PNAS NEXUS

PNAS Nexus, 2025, 4, pgaf267
https://doi.org/10.1093/pnasnexus/pgaf267
Advance access publication 21 August 2025
Research Report

Does civilian oversight impact police legitimacy?

Kaylyn Jackson Schiff[a,*,1], Daniel S. Schiff[a,1], Bryant J. Moy[b,1], Joshua McCrain[c,1], Scott M. Mourtgos[d,1] and Ian T. Adams[d,1]

[a]Department of Political Science, Purdue University, West Lafayette, IN 47907, USA
[b]Department of Politics, New York University, New York, NY 10012, USA
[c]Department of Political Science, University of Utah, Salt Lake City, UT 84112, USA
[d]Department of Criminology & Criminal Justice, University of South Carolina, Columbia, SC 29208, USA
*To whom correspondence should be addressed: Email: schiffk@purdue.edu
[1]K.J.S., D.S.S., B.J.M., J.M., S.M.M., and I.T.A. contributed equally to this work.
Edited By Sera Linardi

Abstract

Does the implementation of external oversight in policing improve public perceptions of police legitimacy? Civilian review boards (CRBs) are frequently promoted as mechanisms to enhance the legitimacy of police agencies by providing independent oversight. Despite public support for CRBs, their adoption and effectiveness remain limited, raising concerns about their actual impact on procedural fairness and police legitimacy. This study assesses the role of CRBs in shaping public perceptions by examining various decision-making scenarios involving police chiefs and CRBs. Using a survey experiment fielded to 2,503 respondents, we investigate whether CRBs enhance legitimacy when they either coincide with or conflict with police chiefs' determinations in cases of officer misconduct. Our findings suggest that while CRBs may enhance perceptions of procedural fairness for some, particularly those with negative views of police, their involvement does not generally increase legitimacy. In fact, when CRBs conflict with police chiefs, they may diminish public trust in both policing and civilian oversight and further entrench politically polarized attitudes towards policing. These results provide empirical evidence to support concerns that CRBs might not fulfill their intended role in enhancing police legitimacy, especially in cases of institutional disagreement.

Keywords: policing, civilian oversight, survey experiment, public attitudes, legitimacy

Significance Statement

Civilian review boards (CRBs) are often proposed to restore public trust in policing, yet their impact on legitimacy remains unclear. In a preregistered survey experiment, we find that merely introducing CRBs does not generally raise perceptions of procedural fairness. In fact, when CRBs and police chiefs disagree on whether misconduct occurred, trust in both institutions erodes, offsetting any gains that might come from consensus. These results expose a core paradox for reformers and police executives: while CRBs can reassure some skeptics, their presence also risks undermining legitimacy if external oversight conflicts with internal authority. By highlighting these potential consequences, our findings sharpen debates on how best to design and implement effective police oversight.

Introduction

Civilian review boards (CRBs) bring nonpolice actors into the process of evaluating police conduct and public complaints, thereby injecting a measure of external accountability into often insular and opaque disciplinary proceedings (1). Advocates and scholars have promoted CRBs as a central reform to bolster transparency, credibility, and public trust in policing (2). Yet in the United States, CRBs remain uncommon and frequently underpowered (3, 4). This is surprising given the public's clear appetite for oversight: large majorities favor both moderate (68%) and robust (60%) CRB structures, and favor CRBs in every US state (5). The discord between steady demand and tepid institutional uptake invites a central question: if CRBs hold such intuitive appeal for citizens, why do police agencies so rarely implement them at full strength?

Proponents maintain that CRBs can restore public trust in policing by introducing a measure of independent oversight and transparency (2, 3). Yet evidence for this claim remains limited. Current evaluations, often hampered by CRBs' narrow mandates and scarce data, have not yielded clear proof that these bodies meaningfully enhance public perceptions of legitimacy—that police conform to established rules and/or laws and act in the public interest. Recent research, for example, finds no significant changes in crime clearances, officer-involved killings, or racial disparities in enforcement following the adoption of CRBs in major US cities (6), nor are CRBs associated with changes in arrests, crime, or use of force (7). Despite sustained public enthusiasm for the idea, police executives have long resisted ceding disciplinary authority to outsiders (5, 8), even as they seek to shore up the very legitimacy that CRBs promise to deliver (9, 10).

Competing Interest: The authors declare no competing interests.
Received: January 30, 2025. **Accepted:** July 31, 2025

These patterns suggest a more complex strategic calculus than simple concern over autonomy. Police chiefs, as leaders of powerful and politically salient institutions (9), may fear that CRBs introduce unpredictability and conflict. Chiefs are institutionally inclined to clear most misconduct allegations (11), yet the presence of a CRB could highlight disagreements that cast doubt on the overall fairness of the process (12). In particular, a lenient CRB might appear either impotent or co-opted, eroding public confidence in the idea of genuine civilian oversight. Similarly, if CRBs appear largely symbolic—serving to satisfy external demands but remaining peripheral to actual discipline—this decoupling from substantive outcomes further damages trust in both the oversight body and the police (10, 13). Moreover, it is also possible that public opinion on legitimacy is obdurate and driven primarily by political polarization (14). Our investigation of civilian oversight additionally speaks to broader debates about institutional legitimacy. Studies of the Supreme Court demonstrate competing views on whether legitimacy depends on agreement with specific rulings (15–17) or persists through a reservoir of goodwill despite disagreement with outcomes (18, 19).

We empirically examine these possibilities by assessing whether CRBs enhance public perceptions of police legitimacy under different decision-making conditions. We define legitimacy in terms of the public's sense that both police and civilian oversight institutions are acting fairly, transparently, and in the public interest. This approach builds on research showing that public views of criminal justice actors reciprocally shape and are shaped by institutional behavior (20–24). Drawing on a preregistered survey experiment with a nationally representative US sample, we compare scenarios where the police chief acts alone to those where a CRB also weighs in. We vary whether the CRB and chief agree or conflict in their determinations of misconduct. By systematically manipulating these conditions, we provide new insights into whether CRBs genuinely bolster public faith in policing, or risk doing the opposite. Specifically, we find that while CRBs may reassure some skeptical citizens, agreement between CRBs and police leaders does not reliably increase perceived police legitimacy, whereas disagreement significantly reduces it. Further, we find that support for CRBs is polarized; individuals tend to side with whichever actor's determination matches their preferred outcomes, regardless of the institutional arrangement.

Civilian oversight and perceptions of police investigations

Our study tests several hypotheses regarding how citizens view investigations into officer misconduct that incorporate a role for civilian oversight. First, building on the idea that "the very existence of citizen oversight may well enhance public trust in the police" ((25), p. 102), we expect that citizens will judge investigations to be more fair and will have greater trust in the process (i.e. greater perceptions of procedural legitimacy) when CRBs are included in investigations of police misconduct, compared to when they are absent. Our "CRBs Boost Procedural Legitimacy Hypothesis" states that the presence of CRBs, as independent investigatory authorities, will boost perceptions of procedural legitimacy for police investigations.

Second, we expect that the public will hold heightened trust in the police when potentially conflicting actors agree on the final determination of an investigation, as this provides additional signals of credibility (26). In particular, we expect police chiefs to benefit from the support of the ostensibly independent CRBs. Public perceptions of police legitimacy should be highest when the two institutions match regarding their decision, and lowest when they do not match and chiefs are perceived to be too lenient or biased. Our "Chiefs Benefit from Consensus Hypothesis" states that agreement between police chiefs and CRBs regarding investigations of police misconduct increases, and conflict reduces, perceptions of police legitimacy.

Next, we turn to perceptions of CRB legitimacy rather than police legitimacy. We expect that perceived CRB legitimacy is dependent upon the direction of mismatch between the determinations of chiefs and CRBs ("Independent Authority of CRBs Hypothesis"). As a result of the perceived conflict of interest of police chiefs in performing internal investigations, and the sense that CRBs should function as independent "checks" on policing agencies that may otherwise be opaque or biased (26), we expect that the public will be more likely to view CRB decisions as legitimate when they go against the presumably internally favorable determinations of the police chief (i.e. lenient decisions). When the CRB is relatively more harsh (finding misconduct when chiefs do not), perceived CRB legitimacy will increase. However, we expect the opposite effect for CRBs that might be perceived as the relatively more lenient actor and thus as a "captured" institution (27, 28). When the CRB is relatively more lenient (finding no misconduct when chiefs find misconduct), perceived CRB legitimacy will decrease.

Finally, we expect that members of the public will have stronger preferences over outcomes than processes (29–31). That is to say, the public will be more likely to view police officers under investigation as guilty of misconduct, due to public distrust of police and suspects of investigations, though we do expect partisan heterogeneity. In the case of police chiefs in particular, we anticipate the public will view police chiefs willing to make a determination against a member of their agency (i.e. "breaking the blue wall of silence") as making a more costly and controversial decision (11), one which deserves additional respect. Our "Presumption of Guilt Hypothesis" states that police chiefs and CRBs who determine that misconduct did occur will receive a greater payoff in terms of public perceptions than police chiefs and CRBs who determine that misconduct did not occur.

Experimental design

To test our hypotheses, we fielded a survey experiment in late August 2024 on 2,503 American adults recruited through Prolific.[a] The sample is representative of the US population in terms of sex, age, and ethnicity.[b]

All participants received a vignette describing an investigation into the actions of a police officer. We randomly assigned participants to one of six different versions of the vignette. Randomization ensures that each respondent has an equal chance of receiving any of the treatments, which handles self-selection into information about CRBs. The sample was balanced across experimental conditions as reported in Table S2. Two vignettes served as the control conditions and explained only that there was a department-led investigation into the officer's behavior and that the police chief had announced the results (either that the officer's actions were appropriate or inappropriate). For the remaining four vignettes, participants were additionally told that the local civilian review board had conducted an independent investigation and had reached a determination. Two treatment vignettes presented matching determinations announced by the police chief and CRB (agreement either that the officer's actions were appropriate or inappropriate). The final two treatment vignettes instead presented mis-matching determinations

between the police chief and CRB (the two conditions examine both possible directions of disagreement). We additionally randomly assigned participants to read about one of three different forms of potential misconduct, ranging in severity from rude language to use of a taser to use of a firearm. However, for the main tests of our hypotheses, we aggregate responses across types of misconduct.

After reading the assigned vignette, we asked questions to assess participants' views of the legitimacy of the process and actors involved. To assess procedural legitimacy, we asked participants to rate their agreement with the statement: "I trust that the allegations in this case were investigated fairly."[c] Then, to understand perceptions of police legitimacy and CRB legitimacy, we asked about agreement with two separate statements with the same structure: "I trust the [Chief or CRB] to make fair decisions in investigations like this in the future."[d] Finally, to enable an examination of "outcome legitimacy," we asked participants to select between two options indicating either that they believed the officer's actions to be appropriate or inappropriate.

To analyze results, we use ordinary least squares (OLS) regression to identify treatment effects comparing the relevant groups and using the associated outcome measure for each hypothesis. For additional information about our experiment, including the wording of the vignettes shown to participants and details about our analysis strategy, please see the Materials and methods section.

Results

CRBs largely do not boost procedural legitimacy

As shown in Table 1, we find no statistically significant evidence that the presence of a CRB increases aggregate perceptions of procedural legitimacy for investigations into alleged police misconduct. We additionally find no significant interactive effects with the type of misconduct, respondent partisanship, or respondent race.

However, as shown in Fig. S1, we find that there are significant differences between individuals with warmer and colder feelings toward the police ($P = 0.002$). CRB presence increases perceptions of procedural legitimacy for individuals that strongly dislike the police (lower ratings on the feeling thermometer[e]) and has either no impact or a small negative impact on perceptions for those with warm feelings towards the police.

Additionally, as displayed in Fig. S2, we find some evidence that perceptions of procedural legitimacy are higher for individuals who report that a CRB exists in their own community,[f] but the effect is not statistically significant ($P = 0.183$). This provides some support for prior work showing increased support and willingness to allocate additional powers to CRBs amongst police chiefs and sheriffs in areas with existing CRBs (5). While empirically demonstrating this is beyond the scope of this article, the establishment of oversight institutions may generate positive feed-forward effects whereby both members of the public and government officials with real-world exposure to a local CRB come to develop more positive feelings regarding their legitimacy.

Overall, then, while we do not find average boosts to procedural legitimacy resulting from presence of a CRB, it would be premature to conclude that this null result holds for all subgroups and contexts. Different manifestations of civilian oversight, increased awareness or experience of the public with these institutions, and different patterns among subgroups could boost legitimacy. Yet, our primary results suggest that the mere presence of CRBs is not a panacea for boosting legitimacy amongst the general public.

Table 1. Impact of CRB presence on perceived procedural legitimacy

	Likert outcome		Binary outcome	
CRB present	0.027	0.024	−0.002	−0.003
	(0.045)	(0.041)	(0.021)	(0.020)
Num.Obs.	2,503	2,499	2,503	2,499
R^2	0.0001	0.157	0.000004	0.103
Covariates	No	Yes	No	Yes

[a]$P < 0.1$, [b]$P < 0.05$, [c]$P < 0.01$, [d]$P < 0.001$.
The results are based on comparing the two control groups pooled (vignette did not include a CRB) to the four treatment groups pooled (vignette included information about a CRB). The variable "CRB present" indicates receipt of one of the four treatment vignettes in which a CRB was present and included in the investigation into alleged misconduct.

Conflict hurts police legitimacy more than consensus helps

Figure 1 shows the impact of CRB determinations on perceived police legitimacy compared to two alternative decision-making situations: when the chief alone determines that misconduct did not occur, or when the chief alone finds misconduct. In both cases, having a CRB match the chief's determination offers no significant boost to perceived police legitimacy ($P = 0.069$ and $P = 0.754$, respectively). However, having a CRB go against the chief's determination significantly undermines perceptions of police legitimacy (−0.35 ($P < 0.001$) and −0.41 ($P < 0.001$) points on the Likert scale for chief determinations that misconduct did not or did occur, respectively). Moreover, we find no significant interaction effects with the type of misconduct, respondent partisanship, respondent ideology, respondent race, respondent feelings toward the police, or respondent awareness of a CRB in their own community. That is, we find that the cuts to perceived police legitimacy are relatively universal and that no subgroups offer boosts.

We also assess whether conflict hurts police legitimacy more than consensus with a CRB helps. We find a significant and large police legitimacy loss of 0.35 points on average on the Likert scale (0.31 SD) when a CRB says that misconduct occurred after a chief's determination of no misconduct. In contrast, we find a smaller and insignificant police legitimacy gain of 0.12 points on average on the Likert scale (0.11 SD) when the CRB agrees with the chief's determination of no misconduct. We believe that this evidence provides part of the explanation for why police chiefs are resistant to implementing civilian oversight institutions such as CRBs.

CRB legitimacy is hurt by conflict and is politically polarized

We find conflicting evidence for our hypothesis regarding the independent authority of CRBs. As shown in Fig. S3, we find no average boost to CRB legitimacy when CRBs alone say that misconduct occurred compared to situations in which both the chief and CRB determine that misconduct did not occur. As a reminder, the logic here is that we are holding fixed the determination of the chief—in this case that misconduct did not occur—and then comparing CRB disagreement to agreement with that determination. Instead, we find a significant negative effect (−0.15 points on the Likert scale, 0.15 SD, $P = 0.028$), suggesting reduced aggregate perceptions of CRB legitimacy for this form of disagreement. In addition, we find a larger significant reduction in perceived CRB legitimacy

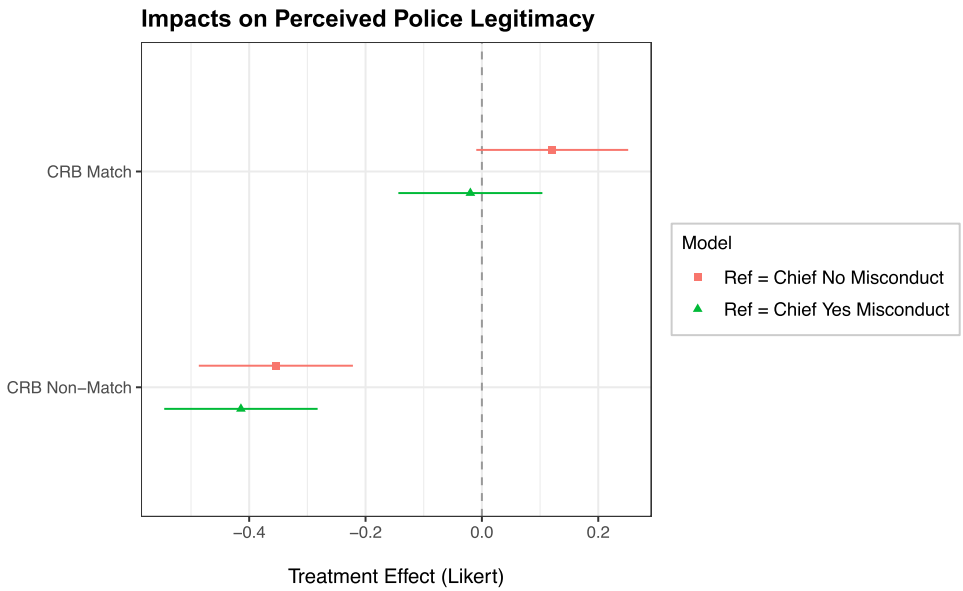

Fig. 1. Impacts on perceived police legitimacy. The results displayed in the figure are based on comparing pairs of treatment groups. First estimate: Chief and CRB both find no misconduct (match) vs. Chief alone determines no misconduct (control). Second estimate: Chief and CRB both find misconduct (match) vs. Chief alone determines misconduct (control). Third estimate: CRB finds misconduct and Chief determines no misconduct (nonmatch) vs. Chief alone determines no misconduct (control). Fourth estimate: CRB finds no misconduct and Chief determines misconduct (nonmatch) vs. Chief alone determines misconduct (control).

when CRBs alone say that misconduct did not occur compared to consensus between the chief and CRB that misconduct occurred (−0.38 points on the Likert scale, 0.39 SD, $P < 0.001$). Thus, like police legitimacy, we find that aggregate perceptions of CRB legitimacy are also negatively impacted by conflict.

However, we find evidence of heterogeneous effects by partisanship, ideology, and feelings toward the police. Warmer feelings toward the police, Republican identification, and conservative ideology are all associated with significant declines in perceptions of CRB legitimacy when the CRB disagrees and finds misconduct, compared to agreement that misconduct did not occur. The top panel of Fig. 2 shows an example of this trend based on ideology. For liberals (scores of 1 and 2 on the 7-point ideology scale), there is a significant boost to perceived CRB legitimacy when the CRB disagrees and determines that misconduct occurred, but for moderates and conservatives (scores of 4 to 7 on the 7-point scale), there is instead a significant decrease ($P < 0.001$). Thus, there is some evidence for our hypothesis regarding increased legitimacy for relatively harsh CRBs, but only amongst individuals who identify as Democrats or liberals, or who have negative feelings toward the police.

However, the bottom panel of Fig. 2 shows the opposite relationship for CRBs that disagree and determine that misconduct did not occur. Liberals and moderates exhibit reduced perceptions of CRB legitimacy, but strong conservatives view contrarian CRBs more favorably when they find that misconduct did not occur ($P < 0.001$), likely driven by greater favorability toward the police ($r = 0.47$). On balance, then, we find overall reduced perceptions of CRB legitimacy when there is conflict between the determinations of CRBs and police chiefs, but there is also heterogeneity such that strong Democrats/liberals are more favorable toward CRBs that independently find misconduct and strong Republicans/conservatives are more favorable toward CRBs that independently validate police actions. This perhaps suggests preferences driven more by outcomes than institutions.

Misconduct determinations impact perceived legitimacy

As shown in Table 2, we find some evidence that determinations that misconduct occurred boost perceived legitimacy. Chief determinations that misconduct occurred significantly increase perceptions of police legitimacy by an average of 0.42 points on the Likert scale (0.37 SD, $P < 0.001$). As shown in columns 2 and 3 of Table 2, this is mostly driven by very large increases amongst Democrats, while there are no significant or substantive changes in Republican perceptions of police legitimacy. For CRBs, determinations that misconduct occurred also slightly increase average perceptions of CRB legitimacy (0.11 points on the Likert scale, 0.12 SD, $P = 0.016$). However, this small average effect masks significant relatively equal and opposing effects amongst Democrats and Republicans, shown in columns 5 and 6. For Democrats, CRB determinations that misconduct occurred significantly improve perceptions of CRB legitimacy (0.45 points on the Likert scale, 0.46 SD, $P < 0.001$), but for Republicans, such CRB determinations reduce perceptions of CRB legitimacy to a similar degree (−0.42 points on the Likert scale, −0.44 SD, $P < 0.001$).

Preferences for outcomes trump preferences for institutions

As for perceptions regarding the legitimacy of the outcomes, we find very high baseline agreement with the chief (92% for Democrats and 82% for Republicans) in cases where the chief alone makes a determination that misconduct occurred. Agreement is lower for Democrats (61%) and higher for Republicans (86%) when the chief alone determines that misconduct did not occur. When the chief and CRB concur, agreement with their determination is very high (95% for Democrats and 84% for Republicans) when they find that misconduct occurred. When they concur that misconduct did not occur, agreement is still quite high amongst both Democrats (72%) and Republicans (88%). Table S1 presents these findings.

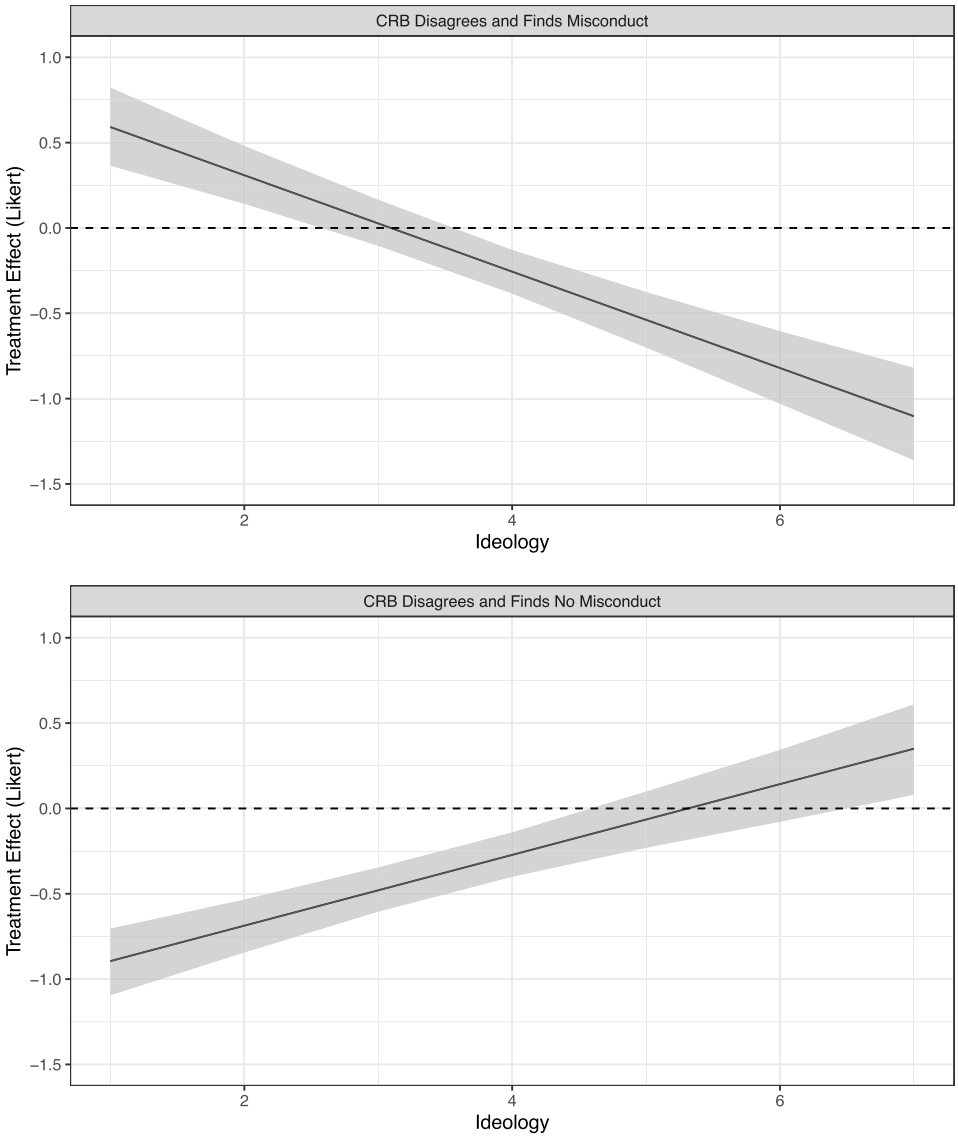

Fig. 2. Ideology and perceptions of CRB legitimacy when CRBs disagree with chiefs. The results displayed in the figure are based on comparing pairs of treatment groups. Top panel: CRB finds misconduct and Chief determines no misconduct (nonmatch) vs. CRB and Chief both find no misconduct (match). Bottom panel: CRB finds no misconduct and Chief determines misconduct (nonmatch) vs. CRB and Chief both find misconduct (match).

When the chief and CRB disagree, Democrats side with the actor who found that misconduct occurred, while Republicans side with the actor who found that misconduct did not occur. That is, when the CRB finds misconduct but the chief does not, 69% of Democrats agree with the CRB, but 61% of Republicans agree with the chief. When the chief finds misconduct but the CRB does not, 70% of Democrats agree with the chief, but 66% of Republicans agree with the CRB. This again suggests preferences for outcomes over institutions.

Discussion and conclusion

This study reports the results of a preregistered survey experiment based on data from a nationally representative sample. We assess whether the inclusion of civilian oversight buttresses public perceptions of procedural legitimacy, police legitimacy, and CRB legitimacy. Our findings are sensitive to standard critiques of survey experiments, such as simplification inherent in informational vignettes and cross-sectional research, potential self-reporting bias, and stronger internal as compared to external validity. Nevertheless, the findings allow us to weigh in on a question of significant public and policy importance.

Overall, our findings present some cause for concern for those interested in promoting civilian oversight in the form of CRBs to boost the perceived legitimacy of police institutions. Generally, CRBs may not elevate aggregate attitudes toward procedural legitimacy. Indeed, only respondents with the coldest baseline attitudes toward the police or the few that are aware of a CRB in their own community are reassured by the presence of a CRB, whereas a much larger share of individuals with more neutral or positive baseline views of police or with no knowledge or awareness of CRBs experience little to no boosts to their perceptions in procedural legitimacy.

These patterns reflect the strategic landscape that police executives confront when weighing the adoption or endorsement of CRBs. We initially anticipated that chiefs would reap legitimacy gains when their decisions aligned with those of a CRB, but that conflict between these actors would erode trust. Yet our findings

Table 2. Impact of misconduct determination on perceived legitimacy

	Police legitimacy			CRB legitimacy		
	All	Dem	Repub	All	Dem	Repub
Chief finds misconduct	0.423[d]	0.730[d]	−0.075			
	(0.039)	(0.053)	(0.067)			
CRB finds misconduct				0.114[b]	0.452[d]	−0.424[d]
				(0.047)	(0.061)	(0.083)
Num.Obs.	2,499	1,388	824	1,659	915	555
R^2	0.261	0.297	0.150	0.055	0.123	0.080

[a]P < 0.1, [b]P < 0.05, [c]P < 0.01, [d]P < 0.001.
The variable "Chief Finds Misconduct" indicates receipt of one of the three vignettes in which the Chief finds misconduct, compared to the pooled group that received one of the three vignettes in which the Chief finds no misconduct. The variable "CRB finds misconduct" indicates receipt of one of the two vignettes in which the CRB finds misconduct, compared to the pooled group that received one of the two vignettes in which the CRB finds no misconduct. The sample size for the "CRB finds misconduct" results are reduced because only about four out of six of respondents received vignettes including a CRB.

Table 3. Design of experiment

Condition	Type of event	Chief decision	CRB decision
1 control-no misconduct	Misconduct: language, taser, shooting	Misconduct did not occur	N/A
2 control-misconduct	Misconduct: language, taser, shooting	Misconduct did occur	N/A
3 match-no misconduct	Misconduct: language, taser, shooting	Misconduct did not occur	Misconduct did not occur (Match)
4 nonmatch-no misconduct	Misconduct: language, taser, shooting	Misconduct did not occur	Misconduct did occur (Nonmatch)
5 match-misconduct	Misconduct: language, taser, shooting	Misconduct did occur	Misconduct did occur (Match)
6 nonmatch-misconduct	Misconduct: language, taser, shooting	Misconduct did occur	Misconduct did not occur (Nonmatch)

show no upside from consensus and a significant legitimacy penalty from disagreement. This asymmetry is particularly problematic when CRBs question lenient judgments, since chiefs face built-in incentives to clear officers of alleged misconduct (11). In other words, the introduction of a CRB does not promise clear benefits in terms of enhancing police legitimacy, and may bring real costs when its conclusions do not harmonize with the chief's own.

External civilian review emerges as a polarizing rather than unifying force. Like other political arenas, the public often judges institutional arrangements more by the results they deliver than by the fairness of their underlying procedures (30, 33). In our data, Democrats and those with initially low regard for the police view CRBs as more legitimate when they render tougher decisions, whereas Republicans and individuals who hold more positive views of law enforcement respond negatively to the same moves. The opposite pattern emerges when CRBs behave leniently: what enhances legitimacy among conservatives reduces it among those on the left. Rather than generating broad-based gains in perceived legitimacy, CRBs thus reinforce existing divides, reflecting broader national trends in political polarization driving public opinion towards political institutions. For CRBs, this suggests that their legitimacy rests heavily on outcome alignment rather than procedural idealism, and reflects deeper assumptions about guilt, innocence, and the fundamental nature of police misconduct.

Our findings remain relevant even if the typical respondent lacks detailed knowledge about civilian review boards. In reality, most citizens know relatively little about many criminal justice institutions, yet public judgments of policy interventions often emerge from limited, context-specific information (34, 35). While our study holds the form and function of CRBs constant—assuming they possess the capacity for independent investigations—the interpretation of "independence" likely varies significantly across contexts and CRB compositions. Moreover, our study holds fixed the (in)existence of a CRB in the respondent's mind. Future research should collect observational data on CRB existence and examine factors associated with both CRB adoption and public opinion toward CRBs.

Relatedly, future work should investigate how variations in CRB presence, authority, and investigatory power influence public perceptions of legitimacy. It is reasonable to hypothesize that CRB authority levels and perceived legitimacy are interconnected, with different configurations of CRB power producing specific effects. By presenting concrete scenarios illustrating CRB decision-making, we effectively approximate the conditions under which ordinary individuals would confront these institutions in practice. The resulting patterns—such as the absence of universal legitimacy boosts and the polarized responses across ideological groups—underscore that even modest cues suffice to influence opinions. Rather than requiring intricate understanding of CRB authority, the public reacts to the prospect of independent oversight in ways consistent with their baseline attitudes toward policing and their inclination to trust or doubt governmental checks. This suggests that our conclusions are robust to the reality of low-information publics and reflect how these reforms would likely play out in politically diverse environments where most citizens never become deep experts in policy detail.

These results highlight the risks of distributing disciplinary authority across multiple actors, especially from the perspective of police executives. Conflict between CRBs and chiefs does more than fail to enhance legitimacy; it actively undermines it, with the greatest damage occurring when one party appears unduly lenient. In contrast, consensus yields at best modest and often polarized gains. To the extent that a primary objective of expanding civilian oversight is to strengthen police legitimacy, our findings indicate that CRBs are a net liability rather than an asset. This asymmetry helps explain why police leaders, wary of these downside risks, often resist fully embracing civilian oversight (5, 12). Yet, we emphasize that CRBs may elevate perceived legitimacy for citizens with lower baseline trust in the police—often precisely the intended audience of reform efforts—and may serve functions and move key outcomes beyond public opinion.

Materials and methods

The survey experiment for this study was reviewed and deemed exempt by the IRBs for Purdue University (Protocol IRB-2024-700), New York University (Protocol IRB-FY2024-9178), the University of Utah (Protocol IRB-00179081), and the University of South Carolina (Protocol Pro00137443). Informed consent was obtained from all participants. The authors additionally preregistered a preanalysis plan for their study with the open science framework (OSF) before fielding the experiment.

Before fielding the study, we conducted simulations to estimate the power of the experimental design for a component of hypothesis H1 across a range of sample sizes. In particular, we conducted simulations assuming a standardized treatment effect of 0.125 for the pooled matching conditions compared to the pooled control groups with no CRB. In order to be powered to detect positive significant effects (such as boosts to perceived legitimacy) when comparing the matching conditions to the control conditions, we estimated that we would need a sample size of 3,000. This informed our anticipated sample size of 3,000. As a result of increased payments needed for participants due to survey durations longer than expected, we were only able to obtain 2,500 participants with the given budget.

Table 3 provides additional information about the six treatment groups in our experimental design. Participants were randomly assigned with equal probability to one of six possible vignettes. The conditions are labeled as well as numbered, with the labels distinguishing between control, matching, and nonmatching vignettes, as well as the direction of the chief's decision (either "no" or "yes" to whether misconduct occurred).

The wording of the vignette shown to participants is as follows:

> "An investigation was recently undertaken into conduct by Officer Jones, a 12-year veteran of the police department.
>
> In a recent incident, Officer Jones [used profanity directed toward a subject — used his taser against a subject — discharged his firearm against a subject] while on duty. At the time of the event, Officer Jones was attempting to serve an arrest warrant.
>
> A department-led investigation sought to evaluate whether Jones's use of [profanity — his taser — his firearm] complied with departmental policy. Department policy allows [profanity — taser use — firearm use] under some conditions.
>
> Following the investigation, Police Chief Smith announced the results, which found Officer Jones's actions to be [appropriate/inappropriate].
>
> [If applicable:] In addition to the Chief's findings, the local Civilian Review Board (CRB) conducted their own investigation and [also/instead] found Officer Jones's actions to be [appropriate/inappropriate]."

The full survey instrument (complete wording and order of all questions asked in our survey) is available in the Supporting Information.

To provide further detail on our analysis strategy, we explain in depth how we assessed our first hypothesis, H1. We use a similar approach to analyze results for our other hypotheses, but the detailed regression specifications can be found in our preanalysis plan with OSF for those who are interested.

To address H1, the CRBs Boost Procedural Legitimacy Hypothesis, we pool the four treatment conditions in which a CRB is present and makes a determination and compare them to the pooled conditions without a CRB (two control conditions). We use the following regression specification:

$$\text{Procedural legitimacy} = \beta_0 + \beta_1 \text{CRB} + \gamma \mathbf{X} + \epsilon,$$

where Procedural legitimacy references the procedural legitimacy outcome, CRB is an indicator for receiving a vignette with a CRB determination, $\mathbf{X}$ refers to the vector of the covariates, and ϵ refers to the error. The reference group is composed of those who received vignette treatments without a CRB determination. As per H1, we expect β_1 to be positive (e.g. greater procedural legitimacy with CRBs as an independent investigatory authority).

In our analysis, we control for the following demographic covariates: gender, race, age, education, household income, partisanship, ideology, feelings toward the police (police thermometer), perceptions of crime, and prior experience (for self and friends/family) with police use of force.

Notes

[a] Thirty-three respondents completed our survey in September or early October due to difficulty filling quotas for the representative sample. All respondents completed the survey by 2024 October 2.

[b] We included attention checks in our survey, but we did not boot respondents who failed the attention checks, or exclude them from the main analysis, as is best practice (32). 84% of respondents in our sample answered two instructional manipulation check questions correctly, and the Supporting Information presents results—which are consistent with the main results—for this attentive subset as Table S3, Fig. S4, Fig. S5, and Table S4.

[c] We measured all outcomes—except the outcome legitimacy measure—using a five-point Likert scale from Strongly disagree to Strongly agree.

[d] Note that the statement about CRBs was only shown to participants who received a vignette mentioning a CRB.

[e] Note that the first quartile police feeling score is 35 on the 100 point scale, and the median is 65.

[f] Specifically, 12% of respondents (290/2,503) report awareness of a CRB in their own community; though we cannot verify from our data if these reports are accurate. While this boost to legitimacy associated with self-reported awareness of a local CRB could be driven by other factors, note that awareness of a CRB in one's own community is not strongly correlated with either Democratic identification ($r = -0.04$) or feelings toward the police ($r = 0.08$).

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

K.J.S., D.S.S., J.M., and I.T.A. received funding from the National Institute of Justice (award number 15PNIJ-23-GG-05487-RESS) to complete this research.

Author Contributions

K.J.S., D.S.S., B.J.M., J.M., S.M.M., and I.T.A. all contributed to the conceptualization, experimental design, analysis, writing, and editing.

Data Availability

The data and replication materials for this study, along with the preregistration, are available via the Open Science Framework (OSF) at https://osf.io/93uep/.

References

1 Walker S. *Police accountability: the role of citizen oversight*. 1st ed. Wadsworth Publishing, Australia, Belmont, CA, 2000.
2 Ferdik FV, Rojek J, Alpert GP. 2013. Citizen oversight in the United States and Canada: an overview. *Police Pract Res*. 14(2):104–116.
3 De Angelis J, Rosenthal R, Buchner B. 2016. Civilian oversight of law enforcement: assessing the evidence. U.S. Department of Justice, Office of Justice Programs, NCJ 250287.
4 Rocha Beardall T. 2022. Police legitimacy regimes and the suppression of citizen oversight in response to police violence. *Criminology*. 60(4):740–765.
5 Adams IT, McCrain J, Schiff DS, Schiff KJ, Mourtgos SM. 2024. Police reform from the top down: experimental evidence on police executive support for civilian oversight. *J Policy Anal Manag*. 44(2):403–427.
6 Krishnamurthy A. 2024. Handcuffed: the limited effects of civilian oversight on police behavior.
7 Kim J. 2022. The effect of civilian oversight on police organizational performance: a quasi-experimental study. *Am Rev Public Adm*. 52(5):382–397.
8 Chanin J, Espinosa S. 2016. Examining the determinants of police department transparency: the view of police executives. *Crim Justice Policy Rev*. 27(5):498–519.
9 Crank J, Langworthy R. 1992. Institutional perspective on policing. *J Crim Law Criminol*. 83(2):338.
10 Rubin AT, Vaughn PE, Rudes DS. 2024. Neo-institutional analyses of criminal legal organizations and policies. *Annu Rev Law Soc Sci*. 20(1):331–352.
11 Clark TS. 2024. A political-economic model of police administration. *J Polit*. 86(4):1462–1478.
12 Mourtgos SM, Adams IT, McCrain J, Schiff KJ, Schiff DS. 2024. Speaking of reform: experimental insights into influencing police executives' perspectives on civilian oversight. *Criminol Public Policy*. 23(3):689–720.
13 Meyer JW, Rowan B. 1977. Institutionalized organizations: formal structure as myth and ceremony. *Am J Sociol*. 83(2):340–363.
14 Schiff KJ, *et al*. 2024. Police shooting statistics and public support for police reforms. *J Exp Polit Sci*. 11(3):288–299.
15 Bartels BL, Johnston CD. 2013. On the ideological foundations of supreme court legitimacy in the American public. *Am J Pol Sci*. 57(1):184–199.
16 Christenson DP, Glick DM. 2015. Chief justice Roberts's health care decision disrobed: the microfoundations of the supreme court's legitimacy. *Am J Pol Sci*. 59(2):403–418.
17 Nicholson SP, Hansford TG. 2014. Partisans in robes: party cues and public acceptance of supreme court decisions. *Am J Pol Sci*. 58(3):620–636.
18 Gibson JL, Nelson MJ. 2015. Is the US supreme court's legitimacy grounded in performance satisfaction and ideology? *Am J Pol Sci*. 59(1):162–174.
19 Caldeira GA, Gibson JL. 1992. The etiology of public support for the supreme court. *Am J Pol Sci*. 36(3):635–664.
20 Peyton K, Sierra-Arévalo M, Rand DG. 2019. A field experiment on community policing and police legitimacy. *Proc Natl Acad Sci U S A*. 116(40):19894–19898.
21 Porumbescu GA, Neshkova MI, Huntoon M. 2019. The effects of police performance on agency trustworthiness and citizen participation. *Public Manag Rev*. 21(2):212–237.
22 Sunshine J, Tyler TR. 2003. The role of procedural justice and legitimacy in shaping public support for policing. *Law Soc Rev*. 37(3): 513–547.
23 Wood J, Gannon TA. *Public opinion criminal justice: context, practice and values*. Taylor and Francis, Hoboken, 2013.
24 Keyes VD, Keyes L. 2023. A scoping review of police civilian oversight literature 1992–2022. *Policing Int J*. 46(4):601–622.
25 Worden RE, Bonner HS, McLean SJ. 2018. Procedural justice and citizen review of complaints against the police: structure, outcomes, and complainants' subjective experiences. *Police Q*. 21(1):77–108.
26 Kerstetter WA, Rasinski KA. 1994. Opening a window into police internal affairs: impact of procedural justice reform on third-party attitudes. *Soc Justice Res*. 7(2):107–127.
27 Peltzman S. 1976. Toward a more general theory of regulation. *J Law Econ*. 19(2):211–240.
28 Dal Bó E. 2006. Regulatory capture: a review. *Oxf Rev Econ Policy*. 22(2):203–225.
29 Arvai JL, Froschauer A. 2010. Good decisions, bad decisions: the interaction of process and outcome in evaluations of decision quality. *J Risk Res*. 13(7):845–859.
30 Esaiasson P, Persson M, Gilljam M, Lindholm T. 2019. Reconsidering the role of procedures for decision acceptance. *Br J Polit Sci*. 49(1):291–314.
31 Worden RE, McLean SJ. *Mirage of police reform*. University of California Press, 2017.
32 Berinsky AJ, Margolis MF, Sances MW. 2014. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am J Pol Sci*. 58(3):739–753.
33 Hibbing JR. *Stealth democracy: Americans' beliefs about how government should work*. Cambridge University Press, 2002.
34 Lupia A. 1994. Shortcuts versus encyclopedias: information and voting behavior in California insurance reform elections. *Am Polit Sci Rev*. 88(1):63–76.
35 Mourtgos SM, Adams IT. 2020. Assessing public perceptions of police use-of-force: legal reasonableness and community standards. *Justice Q*. 37(5):869–899.