effects on democracy (Arceneaux and Wielen, 2017; however, see Achen and Bartels 2016). Some work has found that policy preferences are formed from the top-down, with individuals updating beliefs based on their broader affinity for elites (Carmines and Stimson, 1989; Layman et al., 2010; Layman and Carsey, 2002; Lenz, 2013). For more complex topics, citizens may require elite prompts to explicitly make the connection between fact and policy (Koçak, N.d.). If citizens rely on elites to process facts and come to policy positions based on these facts then treatments based on this mechanism should prompt participants to adopt the “corrected” position of a co-partisan leader with essentially no cognitive engagement.

A more complex mechanism relies on the well-established phenomenon of group conformity (Chen, Shechter and Chaiken, 1996; Petty and Wegener, 1998). Put simply, people are driven by a desire to conform to group opinion, even in situations when the group has taken a position that contradicts the direct observations of a given citizen (Asch, 1951). For a citizen to hold a contra-group position on a given policy it would have to be of such great importance to the individual to overrule the negative effects of deviating from the group (Abrams et al., 2000). If a desire to align with partisan and/or demographic groups drive policy positions, then highlighting deviation from a group should prompt citizens to realign
their preferences (Deutsch and Gerard, 1955; Johnson and Eagly, 1989). Outside of direct social-psych there is considerable evidence in political science showing that compliance with group norms affects political decision-making across a variety of issue areas, as well as voting (Gerber, Green and Larimer, 2008; Paluck, 2011; Panagopoulos, 2014; Tankard and Paluck, 2016, 2017; Viscusi, Huber and Bell, 2011; White, Forthcoming). The preconditions for this theoretical mechanism are stringent as citizens must: 1) recognize that they are deviant from group attitudes; 2) value group membership sufficiently to reconsider their positions, and finally 3) replace their prior position with the group’s normative position. This is distinct from the elite cue mechanism outlined above, where we would expect people to trust the elite without consideration of their prior attitude or even recognize that the elite’s opinion deviates from their own. It is important to note, however, that while the general social-psychology theories of group identity do not predict differential effects by race, it is possible that the racialized nature of criminal justice policy and the phenomenon of linked fate—cohesion among members of the African American community (Dawson, 1995)—may mean that group pressure might produce heterogenous results for African Americans relative to whites (Eckhouse, 2019; White and Laird, 2020).

A yet more complex mechanism requires citizens to evaluate the evidence the opposition relies upon to support their policy position (Kirmani and Campbell, 2009). Such perception-taking experiments prompt individuals to consider why those in the out-group take specific positions and actions (Kirmani and Campbell, 2009; Galinsky et al., 2008), potentially leading to greater understanding of the out-group and thereby causing opinion shifts. Incentivizing evaluation of the opposition’s thought process should, it is theorized, prompt respondents to engage in cognitive empathy whereby they deeply reflect on the logic and reasoning relied upon by the out-group (Smith, 2006). This design represents the high-point of required cognition in the experimental interventions we evaluate in this paper.

Because groups are increasingly defining how Americans think about politics (Mason
group identity-based interventions are perhaps among the most likely to prompt policy opinion change. By testing whether group-centric appeals can compel individuals to update policy views, we can shed light on the mutability of criminal justice policy positions.

Misperceptions of crime

Beginning in the early 1990s, crime in the U.S. began a sharp and precipitous decline, both in the aggregate and within major cities nationwide. Though the causes of these improvements remain disputed (Levitt, 2012), their magnitudes are astounding. In one of the most striking examples, New York City, which saw more than 2,000 murders per year in the early 1990s, now regularly experiences roughly 300—a trend that has continued virtually uninterrupted despite wide variance in the implementation of aggressive policing tactics like “Stop, Question and Frisk” (Mummolo, 2018) and significant population growth over the same period.

Yet according to the average citizen, America continues to grow more violent and lawless year after year. For over a decade, Gallup has asked the following question:

Is there more crime in the U.S. than there was a year ago, or less?

Average responses to this item appear in the top left panel of Figure 1 alongside national violent, property and homicide rates according to the FBI over the same period. As the figure shows, though crime has fallen dramatically across these various measures, the share of Americans who believe crime is rising has not aligned with these conditions. In many years, large majorities of Americans perceive that crime rose (more than 70% in 2015) even when it fell.
Figure 1: **Perceptions vs. Actual Crime.** The figure displays the share of American adults who told Gallup there was more crime in the United States than the previous year (top left panel) along with FBI measures of the national violent, property and homicide rates over the same period. Blank sections of top left time series series are years when the crime perception question was not fielded by Gallup.

In theory, there are two broad explanations for the trends in opinion displayed above. One is that Americans encounter information about crime trends, but for various reasons, fail to accept it as true. The other is that Americans simply do not encounter such information, on average. In the following section, we outline an experimental design to parse these two explanations, as well as test the ability of corrective information about national crime rates to influence criminal justice policy preferences.

**Research design: correcting misperceptions**

We first conduct a series of survey experiments to test whether corrections improve respondents’ factual knowledge and change related policy preferences. We build on previous
studies of belief accuracy when measuring perceptions of crime. Ansolabehere, Meredith and Snowberg (2013) recommend asking for quantitative values of commonly known metrics or, for more complex political issues, benchmarking against specific values in the survey question itself. We draw on these lessons, but focus on perceptions of over-time national trends rather than point-in-time levels. This focus has several advantages. For one, because point-in-time measures give no purchase on whether a given policy is improving conditions, perceived changes are believed to be central to mechanisms behind retrospective voting (Bar-tels 2002; Hopkins 2011; Fiorina 1981, 1978; Healy and Lenz, 2014). In addition, we focus on national crime trends both because these statistics are more often reported and discussed than local trends, and because evidence suggests that citizens increasingly focus on national issues even in local political contexts (Hopkins, 2018). In addition, asking individuals to judge the direction of changes in these metrics rather than the exact level at a point in time eliminates the need to establish a subjective accuracy bandwidth when coding responses. To measure perceptions of national crime trends, we included the following item in our survey experiments:

A “homicide” is the willful (non-negligent) killing of one human being by another.

The national homicide rate is the number of homicides per 100,000 people in the United States.

Was the homicide rate in the U.S. in 2015 larger or smaller than it was in 2000?

Measuring the accuracy of public perceptions requires the researcher to know the true

1 We use as our definition of a homicide the Bureau of Justice Statistics definition: “Homicide, as defined here, includes murder and nonnegligent manslaughter, which is the willful killing of one human being by another.” We provide this definition to all treatment groups, to avoid any confusion about the technical meaning of homicide. The order of the words “larger” and “smaller” was randomized across respondents, as was the order of the response options which were “larger,” “smaller” and “I don’t know.”
state of the world\textsuperscript{2} In our first three experiments, conducted between March 2017 and May 2018, we ask respondents to consider the change in the homicide rate between 2000 and 2015 because there was clear movement in the rate during this time. The national homicide rate fell from 5.5 to 4.9 (-11%) according to FBI data, a sizable decrease\textsuperscript{3} These agency estimates surely contain some error, but because we have clearly identified the metric of interest in our survey items, we can code responses in terms of whether they align with the best estimates of those same metrics according to authoritative sources—our working definition of an “accurate” perception in this study. Further, this technique is robust to the presence of measurement error by the FBI at points in time. Because we are asking about perceptions of a change, measurement error could be present every year, but would have to change enough over time to flip the sign in the difference between 2000 and 2015 to invalidate our choice of benchmarks. The close correspondence between the FBI and CDC homicide data over time displayed in Appendix ?? Figure ??, as well as the magnitude of the estimated differences in the crime rates between these two years, make this unlikely.

In addition to asking respondents about a verifiable quantity, we also include in one survey a financial incentive for accurate responses for a randomly chosen set of participants to assess the prevalence of insincere responses \cite{Bullock et al., 2015}. In addition, we include a distractor task in between the provision of corrective information and measurement of belief accuracy, in order to ensure that recall of facts endures beyond a trivial length of time (i.e., the time elapsed since answering the question immediately prior). We provide more

\textsuperscript{2} Confidence intervals are constructed using robust (“HC1”) standard errors throughout. In generating all experimental results, we code responses of “I don’t know” for the perception item as incorrect, since dropping responses in the experimental context risks post-treatment bias. Results are robust to this coding decision; see Appendix ??.

\textsuperscript{3} To remain contemporary, we update these figures to consider the same change between 2007 and 2017—a 7% drop—in the second three experiments, conducted between March 2019 and February 2020).
The main objective of our first three survey experiments was to gauge whether individuals would update beliefs about crime trends when presented with accurate information. The first study was conducted on samples drawn from Amazon Mechanical Turk, while all other experiments use Qualtrics. However, as the table shows, we also presented this information in a variety of ways to gauge the robustness of our findings to fluctuations in the information environment (Peterson 2017). These variants included the presence of copartisan elites offering competing information, and situating crime facts within longer news articles documenting an episodic crime. These variants exerted detectable effects on the degree to which individuals updated their factual beliefs about crime, which we document in Appendix Figures, and. However, because our overall conclusion is largely unaffected by these changes, we pool over all information treatments when presenting the effects of corrective information in our main analysis. In all three surveys, respondents were randomly assigned to either: (1) receive a brief report on a European soccer match (the control condition); or (2) receive information on the change in homicides between 2000 and 2015 according to the FBI data. This information treatment was worded as follows:

“According to the Federal Bureau of Investigation (FBI), the homicide rate in the U.S. was 5.5 homicides per 100,000 people in 2000, but was down to 4.9 homicides per 100,000 people in 2015.”

If individuals are resistant to such facts, we expect inaccurate beliefs to persist in the face of this information. If, however, respondents simply rarely encounter such information, supplying these facts may exert large effects. In the following section, we document the

---

4 In all surveys, where minor deception was used, such as when presenting treatments as news articles, respondents were debriefed at the end of the survey.

5 For all Qualtrics samples (studies 2-6) we utilize quota sampling on age, race and gender based on CPS benchmarks.
powerful corrective effect of this information treatment. While focusing on correcting misbeliefs regarding changes in homicide rates centers our investigation on a highly salient crime metric, and has several advantages for measurement, we also test the influence of providing other relevant facts, namely, the U.S.’s outsize share of the world’s prison population, in our elite cue-taking experiment below. This helps to ensure our results are not driven by a focus on a particular type of factual correction.

**Results: correcting misperceptions**

These first three studies confirm that, consistent with the Gallup figures reported in Figure 1, large shares of the public hold inaccurate beliefs about crime trends. We verify the prevalence of misperceptions by analyzing responses in the control conditions in our experiments. Pooling studies 2 and 3 We find that only 22% of participants responded correctly to the item asking about the direction of recent change in the national homicide rate—a strikingly low share, considering the dramatic reduction in homicides between 2000 and 2015.[6]

To ensure that this pattern is not driven by insincere responses, Study 3 includes a condition in which respondents are offered an additional financial incentive for responding correctly ($0.25), which Bullock et al. (2015) found to reduce similar biases. Comparing respondents within the control condition, financial incentives appear to lead to only a small decrease in rates of misperception compared to the non-incentive baseline, (the estimated difference is 5.13 on a scale that ranges between 0 and 100, 95% CI [0.1, 10.2]). We therefore

---

6 Notably, this figure is comparable to the share of Americans who responded either “same” or “less” to the Gallup item asking about the previous year’s crime rate in 2015, which was 26% (McCarthy, 2014).

7 To avoid prompting respondents to look up answers online in order to obtain the extra $0.25, all respondents were alerted prior to the perceived crime question that they would only have 30 seconds to respond to the item.
pool over the incentive and non-incentive conditions in remaining tests.

Table 1: Information Treatment Effects

<table>
<thead>
<tr>
<th></th>
<th>Study 1</th>
<th>Study 2</th>
<th>Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime Statistics</td>
<td>0.455***</td>
<td>0.546***</td>
<td>0.424***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.027)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.437***</td>
<td>0.181***</td>
<td>0.234***</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.017)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,571</td>
<td>1,942</td>
<td>4,242</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01; ***p<0.001

Table 1 displays the effects of the corrective information treatments across these three experiments. As the intercept estimates show, about 44% of respondents in the control condition in Study 1 (M-Turk sample) correctly indicated that homicide rates have fallen. This figure drops substantially in the more representative Qualtrics samples in Studies 2 and 3, which show that 18% and 23%, respectively, correctly signed the direction of homicide trends. While the base rate of correct perceptions varies across these studies, the effect of information treatments is more closely aligned. Across three iterations, supplying respondents with information about homicide trends improves accuracy by between 42 and 55 percentage points. Contrary to concerns that Americans may, for a variety of reasons, be resistant to factual information, respondents across all three studies readily updated their beliefs when supplied with facts. These large increases persist in the face of competing information by elites (48.2, 95% CI [42.7, 53.6]) and coverage of episodic crimes (27.5, 95% CI [23.5, 31.3], see Appendix Figures ??, ??, and ?? for additional information).

See the Appendix for treatment effects within subgroups of the data, including political parties and racial and ethnic categories.
Policy variables measured on seven-point scale. Figures show average percentage-point differences in responses between the information treatment and control condition. Bars denote 95% confidence intervals.

Given the clear increase in the share of respondents who correctly perceived trends in conditions as a result of our experimental interventions, we might expect an associated change in closely related policy stances. For example, if people come to learn that the national homicide rate has been falling, they may be less likely to support “tough on crime” initiatives (Gilens 2001). Conversely, they may also conclude that existing policies are working, and support for them may intensify. If factual beliefs inform policy stances, we should expect to see changes of some kind following corrections that exerted such large effects on beliefs.

Figure 2 displays the effect of each homicide rate treatment on support for various criminal
justice policies, as well as related attitudes (e.g. confidence that police will keep one safe and whether the respondent plans to purchase a gun). But across a host of policies and attitudes, we rarely see discernible effects. Receiving information on the falling homicide rate appears to reduce the probability of the respondent indicating they plan to buy a gun in Study 1, perhaps because they come to believe they are safer than they had previously thought. But these effect estimates are imprecise (-7.69, 95% CIs [-15.17 -0.21]), and the result failed to replicate in Studies 2 and 3. In general, we find virtually no evidence that corrective information about crime trends alters related policy attitudes.

**Testing the influence of group identity-based appeals**

As outlined above, research in social psychology suggest that prompting individuals to consider the opinions of salient groups—both opponents and allies—can lead to persuasion. It is also possible that such appeals interact with the provision of corrective information to produce even larger opinion shifts. To test this, we conducted three additional experiments.

**Study 4: Co-partisan elite cue mechanism**

We start with the simplest means to prompt a connection between facts and policy positions—exposure to an elite making the direct logical connection. We conducted a survey experiment via Qualtrics on 1,098 respondents. In this study we included five treatment cells: (1) a control (a brief description of a European soccer match); (2) an information treatment, where respondents read FBI statistics on violent crime trends; (3) an elite statement treatment, where a co-partisan Senator stated that they do not support certain criminal justice measures; (4) an elite statement and statistics treatment, which included the FBI statistics on violent crime and a co-partisan Senator stating they do not support certain criminal justice measures without referencing statistics; and (5) a reasoned elite explanation and statistics treatment, which included the FBI statistics coupled with a statement by a co-partisan Senator arguing that falling crime reduces the need for certain criminal justice measures.
Importantly, conditions with elite cues also featured co-partisan Senators providing an additional type of factual intervention apart from homicide rates, stating: “The U.S. now has about 30% more people in prison than China, despite having just 23% of China’s population. It is time to relax—or eliminate entirely—many of these tough-on-crime policies.” This feature helps to insure that our previous results are not limited to the provision of homicide rates in particular. If citizens simply rely on elite partisan cues to form policy preferences, any of the conditions in which the co-partisan Senator states their position should be sufficient to induce opinion change. However, our design also allows for the possibility that relevant facts may show enhanced persuasive power when marshalled by an elite to make an argument.

We selected two elites—Elizabeth Warren (D-MA) and Mitch McConnell (R-KY)—following a pretest on Amazon Mechanical Turk assessing the likability and perceived partisanship of Senate leadership. Following these interventions, participants completed a battery of questions measuring views on criminal justice policies. (See Appendix ?? for details on pretest and Appendix ?? for full stimuli.)

First, as with the prior three studies, participants learned the corrective fact about crime trends in the US. Relative to control, the accuracy of beliefs about crime trends improved by approximately 55 percentage points in conditions providing corrective information (p<.001 in all conditions). On the contrary, providing an elite cue without any corrective information has no effect on perceptions of crime (β = 0.02, 95% CI [-0.06, 0.07]). This is noteworthy, because it conflicts with the notion that individuals may use elite cues to update their views of the facts, which we find no evidence for in this setting ([Achen and Bartels, 2016]) (see Appendix Table ?? for full results).

As Figure 5 shows, there is no significant effect of any of the treatments on attitudes toward a wide range of criminal justice policies once we correct for multiple testing[9] including those explicitly mentioned by the trusted co-partisan U.S. Senators (“three strikes” laws and

[9] Because we did not make specific predictions about effects on policy outcomes in this experiment, we employed standard multiple testing corrections after conducting the tests
extending voting rights to convicted felons) and a host of substantively related policies like mandatory minimum sentencing. Simply offering an elite endorsement of a policy—even when supported with factual information—is insufficient to drive opinion change. When co-partisan elites make assertive arguments about criminal justice policy, or use facts to buttress those positions, we see virtually no policy preference updating.

**Study 5: Group pressure mechanism**

We now proceed to test designs that involve more cognitive engagement from respondents than merely following an elite heuristic cue. Although the task in this and the following study are unlikely to occur in the natural world, they offer an informative test of the significant amount of work required to prompt participants to update their policy positions. Given the observed difficulty of changing policy views in our previous experiments, we also narrowed the focus in studies 5 and 6 to a single issue—mandatory minimum sentencing—which has received considerable focus from criminal justice reform advocates (Free Jr., 1997). By prompting individuals to engage in more rigorous contemplation with all focus on a single issue, we are providing theoretically optimal conditions for opinion change to occur.

In study 5, we test the group pressure mechanism, which holds that awareness of deviation from group opinion prompts citizens to alter their position in order to conform to the group and to minimize social distance (e.g., [Asch, 1951]). Our sample was randomly split between four conditions: a pure control (no information), an information condition where participants learned a corrective fact; a social pressure condition where participants received information stating that a majority of participants with the demographic profile of the respondent (participants with the same race, gender and political party) held a view on summarized in Figure 5. Significant effects on prison spending attitudes for the elite cue and corrective statistics with elite cue conditions did not survive corrections: Bonferroni corrected p-values for those tests are both roughly 0.45. The less conservative Benjamini-Hochberg method yields revised p-values of roughly 0.23.
Policy variables measured on seven-point scale. Figures show average percentage-point differences in responses between the information treatment and control condition. Bars denote 95% confidence intervals. Statistically significant results on attitudes on prison spending do not survive multiple testing corrections.

mandatory minimum sentencing that was contrary to the respondent’s view; and the same social pressure cue coupled with a corrective fact. Attitudes toward a slate of criminal justice policies were captured before and after treatment. See Appendix Section ?? for full stimuli.

This design not only presents corrective information, but shows how others from a similar background responded to the information. This prompts respondents to evaluate why their

This intervention involved a minor deception; information about in-group member views was contrived. Respondents were debriefed at the end of the survey.
policy position is deviant from “typical” individuals through cognitive introspection. The purpose of this study is to test whether pressure to conform to an in-group prototype can prompt individual-level policy opinion change. We did not push respondents in a uniform direction. Rather, we pushed respondents to take the counter position to what they reported in a pre-treatment battery of opinion questions. In other words, our test was designed to measure whether opinion change occurs at all, and not to assess its direction or magnitude. In our main analysis we therefore use a dichotomous measure of change as our dependent measure.

Consistent with our prior results, respondents in both the corrective statistic and social pressure with the corrective statistic conditions were significantly more likely to correctly answer the factual question on the crime trend relative to control—33 percentage points (95% CI [0.27, 0.38]) and 46 percentage points (95% CI [0.41, 0.52]) respectively. Again, no such effect was observed in the social pressure condition where corrective information was not presented, with correct responding increasing by .5 percentage points (95% CI [-0.04, 0.05]; see Appendix Table ?? for full results.) However, Figure 4 shows that similar to Study 4, offering corrective information alone has no effect on attitude change, with 3.59% changing opinion in this arm relative to control (95% CI [-2.63, 9.81]). Social pressure on its own also has no effect on attitude change relative to control, with 1.92% changing opinion on mandatory minimums (95% CI [-4.19, 8.04]). However, supplying respondents with both corrective information and the social pressure stimuli prompted a statistically significant reevaluation of mandatory minimums, with 6.40% of respondents changing opinion relative to the pure

11 While it was not our intended focus, we report results when analyzing the absolute magnitude of policy opinion change in Appendix Table ?? . Overall we see a similar pattern across conditions, though the magnitude of opinion shifts is generally small and the significant effect we observe—social pressure with information—is no longer significant ($\beta = .02$, 95% CI [-0.01, 0.05]).

22
Policy variables measured on seven-point scale. Figure shows average percentage-point differences in absolute opinion change between the treatments and the control condition. Bars denote 95% confidence intervals.

control (95% CI [0.23, 12.57]), though we cannot distinguish the latter two effects from one another (p-value on the difference = .14). As we noted social pressure is a generic phenomenon that should work across kinds of groups (Chen, Shechter and Chaiken [1996; Petty and Wegener [1998]), but criminal justice policy is deeply racialized in the United States. Consistent with recent work (see White and Laird [2020a; Eckhouse [2019]), we find that African American respondents appear less responsive to our treatment, though we are under-powered to test for differences between African American and white respondents (see Appendix Table ??). Of the respondents in the social pressure and information condition, 31.57% moved their opinion in the assigned direction (against their pre-treatment position), 57.30% maintained their pre-treatment opinion and 11.11% moved to adopt a stronger position consistent with their pre-treatment opinion, overall this movement is marginally significant relative to pure control ($\chi^2 = 4.71, p = .09$)\textsuperscript{12}

\textsuperscript{12} Unlike Study 4, we do not employ multiple testing corrections when analyzing results for studies 5 and 6 since both were explicitly designed to move opinion on mandatory minimum sentencing, the policy discussed explicitly in the intervention.
Study 6: Out-group Perspective-Taking

This final study evaluates whether prompting individuals to consider the views of an opposing group can cause opinion change. We tasked respondents with generating evidence against their position on mandatory minimum sentences in the structure of a persuasive writing assignment. Considering the reasoning and state of those who disagree should, theory argues, prompt reevaluation of attitudes (e.g., Kirmani and Campbell, 2009). This task is the most burdensome cognitively of the set we test as it requires participants to spend considerable time (enough to draft a paragraph) pondering the oppositions’ argument.

Prior to the perspective-taking task, all respondents completed a battery assessing attitudes toward a slate of criminal justice policies. The sample was split such that participants were randomly assigned to one of four treatment conditions: pure control, a condition where a corrective fact was offered, a task designed to encourage the construction of a high-quality argument against their previously disclosed position on mandatory minimum sentences, and the same perspective-taking task coupled with corrective information. In the last two conditions, participants who initially supported mandatory minimums were asked to try to convince someone who supports mandatory minimums to oppose them, while someone who opposes mandatory minimums was tasked with convincing respondents who oppose them to support them. To incentivize participants to make high quality arguments, those assigned to the perspective-taking task were offered a chance to win a $20 cash bonus prize based on the quality of their work in the task. (See Appendix Section ?? for full stimuli.)

Compliance with the perspective-taking task assignment was high, with 81.7% of respondents providing a valid response to the prompt (to compute this metric we counted responses that made an argument about mandatory minimums, and excluded individuals who objected to the task or provided irrelevant text; see Appendix Table ?? for examples). To present the most conservative result we include non-compliers and estimate an intent-to-treat effect
The average response was several sentences and included 41.6 words. We again find that the provision of corrective facts substantially increases the accuracy of crime perceptions by 46 points when corrective facts are given on their own (95% CI [0.41, 0.52]); and 30 points when facts are coupled with the perceptive-taking task (95% CI [0.24, 0.35]). The stand-alone perceptive-taking condition did not alter crime perceptions, with movement of 3 points (95% CI [-0.01, 0.08]; see Appendix Table ??).

To evaluate effects on policy preferences, as with the prior experiment, we use a binary measure of whether opinion changed at all as the dependent measure in our models.\footnote{The largest source of non-compliance was an explicit refusal from participants to argue against their own views.}

Figure ?? shows that consistent with the prior two studies, merely offering corrective information has no effect on opinion change, with an estimated 3.59% of respondents changing their opinion relative to pure control (95% CI [-2.63, 9.82]). However, the perspective-taking task, both on its own and when combined with corrective information, changed attitudes relative to pure control. Of those assigned to the perspective-taking condition 13.80% (95% CI [7.79, 19.82]) changed their opinion on mandatory minimums, while 13.80% (95% CI [7.77, 19.84]) of those assigned to the reflection with information condition changed their opinion. Of those assigned to either the reflection condition or the reflection with information condition 36.01% moved their opinion in the assigned direction (contrary to their pre-treatment position), 49.91% maintained the same position and 14.08% adopted a more hard-line position.

Taken together, experiments 4-6 show that merely consuming facts or receiving elite messages is not sufficient to drive opinion change on criminal justice policy. However, our\footnote{Using amount of opinion change as the dependent measure produces comparable results with both perspective-taking treatments delivering significant opinion revision. See Appendix Table ??}
Policy variables measured on seven-point scale. Figure shows average percentage-point differences in absolute opinion change between the treatments and the control condition. Bars denote 95% confidence intervals.

results provide evidence that social pressure related to salient in-groups, and engaging in the construction of reasoned arguments related to policy preferences, both generate significant changes in the attitudes of respondents. While our results are too imprecise to distinguish the effects of the group-based appeals with and without the addition of corrective information, we show conclusively that corrective information on its own does not alter policy views in this issue area. Only in the presence of strong group-based appeals do we see any evidence of opinion change, and the added effect of correcting misperceptions is ambiguous at best. In addition, though we highlight effects pertaining to mandatory minimum sentencing for studies 5 and 6 since that was the focus of our interventions, additional tests show that opinion remained stable across a range of closely related policy issues, with a spillover only occurring on two of six issues in and in only one condition (from a full set of 32 additional combinations of issues and conditions; see Appendix Tables ?? and ??). Moreover, these spillovers do not survive multiple testing corrections. This lack of spillover underscores the difficulty of political persuasion, suggesting those wishing to compel citizens to update their views must craft numerous appeals tailored to specific policy proposals.
Discussion and Conclusion

As with all calls for meaningful policy change, the road to systemic criminal justice reform faces myriad obstacles, including but not limited to the challenge of garnering support in the mass public. This task is complicated in the case of criminal justice by Americans’ widespread mistaken beliefs about the closely related issue of crime. Although crime rates have fallen dramatically in recent decades, most Americans believe they are rising. This trend has potentially important political implications: by many prominent accounts, effective retrospective voting requires accurately evaluating social conditions, “law and order” platforms may remain popular due to faulty perceptions of crime, and research suggests that security threats make individuals more willing to relinquish civil liberties (Davis and Silver, 2004). It is therefore critical to evaluate whether misperceptions can be corrected, and whether these corrections have meaningful impacts on related policies.

Our results suggest that the prevalence of misperceptions of national crime is largely a byproduct of the lack of exposure to factual information about crime. Contrary to several prior studies which find it difficult to correct misperceptions (Kuklinski and Hurley, 1994; Kuklinski et al., 1998; Nyhan and Reifler, 2010, 2016), we find that citizens readily update their beliefs in response to crime statistics attributed to authoritative sources. However, informing individuals about broad improvements in crime rates does not affect theoretically proximal criminal justice policy preferences. Even when trusted copartisan elites make explicit links between changes in crime rates and policy positions, the provision of corrective information fails to trigger movement in policy preferences. Given the widespread nature of misperceptions of crime, and the ease with which these mistaken beliefs are corrected, the inability of corrective information to alter related policy views suggests citizens do not rely on perceptions of the objective status of social conditions when evaluating policies. To be clear, our experiments cannot rule out the influence of other sources of information, such as personal experiences with crime. But our results do strongly suggest that misperceptions of
objective measures of social conditions do not influence policy views.

We find consistent evidence that opinion change is possible, however, when encouraged through interventions derived from an alternative theory of policy preference formation centering on group identity. Alerting individuals that their views are out of step with other members of their race, party and gender causes roughly 31.57% of individuals to rethink their position on mandatory minimums in the assigned direction. And by compelling participants to think deeply about the arguments of the opposition, we see modest shifts in opinion among close to 36.01% of respondents. Despite widely held misperceptions about crime trends, learning accurate facts does little to bolster this effect. In addition, group-based appeals show virtually no effect on perceptions of crime, which clashes with assertions that such appeals distort people’s views of relevant facts. Interventions focused on salient group’s opinions led to measurable levels of persuasion, but perceptions of crime appear largely irrelevant to criminal justice policy preferences.

Our analysis comes with an important caveat: all of these experiments were conducted prior to the killing of George Floyd and subsequent social unrest, which prompted considerable discussion of criminal justice policy reform. It is difficult to theorize whether these events, however, hardened preexisting opinions among most Americans or made interventions that involve factual corrections more or less effective. Criminal justice is a rapidly evolving issue space, one that began to receive renewed, intense scrutiny following the 2014 killing of Michael Brown in Ferguson, MO and ensuing protests, which occurred prior to our surveys. Future work should continue to evaluate the effectiveness of the interventions we explore here as the context around these policy issues continues to evolve.

Nevertheless, the patterns we observe, which remain consistent across multiple samples, have important implications, both for scholarly understanding of how citizens evaluate policies, and the prospects of campaigns and interest groups to garner support for their issue positions. For one, our results offer a simple explanation for the previously puzzling set
of results that found citizens were both willing to update beliefs about the status of social conditions, but unwilling to rethink their stances on the very policies meant to address those conditions (e.g. Green et al., 2011; Hopkins, Sides and Citrin, 2019; Lawrence and Sides, 2014). As our experiments make clear, no such updating occurred because individuals rely on different informational cues and cognitive processes when forming their views on policy, among them deep consideration of the policy views of groups salient to them. But while we show persuasion is possible, we also demonstrate that it remains a daunting task, as the interventions which moved policy views required substantial focus and cognitive effort, and caused small amounts of opinion change for a minority of respondents on only one explicitly mentioned issue.

Political observers concerned about widespread misperceptions have long asserted they could serve to distort the policymaking process. On the contrary, our results suggest that in the area of criminal justice policy, such facts, at least when presented in isolation, are often irrelevant for the updating of policy preferences in the mass public. At the same time, the types of interventions capable of compelling citizens to rethink their policy views are complex, costly, and effective only at the margins, suggesting a difficult road for advocates of system-wide criminal justice reform.

References


Ansolabehere, Stephen, Marc Meredith and Erik Snowberg. 2013. “Asking About Numbers:


Iyengar, Shanto, Gaurav Sood and Yphtach Lelkes. 2012. “Affect, not ideology social


Layman, Geoffrey C., Thomas M. Carsey, John C. Green and Richard Herrera. 2010. “Ac-


Madrigal, Alexis C. 2017. “What Facebook Did to American Democracy And why it was so hard to see it coming.” *The Atlantic*.


Soss, Joe and Vesla Weaver. 2017. “Police are our government: Politics, political science,


