SUPPORTING INFORMATION FOR

"WHY DO PEOPLE COOPERATE WITH THE POLICE AND CRIMINAL COURTS? A TEST OF PROCEDURAL JUSTICE THEORY IN 30 COUNTRIES"*

JONATHAN JACKSON, JOUNI KUHA, BEN BRADFORD, and MIKE HOUGH

Comment [SVV1]: AU: Please note that this file serves as a camera-ready copy. Unlike the text, which still needs typesetting, this file will be uploaded online as shown with any comments/tracked changed removed. So please review the file as a final product.

Although the main text and figures/tables must match CRIM style, the material in the supporting information only needs to be formatted consistently with itself.

* Published in *Criminology*, volume 63, issue 3, 2025.

© 2025 American Society of Criminology doi: 1.1111/j.1745–9125.2011.00245.x

CONTENTS

S1: SURVEY MEASURES
S2: COMMENTS ON SURVEY DEVELOPMENT AND COMPARABILITY
S3: STATISTICAL MODELING: SPECIFICATION AND ESTIMATION
S4: TABLES OF RESULTS FROM THE ESTIMATED MODELS
S5: ADDING POLICE-INITIATED CONTACT TO THE MODELS
S6: FURTHER COMMENTS ON POLICE EFFECTIVENESS AS PREDICTOR OF COOPERATION
S7: SUPPLEMENTARY MODELS FOR SOUTH AFRICA
S8: R CODE AND FULL COUNTRY-SPECIFIC MODEL DETAILS HAVE BEEN UPLOADED AS A SEPARATE PDF FOR PEER REVIEW BECAUSE OF THE SIZE (ONCE ACCEPTED, THEY WILL BE PUT ONLINE)

Supplement S1. Survey measures

Table S1. Full wording

ESS	South Africa	US
Perceived police procedural justice:		
Now some questions about when the police deal with crimes like house burglary and physical assault.	*Now some questions about when the police deal with crimes like house burglary and physical assault.*	*Thinking about the police in your community*
[1] Based on what you have heard or your own experience how often would you say the police generally treat people in [country] with respect?	*[1] Based on what you have heard or your own experience how often would you say the police generally treat people in South Africa with respect?*	*[1] How often do the police treat the people with dignity and respect?*
[2] About how often would you say that the police make fair, impartial decisions in the cases they deal with?	*[2] About how often would you say that the police make fair, impartial decisions in the cases they deal with?*	*[2] How often do the police make fair and impartial decisions in the cases they deal with?*
[3] And when dealing with people in [country], how often would you say the police generally explain their decisions and actions when asked to do so?	*[3] And when dealing with people in South Africa, how often would you say the police generally explain their decisions and actions when asked to do so?*	*[3] How often do the police explain their decisions and actions in ways that people can understand?*
Options: Not at all often (1), Not very often (2), Often (3), Very often (4)	*Options: Not at all often (1), Not very often (2), Often (3), Very often (4)*	*Options: Almost never (1), Sometimes (2), Often (3), Almost always (4)*
[1], [2], [3] used as items in factor analysis measurement model of perceived procedural justice.		
Perceived police distributive justice:		
Now some questions about whether or not the police in [country] treat victims of crime equally. Please answer based on what you have heard or your own experience.	*Now some questions about whether or not the police in South Africa treat victims of crime equally. Please answer based on what you have heard or your own experience.*	*…"your community" instead of name of the respondent's country*
[1] When victims report crimes, do you think the police treat rich people worse, poor people worse, or are rich and poor treated equally?	*[1] When victims report crimes, do you think the police treat rich people worse, poor people worse, or are rich and poor treated equally?*	*[1] When victims report crimes do you think that the police… Treat rich people worse than others (0), Treat rich and poor people equally (1), Treat rich people better than others (0)*
Options for [1]: rich people treated worse (0), poor people treated worse (0), rich and poor treated equally (1)	*Options for [1]: rich people treated worse (0), poor people treated worse (0), rich and poor treated equally (1)*	
[2] And when victims report crimes, do you think the police treat some people worse because of their race or ethnic group or is everyone treated equally?	*[2] And when victims report crimes, do you think the police treat some people worse because of their race or ethnic group or is everyone treated equally?*	*[2] When victims report crimes do you think that the police… Treat white people worse than minorities (0), Treat people of different ethnicities equally (1), Treat white people worse than minorities (0)*
Options for [2]: People from a different race or ethnic group than most [country] people treated worse (0), People from the same race or ethnic group as most [country] people treated worse (0), Everyone treated equally regardless of their race or ethnic group (1)	*Options for [2]: White, Indian and Coloured South Africans are treated worse than black South Africans (0), Black South Africans are treated worse than other race groups (0), [equally, as in ESS] (1)*	
Combined measure of distributive justice: [1]+[2]		
Perceived police lawfulness:		

Now one last question about the police and things they may or may not do.	*[No introduction]*	*Do you disagree or agree that:*
[1] How often would you say that the police in [country] take bribes?	*[1] How often would you say that the police in South Africa take bribes?*	*[1] The police take bribes*
Options for [1]: 11-point scale, from Never (0) to Always (10)	*Options for [1]: 11-point scale, from Never (0) to Always (10)*	*Options for [1]: Disagree strongly (5), Disagree (4), Neither (3), Agree (2), Agree strongly (1)*
[2] … please say to what extent you agree or disagree with each of the following statements about the police in [country]… The decisions and actions of the police are unduly influenced by pressure from political parties and politicians.	*[2] … please say to what extent you agree or disagree with each of the following statements about the police in South Africa… The decisions and actions of the police are unduly influenced by pressure from political parties and politicians.*	*[2] The decisions and actions of the police are unduly influenced by pressure from political parties and politicians.*
Options for [2]: Agree strongly (1), Agree (2), Neither agree nor disagree (3), Disagree (4), Disagree strongly (5)	*Options for [2]: Agree strongly (1), Agree (2), Neither agree nor disagree (3), Disagree (4), Disagree strongly (5)*	*Options for [2]: Disagree strongly (5), Disagree (4), Neither (3), Agree (2), Agree strongly (1)*
Combined measure of lawfulness: average of [1] and [2]		
Perceived police effectiveness:		
[1] Based on what you have heard or your own experience how successful do you think the police are at preventing crimes in [country] where violence is used or threatened?	*[1] Based on what you have heard or your own experience how successful do you think the police are at preventing crimes in South Africa where violence is used or threatened?*	*[1] How successful do you think the police are at preventing crimes where violence is used or threatened in your community?*
[2] And how successful do you think the police are at catching people who commit house burglaries in [country]?	*[2] And how successful do you think the police are at catching people who commit house burglaries in South Africa?*	*[2] And, how successful do you think the police are at catching people who commit house burglaries?*
Options for [1] and [2]: 11-point scale, from Extremely unsuccessful (0) to Extremely successful (10)	*Options for [1] and [2]: 11-point scale, from Extremely unsuccessful (0) to Extremely successful (10)*	*Options for [1] and [2]: 11-point scale, from Extremely unsuccessful (0) to Extremely successful (10)*
[3] If a violent crime were to occur near to where you live and the police were called, how slowly or quickly do you think they would arrive at the scene?	*[3] If a violent crime or house burglary were to occur near to where you live and the police were called how slowly or quickly do you think they would arrive at the scene?*	*[3] If a violent crime were to occur near to where you live and the police were called, how soon do you think they would arrive at the scene?*
Options for [3]: 11-point scale, from Extremely slowly (0) to Extremely quickly (10)	*Options for [3]: 11-point scale, from Extremely slowly (0) to Extremely quickly (10)*	*Options for [3]: 11-point scale, from Extremely slowly (0) to Extremely quickly (10)*
[1], [2], [3] used as items in factor analysis measurement model of perceived effectiveness.		
Worry about criminal victimization:		
		Now I would like to ask you about how much (if at all) you worry about specific crimes. How worried are you about …
[1] How often, if at all, do you worry about your home being burgled?	*[1] How often do you worry about your home being burgled?*	*[1] Having your home broken into and something stolen*
[2] How often, if at all, do you worry about becoming a victim of violent crime?	*[2] How often do you worry about becoming a victim of*	*Being mugged or robbed*

	violent crime?	
Options: All or most of the time (4), Some of the time (3), Just occasionally (2), Never (1)	*Options: All or most of the time (4), Some of the time (3), Just occasionally (2), Never (1)*	*Options: Very worried (4), Fairly worried (3), Not very worried (2), Not at all worried (1)*
Combined measure of worry about victimization: average of [1] and [2]		
Normative alignment with the police:		
…please say to what extent you agree or disagree with each of the following statements about the police in [country].	*…please say to what extent you agree or disagree with each of the following statements about the police in South Africa.*	*Do you disagree or agree that:*
[1] The police generally have the same sense of right and wrong as I do.	*[1] The police generally have the same sense of right and wrong as I do.*	*[1] The police generally have the same sense of right and wrong as I do.*
[2] The police stand up for values that are important to people like me.	*[2] The police stand up for values that are important to people like me.*	*[2] The police stand up for values that are important to people like me.*
[3] I generally support how the police usually act.	*[3] I generally support how the police usually act.*	*[3] You generally support how the police act in your community.*
Options: Agree strongly (5), Agree (4), Neither agree nor disagree (3), Disagree (2), Disagree strongly (1)	*Options: Agree strongly (5), Agree (4), Neither agree nor disagree (3), Disagree (2), Disagree strongly (1)*	*Options: Agree strongly (5), Agree (4), Neither agree nor disagree (3), Disagree (2), Disagree strongly (1)*
[1], [2], [3] used as items in factor analysis measurement model of normative alignment.		
Duty to obey the police:		
Now some questions about your duty towards the police in [country]… To what extent is it your duty to…	*Now some questions about your duty towards the police in South Africa… To what extent is it your duty to…*	*Now some questions about the police in your community. Do you disagree or agree that it is your responsibility that:*
[1] back the decisions made by the police even when you disagree with them?	*[1]* was not included in the survey	*[1] You should support the decisions made by police officers even when you disagree with them.*
[2] do what the police tell you even if you don't understand or agree with the reasons?	*[2] do what the police tell you even if you don't understand or agree with the reasons?*	*[2] do what the police tell you even if you don't understand or agree with the reasons?*
[3] do what the police tell you to do, even if you don't like how they treat you?	*[3] do what the police tell you to do, even if you don't like how they treat you?*	*[3] do what the police tell you to do, even if you don't like how they treat you?*
Options: 11-point scale, from Not at all my duty (0) to Completely my duty (10)	*Options: 11-point scale, from Not at all my duty (0) to Completely my duty (10)*	*Options: Disagree strongly (0), Disagree (2.5), Neither (5), Agree (7.5), Agree strongly (10)*
[1], [2], [3] used as items in factor analysis measurement model of obligation to obey.		
Willingness to cooperate with the police and courts:		
Now some questions about what you would do if you were the only witness to a crime.	*Now some questions about what you would do if you were the only witness to a crime.*	*Imagine that you were the only witness to a crime.*
[1] Imagine that you were out and saw someone push a man to the ground and steal his wallet. How likely would you be to call the police? Would you be…	*[1] Imagine that you were out and saw someone push a man to the ground and steal his wallet. How likely would you be to call the police? Would you be…*	*[1] If you saw someone push a person to the ground and steal their purse or wallet how likely would you be to call the police?*
Options for [1]: Not at all likely (1), Not very likely (2), Likely (3), Very likely (4)	*Options for [1]: Not at all likely (1), Not very likely (2), Likely (3), Very likely (4)*	*Options for [1]: Very unlikely (1), Unlikely (2), Likely (3), Very likely (4)*
[2] How willing would you be to identify the	*[2] How willing would you be*	*[2] How willing would you be*

person who had done it? Would you be…	*to identify the person who had done it? Would you be…*	*to identify the person who had committed the crime?*
[3] And how willing would you be to give evidence in court against the accused? Would you be…	*[3] And how willing would you be to give evidence in court against the accused? Would you be…*	*[3] And how willing would you be to give evidence in court against the accused? Would you be…*
Options for [2] and [3]: Not at all willing (1), Not very willing (2), Willing (3), Very willing (4)	*Options for [2] and [3]: Not at all willing (1), Not very willing (2), Willing (3), Very willing (4)*	*Options for [2] and [3]:* Same as for [1], even though question stem has "willing" rather than "likely".
[1], [2], [3] used as items in factor analysis measurement model of willingness to cooperate.		
Public contact with the police:		
[1] In the past 2 years, did the police in [country] approach[1] you[2], stop you or make contact with you for any reason?	*[1] In the past 2 years, did the police in South Africa approach you, stop you or make contact with you for any reason?*	*[1] In the past 2 years, have the police in your community approached you, stopped you or made contact with you for any reason?*
Options for [1]: No (1), Yes (2)	*Options for [1]: No (1), Yes (2)*	*Options for [1]: No (1), Yes (2)*
[2] [Ask if has been approached / stopped / contacted by the police for any reason in past 2 years] *How dissatisfied or satisfied were you with the way the police treated[3] you the last time this happened?*	*[2]* [Ask if has been approached / stopped / contacted by the police for any reason in past 2 years] *How satisfied or dissatisfied were you with the way the police treated you the last time this happened?*	*[2]* [Ask if has been approached / stopped / contacted by the police for any reason in past 2 years] *How satisfied or dissatisfied were you with your experience with the police the last time this happened?*
Options for [2]: Very dissatisfied (1), Dissatisfied (2), Neither dissatisfied nor satisfied (3), Satisfied (4), Very satisfied (5)	*Options for [2]: Very dissatisfied (1), Dissatisfied (2), Neither dissatisfied nor satisfied (3), Satisfied (4), Very satisfied (5)*	*Options for [2]: Very dissatisfied (1), Dissatisfied (2), Neither dissatisfied nor satisfied (3), Satisfied (4), Very satisfied (5)*

[1] Approach for any reason (e.g. to ask you for information or because they suspect you have committed a crime or they need to ask you to do something).
[2] 'You' as in 'the police approached, stooped or contacted the respondent personally'.
[3] 'Treated' in the sense of how the police responded to or dealt with the respondent.

Procedural justice

Our scale of procedural justice captures three key dimensions of fair treatment and decision-making during police-citizen encounters. The first focuses on dignity and respect—whether officers acknowledge individuals' rights and treat them with courtesy. When they do, people are more likely to feel they have been treated fairly; when they do not—when officers stereotype, belittle or dismiss—interactions can feel dehumanising and unjust. The second item assesses neutrality: do officers make decisions based on consistent rules rather than personal bias or prejudice? Neutral decision-making signals that everyone is treated equally under the law. The third item concerns voice—whether people feel able to share their side of the story. Voice does not require formal procedures; even brief opportunities to explain oneself can make encounters feel more fair. Research on police stops suggests that when officers allow individuals to speak before acting, people are more likely to judge the process as fair, regardless of the outcome.

Distributive justice

Distributive justice refers to the fair distribution of policing resources and outcomes—how protection, attention and investigative effort are allocated across different social groups. Our measures capture public perceptions of whether victims receive equal treatment from the police, focusing specifically on differences by wealth and race. By anchoring these items in concrete victim group comparisons, we aim to make the concept more tangible and reduce ambiguity. While designed to tap into material aspects of policing—such as response times, prioritisation of cases and allocation of investigative resources—we do recognise that the phrase *treating people equally* may also evoke notions of interpersonal fairness, including demeanour and decision-making. See section S2 for further discussion.

Effectiveness, lawfulness and fear of crime

Effectiveness was measured using items that reflect key domains of police performance: crime detection, crime prevention and citizen protection, across property and violent crime. Perceived lawfulness was assessed using two items: (1) how often respondents believed police in their country accept bribes; and (2) the extent to which they thought police actions are unduly influenced by political parties or politicians. Fear of crime was measured by asking how frequently respondents worry about two specific risks: having their home burgled and becoming a victim of violent crime (see Jackson & Kuha, 2014 for details).

Legitimacy

Legitimacy was measured along two dimensions (Jackson et al., 2012, 2013). The first, normative alignment, captures whether respondents believe the police act in ways that reflect their own values and expectations of appropriate conduct. This belief signals more than approval—it reflects the perception that police exercise power in a morally appropriate way. When people see police behaviour as consistent with their normative standards, they are more likely to view the institution as rightful, justified, and worthy of support.

The second dimension, duty to obey, reflects a felt obligation to follow police directives—not out of fear, habit or convenience, but because the police are seen as legitimate authorities. This sense of obligation goes beyond passive compliance; it represents an active endorsement of the police's right to determine appropriate behaviour. The concept echoes Raz's (1986: 135) idea of *pre-emptive reasons*: individuals defer not because they always agree, but because they recognise the institution's rightful authority to decide. The items are designed to capture exactly this sense of positive obligation—compliance as a moral, civic or legal duty, rooted in democratic citizenship. See section S2 for further detail.

Willingness to cooperate

Cooperation was assessed through a vignette-based measure, reflecting a latent continuum of engagement with the justice system. At the lower end are those unlikely to report or assist; at the upper end are individuals willing to report a crime, identify the suspect and give evidence in court. This allows us to capture variation in both willingness and depth of cooperative behaviour in response to a hypothetical offence.

Police contact (for supplementary analysis)

In section S5 we supplement our main analyses by incorporating two measures of direct police contact: (1) a binary item asking whether respondents had been approached, stopped or otherwise contacted by the police

in the past two years; and (2) an ordinal measure of satisfaction, in which those with contact rated how satisfied or dissatisfied they were with how they were treated.

REFERENCES

Jackson, J., Bradford, B., Hough, M., Myhill, A., Quinton, P., & Tyler, T. R. (2012). Why do people comply with the law? Legitimacy and the influence of legal institutions. *British Journal of Criminology, 52*(6), 1051-1071. https://doi.org/10.1093/bjc/azs032

Jackson, J., Bradford, B., Stanko, E. A., & Hohl, K. (2013). *Just Authority? Trust in the Police in England and Wales*. Oxon: Routledge.

Jackson, J. & Kuha, J. (2014). 'Worry about Crime in a Cross-National Context: A Model-Supported Method of Measurement Using European Social Survey Data', *Survey Research Methods*, 8, 2, 109-126, doi: 10.18148/srm/2014.v8i2.5457

Raz, J. (1986). *The Morality of Freedom*. Clarendon Press.

Supplement S2. Comments on survey development and comparability

The ESS data come from the *Trust in Justice* rotating module in Round 5 (Hough et al., 2013a, 2013b). Rotating modules are selected through open competition, where academic teams propose new content addressing pressing social and political issues. Each proposal outlines the substantive aims, theoretical frameworks, concepts and indicative measures using the ESS *Question Module Design Template*. Once selected, the module undergoes a multi-stage development process.

For *Trust in Justice*, this process spanned 18 months (Jackson et al., 2011). A multidisciplinary advisory panel reviewed the proposed questions and assessed their reliability and validity using the *Survey Quality Predictor Program*. The questions underwent multiple rounds of revision and reassessment before being presented to ESS National Coordinators for further evaluation of substantive and translation-related issues (see here for detailed documentation of the questionnaire development and testing process). The questionnaire was then tested in a two-nation quantitative pilot study, which included split-ballot multi-trait/multi-method experiments. Pilot data were analyzed for item non-response, factor structure, correlations and translation quality. This was followed by additional expert review and further consultations with National Coordinators. The finalized questionnaire incorporated annotated clarifications for translation and was subsequently translated into all relevant languages before fieldwork began.

To illustrate the process, we focus on two constructs—distributive justice and duty to obey. We highlight these for two reasons. First, the exact item wordings in our study departed slightly from prior approaches. Second, both constructs raise conceptual challenges worth clarifying.

Distributive justice

Procedural justice and distributive justice are typically distinguished as follows:

- Procedural justice concerns the fairness of processes in individual police-citizen encounters:
    - *Decision-making fairness*, whether officers make objective and unbiased decisions when dealing with the public (e.g., *"Do officers make fair and impartial decisions in their interactions?"*)
    - *Interpersonal treatment*, whether officers treat individuals with respect and dignity during encounters (e.g., *"Do officers engage with people in a respectful and professional manner?"*).
- By contrast, distributive justice refers to the fair distribution of legal outcomes and policing resources across aggregate social groups in society:
    - *Equity in legal outcomes*, whether certain groups receive the legal outcomes they deserve while others do not (e.g., *"Are legal outcomes applied fairly across different social groups?"*)
    - *Equitable allocation of policing resources*, whether some groups receive better policing services than others (e.g., *"Do certain social groups experience greater police responsiveness and protection?"*) and whether some groups receive more of the burdens of policing than other groups (e.g., *"Are some social groups more frequently stopped, searched, or surveilled by police than others?")*

This distinction separates (a) fairness in personal encounters from (b) fairness in the structural distribution of policing benefits and burdens. While procedural justice focuses on the micro-level—how officers treat individuals—distributive justice addresses systemic fairness at the macro-level, especially how outcomes are distributed across race, class and other group categories.

As such, measures of procedural justice typically focus on respectful treatment, neutral decision-making and voice provision in encounters between officers and citizens.[4] By contrast, measures of distributive justice typically stress resources and outcomes, mostly in terms of fair allocation across core societal groups. Consider some common survey indicators. Sunshine & Tyler's (2003) used the following in study 1: "(1) ''How often do people receive the outcomes they deserve under the law when they deal with

[4] To be sure, procedural justice perceptions can be aggregated up. People can believe that police *generally* treat people with respect and dignity in individual police-citizen encounters—this is the belief that officers 'do' procedural justice *en masse*. But when aggregated up, procedural justice as a concept does not reference systematic differences between social groups in society—that is the domain of distributive justice.

the police?,'' (2) ''Are the outcomes that people receive from the police better than they deserve, worse than they deserve, or about what they deserve under the law?,'' (3) ''How often do the police give people in your neighborhood less help than they give others due to their race?,'' (4) ''The police do not provide the same quality of service to people living in all areas of the city,'' and (5) ''Minority residents of the city receive a lower quality of service from the NYPD than do whites." In study 2 the measures were: "…respondents whether eight groups received the quality of service they deserved from the police: people like the respondent, people in their neighborhood, minorities in their neighborhood, whites, African Americans, Hispanics, poor people, and wealthy people…Respondents were also asked whether (1) ''The police treat everyone equally regardless of their race,'' (2) ''The police provide better services to the wealthy (reversed),'' and (3) ''They sometimes give minorities less help due to their race (reversed).''[5]

How did we measure distributive justice? At the start of the design process, as rotating module designers we initially proposed three existing distributive justice indicators for the *Trust in Justice* rotating module: (A) 'Provide the same quality of service to everyone,' (B) 'Enforce the law consistently when dealing with all people,' and (C) 'Make sure people receive the outcomes they deserve under the law.' In response, the methodological board raised three main queries. For (A) it was unclear what 'quality of service' encompassed. For (B) and (C) the term 'people' lacked specificity—did it refer to criminals, witnesses, or victims? For (C) the phrase 'receive the outcomes' was ambiguous—did it refer to police decisions, legal punishments or broader judicial outcomes?

After several rounds of consultation, the focus shifted to victims—a clearer, more culturally universal frame—and specifically whether some victim groups were given a better quality of service by the police than other victim groups. The methodological development panel took the view that specifying *victims of crime* added clarity to the type of police-citizen encounter, helping to anchor the scenario in respondents' minds across all 28 countries. Upon further consideration, however, the phrase *same quality of service* was considered too culturally specific and difficult to translate. Because the panel believed that *quality of service* was not a universally understood concept in all countries, particularly outside the UK, the questions moved to *treating people equally.* As a clearer and more universally intelligible alternative, this hoped to ensure cross-national consistency, capturing whether respondents in each of the 28 countries perceived the police as prioritizing or privileging certain groups over others when dealing with and responding to crime victims.

The revised items were: *[1] "When victims report crimes, do you think the police treat rich people worse, poor people worse, or are rich and poor treated equally?" and [2] "When victims report crimes, do you think the police treat some people worse because of their race or ethnic group or is everyone treated equally?".* ESS interviewers were given the following clarifying note, to be used if survey respondents queried what *treat people* means exactly: "*Treat in the sense of how the police respond to and deal with people*".

Blurring lines? Distributive and procedural justice

Despite our efforts to distinguish them, some interpretive overlap between procedural and distributive justice is likely. For instance, the idea that certain groups are *treated worse* may not only signal perceptions of unequal resource allocation (distributive injustice) but also evoke concerns about disrespectful treatment and biased decision-making (procedural injustice)—just scaled up to group level. If people believe that poor victims are treated worse than rich victims, they may be thinking not just of slower response times or less investigative effort, but also of unequal demeanour and unfair judgements.

Two further points are worth emphasising. First, many respondents may assume that fair processes lead to fair outcomes. If police are perceived as disrespectful or biased toward minority victims, people may infer that those groups are also less likely to receive adequate protection or support. In this way, procedural unfairness can be read as a sign of distributive injustice. Second, people may substitute one fairness judgment for another. Evaluating systemic resource allocation is cognitively demanding. It is often easier to

[5] Reisig et al. (2007) asked "Provide the same quality of service to all citizens", "Enforce the law consistently when dealing with all people", "Make sure citizens receive the outcomes they deserve under the law", "Give minorities less help because of their race" and "Provide better services to wealthier citizens". Wolfe et al. (2016) asked whether police in their neighborhood ''give minorities less help because of their race'' and ''provide better services to wealthy citizens''. Jackson et al. (2023) asked "The police treat everyone fairly, regardless of who they are", "The police treat everyone equally", "The police provide the same quality of service to all citizens", and "The police enforce the law consistently when dealing with people."

judge fairness from visible cues—tone of voice, body language, the officer's attitude. Respondents may use these interpersonal signals as proxies for deeper systemic concerns.

That said, we would defend the view that our items primarily capture perceived disparities in outcomes and resources across wealth and race lines. While procedural concerns may inform these perceptions, the questions are explicitly anchored in group-level, structural fairness. The focus is on what happens after a crime is reported—whether police follow up, prioritise, and support victims equitably across groups.

Of course, in the real world, procedural and distributive injustices often go hand in hand. Disrespectful treatment and biased decisions can accumulate into institutional patterns of inequality. Our goal was to capture public beliefs about whether some groups are systematically disadvantaged in how police services are delivered—a core concern for public trust in justice across diverse societies.

Duty to obey

When considering how to measure felt obligation to obey the police, a key priority was to try to distinguish between two distinct types of obligation:

- Instrumentally grounded obligations arise from pragmatism, fear of consequences, or a sense of powerlessness, where compliance is motivated by external necessity rather than an internalized belief in the police's moral authority.
- Normatively grounded obligations reflect a freely chosen, morally driven sense of duty to obey the police, based on ascribed legitimacy and rightful authority.

One concern in the design process was that standard measures of legitimacy could conflate these two motivations (see Bottoms and Tankebe, 2012; Tankebe, 2009; Tankebe, 2013; Johnson et al., 2014; Posch et al., 2021; Van Petegem et al., 2021; Reisig et al., 2023; Reisig & Trinkner, 2024). In particular, survey respondents may agree with statements about obligation for reasons that blend instrumental compliance (e.g., fear of punishment) with normative commitment (e.g., belief in rightful authority).

To address this ambiguity, we designed measures that captured normatively grounded obligation, emphasizing the concept of a moral duty to obey the police. To aid this, ESS interviewers were provided with the following clarifying note to use if respondents asked for clarification on the term duty: "*Duty in the sense of a citizen's moral duty to the state*." This approach aims to align the measurement process with Raz's (1986) concept of pre-emptive consent, in which obligation arises from an internalized commitment to the police's authority rather than from a calculated assessment of potential consequences.

To assess how well the concept translated across languages, we asked the ESS team for a detailed breakdown of how duty—specifically, the idea of a moral duty to obey the police—was referred to in different languages within Round 5 of the ESS. The key question was whether these translations accurately captured legitimacy—that is, people obeying the police because they recognize their rightful authority—or whether they left room for interpretations rooted in instrumental compliance, where obedience stems from fear of punishment or a sense of powerlessness.

We find that many of the translations explicitly conveyed a sense of moral or civic responsibility. In German, *Pflicht* denotes an obligation arising from ethical, moral, or religious reasons, closely aligning with the intended meaning. Similarly, the French *devoir* is defined as a moral duty independent of legal enforcement, reinforcing that the obligation to obey is based on legitimacy rather than coercion. The Russian *долг* carries strong moral connotations, often used in phrases such as *"a man's duty to protect his family."* Other languages, such as Portuguese (*dever*) and Croatian (*dužnost*), also employ terms that clearly denote an obligation rooted in moral or civic values. In some cases, translations went even further to specify moral duty explicitly. For instance, the Hebrew and Arabic versions of the survey included clarifications to emphasize that the duty in question was moral, rather than legal or coercive.

Beyond individual moral duty, some translations frame the concept in terms of *civic duty*—a citizen's responsibility to the state. For example, the Norwegian version used *borgerplikt*, which directly translates to *citizen's duty*. In Russian and Ukrainian, the phrase *гражданский долг* (*civic duty*) was used, reinforcing the idea that duty to the police is rooted in legitimacy within the broader social contract rather than individual fear of punishment. Similarly, the Greek *υποχρέωση* conveys a blend of moral and personal obligation rather than a strictly legal or externally imposed duty.

However, in a few cases, the translations are less precise in distinguishing moral duty from other forms of obligation. The Czech *povinnost* and Polish *obowiązek* can refer to both moral and legal duties, allowing for multiple interpretations. Similarly, the Estonian *kohus* can mean *duty, obligation, or responsibility*, making it more open-ended than some of the stronger moral translations. The Swiss-German term *verpflichtet sein* (*to be bound to*) lacks an explicit moral component. In some languages, particularly Lithuanian and Ukrainian, softer wording is used, such as modal verbs (*should* rather than *must*), which may slightly weaken the normative force of the term.

Despite these minor variations, the overwhelming pattern across languages is that duty was translated in a way that captures moral responsibility rather than coercion. Overall, the translation analysis suggests that, in most cases, the wording preserves the normative, moral dimension of duty rather than implying fear-based obedience. In most languages, respondents likely interpret the question as assessing their genuine belief in the police's right to issue directives, rather than whether they comply due to fear or powerlessness.

We acknowledge, however, a limitation in how obligation to obey was measured in the US survey. While the ESS measures explicitly framed obligation as a normative duty, the US survey relied on traditional measures that asked respondents to what extent they believe they should comply with police directives. These measures may conflate normative motivation (a moral duty to obey) with instrumental compliance (e.g., obeying out of fear of punishment or strategic self-interest).

Future cross-national research should refine measures of obligation to obey by explicitly distinguishing between different motivations for compliance. It is essential to ensure that survey questions across countries and languages accurately capture the internalized normative basis of legitimacy, rather than instrumental forms of compliance rooted in coercion or strategic calculation. By adopting measures that clearly differentiate between duty-based and instrumental obedience, future comparative studies can provide a more precise assessment of the role legitimacy plays in shaping public cooperation with legal authorities across different jurisdictions.

Differences in framing of the police comparing the ESS/South Africa to US

We should note that the framing of *the police* in the overall question set varied slightly across surveys. The ESS and SASAS instructed respondents to consider the police responsible for crimes such as house burglary and physical assault, while the US survey asked about *police in your community*. This distinction was intentional, reflecting the diverse policing structures across countries. In some nations, such as France and Italy, different police forces handle different types of policing, which could create ambiguity if respondents were asked about *the police* in general. By anchoring questions to common neighbourhood-level crimes, we aimed to ensure that respondents across different contexts were thinking about everyday policing—specifically, the officers most likely to engage with the public in routine interactions. In the US, *police in your community* does the same work. While we recognize this as a minor difference in question framing, we believe it ultimately strengthens the validity of cross-national comparisons by reducing ambiguity in respondent interpretation and directing attention to the most relevant aspects of procedural justice within each policing system.

REFERENCES

Bottoms, A., & Tankebe, J. (2012). Beyond procedural justice: A dialogic approach to legitimacy in criminal justice. *Journal of Criminal Law & Criminology*, *102*(1), 119–170.

Jackson, J., Bradford, B., Giacomantonio, C. & Mugford, R. (2023). Developing core national indicators of public attitudes towards the police in Canada. *Policing & Society*, 33, 3, 276-295, https://doi.org/10.1080/10439463.2022.2102757

Jackson, J., Bradford, B., Hough, M., Kuha, J., Stares, S. R., Widdop, S., Fitzgerald, R., Yordanova, M. & Galev, T. (2011). Developing European indicators of trust in justice. *European Journal of Criminology*, 8, 4, 267-285, https://doi.org/10.1177/1477370811411458

Johnson, D., Maguire, E. R., & Kuhns, J. B. (2014). Public perceptions of the legitimacy of the law and legal authorities: Evidence from the Caribbean. *Law & Society Review, 48*(4), 947–978, https://doi.org/10.1111/lasr.12102

Pósch, K., Jackson, J., Bradford, B. & MacQueen, S. (2021). '"Truly free consent"? Clarifying the nature of police legitimacy using causal mediation analysis. *Journal of Experimental Criminology*, 17, 4, 563-595, https://doi.org/10.1007/s11292-020-09426-x

Reisig, M. D., & Trinkner, R. (2024). Distinguishing between normative and non-normative motivations to obey the police: furthering the development of a police legitimacy scale. *Policing: An International Journal*, *47*(1), 50-65, https://doi.org/10.1108/PIJPSM-05-2023-0069

Reisig, M. D., Bratton, J., & Gertz, M. G. (2007). The construct validity and refinement of process-based policing measures. *Criminal Justice and Behavior, 34*(8), 1005-1028. https://doi.org/10.1177/0093854807301275

Reisig, M. D., Trinkner, R., & Sarpong, D. (2023). Measuring normative obligation to obey the police: An empirical assessment of a new police legitimacy scale. *Journal of Criminal Justice*, *86*, 102045, https://doi.org/10.1016/j.jcrimjus.2023.102045

Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law and Society Review, 37*(3), 555-589. https://doi.org/10.1111/1540-5893.3703002

Tankebe, J. (2009). Public cooperation with the police in ghana: Does procedural fairness matter? *Criminology*, *47*(4), 1265–1293. https://doi.org/10.1111/j.1745-9125.2009.00175.x

Tankebe, J. (2013). Viewing things differently: The dimensions of public perceptions of police legitimacy. *Criminology*, *51*(1), 103–135. https://doi.org/10.1111/j.1745-9125.2012.00291.x

Van Petegem, S., Trinkner, R., van der Kaap-Deeder, J., Antonietti, J. P., & Vansteenkiste, M. (2021). Police procedural justice and adolescents' internalization of the law: Integrating self-determination theory into legal socialization research. *Journal of Social Issues*, *77*(2), 336-366, https://doi.org/10.1111/josi.12437

Wolfe, S. E., Nix, J., Kaminski, R., & Rojek, J. (2016). Is the effect of procedural justice on police legitimacy invariant? Testing the generality of procedural justice and competing antecedents of legitimacy. *Journal of Quantitative Criminology*, *32*, 253-282, https://doi.org/10.1007/s10940-015-9263-8

Supplement S3. Statistical modelling: Specification and estimation

In this section we describe the specification, motivation and estimation of the statistical models that were used in our analysis. Let $\boldsymbol{X}_{ji}=(X_{1ji},X_{2ji},X_{3ji},X_{4ji},X_{5ji},X_{6ji},\ X_{7ji},\eta_{1ji},\eta_{2ji})$ denote the exogenous variables, i.e. the ones furthest to the left in the path diagram in Figure 1, for respondent i=1,…,n_j in country j=1,…,J=30. These include the observed covariates (age and dummy variables for female gender and upper secondary and higher education levels; X_{1ji} – X_{4ji} respectively), plus the observed measures of perceived distributive justice (X_{5ji}), lawfulness of the police (X_{6ji}) and the respondent's worry about crime (X_{7ji}), together with the latent variables for procedural justice (η_{1ji}) and effectiveness (η_{2ji}). We denote the rest of the latent variables as η_{3ji} for normative alignment with the police, η_{4ji} for duty to obey the police, and η_{5ji} for willingness to cooperate with the police and the courts. Our model of interest is then of the form

$$\eta_{3ji} = \beta_{30j} + \beta_{3xj}\boldsymbol{X}_{ji} + \varepsilon_{3ji} \qquad (1)$$
$$\eta_{4ji} = \beta_{40j} + \beta_{4xj}\boldsymbol{X}_{ji} + \varepsilon_{4ji} \qquad (2)$$
$$\eta_{5ji} = \beta_{50j} + \beta_{5xj}\boldsymbol{X}^{*}_{ji} + \beta_{53j}\eta_{3ji} + \beta_{54j}\eta_{4ji} + \varepsilon_{5ji} \qquad (3)$$

where the β-terms are parameters (regression coefficients) and the ε-terms are normally distributed residuals with mean 0 and variances $var(\varepsilon_{kji}) = \sigma_k^2$ for k=3,4,5 and one non-zero covariance $cov(\varepsilon_{3ji},\varepsilon_{4ji}) = \sigma_{34}$. Here ε_{5ji} is taken to be independent of $(\varepsilon_{3ji},\varepsilon_{4ji})$ for respondent ji, and $(\varepsilon_{3ji},\varepsilon_{4ji},\varepsilon_{5ji})$ all independent of each other for different respondents. In (3), $\boldsymbol{X}^{*}_{ji}$ omits distributive justice and lawfulness from $\boldsymbol{X}_{ji}$. This is a conventional linear structural equation model (SEM) for the latent responses (η_{3ji}, η_{4ji}, η_{5ji}), corresponding to the model represented in Figure 1. We refer to (1)-(3) as the *structural model* of our analysis. Note that all of the parameters can have different values in every country, as indicated by the subscript j in them.

Each of the five latent variables η_{kji} is measured by three observed items Y_{klij}, for k=1,…,5 and l=1,2,3. The *measurement model* for each of them is of the form

$$Y_{klij} = \tau_{kl} + \lambda_{kl}\eta_{kji} + \delta_{klij} \qquad (4)$$

where the τ-terms (measurement intercepts) and λ-terms (factor loadings) are parameters and the δ-terms are normally distributed residuals with mean 0 and variances $var(\delta_{klji}) = \theta_{kl}^2$, and with all δ_{klij} independent of each other. This is a conventional factor analysis measurement model, separately for each of the five latent variables. Each parameter in (4) is taken to have the same value in every country j. This is the assumption of measurement equivalence of the items across the countries, as discussed further below.

We estimate the models using a "two-step" method of estimation. In its first step, only the parameters of the measurement models are estimated, separately for the measures of each latent variable and omitting all other variables. This means for each of $\eta_{1ji},\ldots,\eta_{5ji}$ we fit a factor analysis model (4), to the sample of all $n = \sum_j n_j$ respondents across all the countries. In each model the latent variable is taken to be normally distributed, with its mean fixed at 0 and variance at 1 for identification of the model. This implies that the scale of each latent variable is defined so that its marginal distribution across all the countries has mean 0 and variance 1. No further identification constraints are required in the rest of the estimation. To match the scaling of the latent variables, we also scaled the values of the observed summary measures of distributive justice, lawfulness and worry about crime so that their means and variances across all the countries were 0 and 1 respectively.

In the second step, the structural model (1)-(3) is estimated, combined with the measurement models (4) but with all the measurement parameters fixed at their estimated values from the first step. The likelihood function in this second step is what it would be for a structural equation model where (1)-(3) and (4) were estimated together, except that the step-1 estimates are substituted for τ_{kl} and λ_{kl} as known values. Thus, only the structural parameters (the βs and σs) are estimated in this step.

The basic idea and general properties of two-step estimation of this kind are described for different types of latent variable models by Bakk and Kuha (2018), Rosseel and Loh (2022), and Kuha and Bakk (2023). Two-step estimates of the structural parameters are consistent and have essentially similar large-sample properties as "one-step" estimates that would be obtained by fitting both structural and measurement parts of the model together. The two-step approach is motivated by two broad considerations. Conceptually,

by separating the estimation of the measurement and structural models, it also separates the effective *definition* of the latent variables from other variables in the structural model. This way, the measurement model is not affected even if the structural model is changed or misspecified. In our analysis, the operational definition of each of the five latent variables is thus determined only by their own indicators, without any contribution (even indirect ones) from the other variables in the model. Practically, two-step estimation can be computationally much less demanding than one-step estimation, especially for complex structural models. In our analysis, one-step estimation would involve fitting all of (1)-(4) to the pooled sample of all the countries, requiring the estimation of 1965 distinct parameters at once. The two-step approach is substantially simpler, because its second step for (1)-(3) can be done separately for each country.

Note that we do not report or discuss conventional model fit statistics (e.g., Root Mean Square Error of Approximation, Comparative Fit Index) that are commonly presented in structural equation modeling. These statistics primarily assess the specification of the *measurement model* component of a joint model. In our case however, the measurement models are specified and estimated separately for each latent variable in the structural model (as shown in Figure 1 of the paper) and then fixed at their estimated values. Each measurement model is a factor analysis for a single latent variable, with distinct survey measures identified *a priori* (as listed in Supplementary Materials S1). Because each of these models includes three indicators, they are saturated, meaning that conventional fit statistics would necessarily indicate perfect fit.

Moreover, conventional model fit statistics would not be particularly informative even if we had estimated the joint model using a one-step estimation approach. In cases where such statistics suggest poor fit—especially in large samples—the only way to adjust the measurement model to improve these indices would be to introduce cross-loadings between different sets of variables (e.g., allowing some indicators of procedural justice to also serve as indicators of police effectiveness) or to incorporate additional error correlations. However, such modifications would be both theoretically problematic and interpretationally unhelpful.

We also calculate estimates of the country-level marginal means of the variables. For the latent variables this is done by fitting for each of them separately a model of the form $\eta_{kji} \sim N(\mu_{kj}, \sigma_{kj}^2)$ in the second step, with the same measurement model from step 1. This is a structural model η_{kji} where its mean μ_{kj} and variance σ_{kj}^2 depend only on the country k. We also denote by μ_{kj} the population means of the observed measures of distributive justice, lawfulness and worry about crime, which are estimated by their means in the sample. These estimates of μ_{kj} for both latent and observed variables were calculated using the survey weights that are available for the surveys.

We used the *lavaan* package (Rosseel 2012) in the R language (R Core Team 2022) to carry out both steps of the estimation (using a likelihood-based approach as described above, rather than the closely related SAM formulation of Rosseel and Loh 2022, which is also implemented in *lavaan*). The weighted estimates of the country means of the variables were calculated using also the the *lavaan.survey* (Oberski 2014) and *survey* packages (Lumley 2010). The R code that was used for the analysis is included in supplementary materials as section S4.

Using the statistical models for comparative analysis

The basic approach of our empirical analysis is to compare estimated regression coefficients of the model (1)-(3) between and within countries, in order to examine the estimated effect sizes in the theoretical model in Figure 1 of the main manuscript across these diverse social, political and legal contexts. We also examine how the estimated means of the variables (μ_{kj}) vary across the countries, and how the regression coefficients correlate with the estimated country-level means μ_{kj} of perceived procedural justice.

For these comparisons to be meaningful, we have to first be able to assume that the variables are defined in the same way across all the countries. The first methodological part of this assumption is the careful cross-national harmonization of the survey questions, as already discussed. Second, for the latent variables we make the assumption of formal measurement equivalence across the countries, meaning that the parameters of the measurement models (4) do not depend on country k. We therefore impose on each latent variable an operational definition which is the same in each country and which is determined by the measurement model for the variable which is estimated from the pooled data for all the countries together. The estimated coefficients of all the variables then refer to the variables thus defined and are comparable across the countries in this sense.

We acknowledge that measurement equivalence could be examined empirically rather than assumed. This would involve estimating measurement models where certain parameters vary across countries,

allowing differences to be assessed using likelihood ratio tests. If significant differences were detected, these variations could be retained in the structural equation model used for the main analysis, accounting for non-equivalence in parts of the model.

However, we did not take this approach because, in cross-national analyses, we find that it often reduces rather than enhances conceptual and face-value interpretability. In large-scale surveys such as the ESS, where data are collected from numerous countries with sizable samples, formal statistical tests almost invariably detect some degree of measurement non-equivalence. Incorporating such adjustments into the fitted models would introduce a key drawback: the definition of a latent variable's scale would then differ across countries. In practical terms, this would mean that respondents from different countries who provide identical survey responses could still be assigned different latent variable values according to the model.

Instead, we assume measurement equivalence, ensuring that the scales remain consistent across countries. This assumption preserves interpretability, as the qualitative meaning of the latent variables aligns with theoretical expectations—for example, more cooperative responses to the three indicators of *willingness to cooperate with the police* correspond to higher values of that latent variable. Quantitatively, the weights assigned to different survey items remain the same across all countries, as they are defined in the first step of our two-step estimation process, where measurement models are estimated jointly for all countries.

Another important type of comparison involves examining the relative strength of different predictors both within and between countries. For example, we may wish to compare the influence of *instrumental motives*—such as perceptions of police effectiveness in crime reduction and personal concerns about victimization risk—against *normative motives*, including procedural justice, normative alignment, and duty to obey. Fully addressing such comparisons is inherently complex, as in any regression analysis, and requires substantive considerations beyond the statistical framework. Our approach is therefore a commonly used, more limited version: we rely on standardized coefficients, which ensure comparability by expressing all variables in similarly defined units of measurement. In our study, standardization is built into the model for all key explanatory variables (excluding background covariates such as age, gender, and education), as they are constructed so that their standard deviations across all countries equal 1. For latent variables, this constraint is imposed when estimating the measurement models, while observed measures—such as perceived lawfulness, distributive justice, and fear of crime—are directly standardized in the same way. Every regression coefficient β of these variables is then in a "fully standardised" form where a difference of 1 unit of standard deviation in this sense in the corresponding explanatory variable is associated with an expected difference of β standard deviation units in the response variable.

For the country means μ_{kj} of individual latent variables, a useful reference value is provided by the overall mean across the countries, which is fixed at 0. For example, a country for which the mean of perceived procedural justice is positive has a higher mean level of this variable than do these countries on average.

REFERENCES

Bakk, Z., & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika, 83*(4), 871-892. https://doi.org/10.1007/s11336-017-9592-7

Kuha, J., & Bakk, Z. (2023). Two-step estimation of latent trait models. *arXiv preprint arXiv: 2303.16101*. https://doi.org/10.48550/arXiv.2303.16101

Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R*. Hoboken, NJ: John Wiley and Sons.

Oberski, D. (2014). lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software, 57*(1), 1-27. https://doi.org/10.18637/jss.v057.i01

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(1), 1-36. 10.18637/jss.v048.i02

Rosseel, Y., & Loh, W. W. (2022). A Structural After Measurement approach to Structural Equation Modeling. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000503

Supplement S4. Tables of results from the estimated models
Note: The estimation code and detailed outputs from the models are shown in separate Supplement S8. They include the estimates for those parameters and standard errors that are not included in the tables in this Supplement S4.

Table S4a. Measurement loadings in estimated factor analysis measurement models

Construct	Item 1	Item 2	Item 3
Procedural justice	.579	.573	.533
Effectiveness	1.737	1.821	1.428
Obligation to obey	2.156	2.728	2.495
Normative alignment	.727	.789	.685
Cooperation	.482	.813	.771

Note: See Supplement S1 for the wordings and response options of the measurement items for each construct

Table S4b. Sample sizes and R^2 statistics in the structural equation models (SEMs).

Country	*n*	Normative alignment	Duty to obey	Cooperation
Austria	2,256	.45	.32	.16
Belgium	1,699	.50	.14	.06
Bulgaria	2,430	.46	.18	.11
Croatia	1,608	.52	.17	.06
Cyprus	1,068	.61	.30	.07
Czech Republic	2,384	.49	.06	.06
Denmark	1,573	.44	.14	.09
Estonia	1,792	.44	.12	.10
Finland	1,875	.40	.21	.13
France	1,719	.57	.13	.06
Germany	3,021	.39	.11	.06
Greece	2,714	.66	.26	.03
Hungary	1,560	.50	.13	.07
Ireland	2,537	.62	.26	.14
Israel	2,241	.39	.13	.08
Lithuania	1,621	.45	.16	.03
Netherlands	1,823	.55	.12	.06
Norway	1,540	.47	.20	.09
Poland	1,746	.45	.13	.04
Portugal	2,149	.33	.09	.03
Russia	2,594	.40	.23	.03
Spain	1,879	.68	.19	.06
Slovakia	1,841	.45	.05	.07
Slovenia	1,381	.48	.08	.05
Sweden	1,487	.59	.15	.08
Switzerland	1,501	.55	.05	.05
Ukraine	1,923	.28	.14	.05
United Kingdom	2,333	.57	.18	.10
South Africa	2,388	.37	.20	.07
US	1,603	.70	.20	.14

Table S4c: Estimated regression coefficients in the structural equation models (SEMs) for normative alignment with the police, separately for each country.

Country	Procedural justice	Distributive justice	Lawfulness	Effectiveness	Fear of crime
Austria	.459	.190	.045	.046	-.060
Belgium	.636	.064	.062	.191	.001
Bulgaria	.481	.079	.029	.185	-.072
Croatia	.612	.040	.021	.145	.010
Cyprus	.534	.033	.122	.293	.100
Czech Republic	.556	.074	-.018	.246	.005
Denmark	.566	.025	.086	.132	-.020
Estonia	.507	.001	.078	.168	.064
Finland	.481	.033	.074	.142	.027
France	.502	.061	.018	.401	.125
Germany	.433	.075	.023	.198	.015
Greece	.543	.063	.082	.247	.022
Hungary	.626	.054	.038	.166	.032
Ireland	.601	.047	.054	.221	.018
Israel	.385	.036	.212	.210	.062
Lithuania	.561	.033	-.053	.168	.046
Netherlands	.581	.016	.018	.336	.031
Norway	.557	-.019	.045	.158	.041
Poland	.397	-.016	.042	.246	.078
Portugal	.328	.022	.039	.156	-.034
Russia	.344	.057	.084	.286	-.019
Spain	.764	.004	.065	.137	-.022
Slovakia	.269	.031	.103	.385	-.006
Slovenia	.621	.038	-.026	.132	-.024
Sweden	.635	-.032	.041	.149	-.018
Switzerland	.567	.050	.090	.179	-.007
Ukraine	.465	.073	-.163	.252	.054
United Kingdom	.577	.078	.049	.196	.016
South Africa	.236	.138	-.029	.378	.021
US	.494	.044	.098	.248	.024

Table S4d: Estimated regression coefficients in the structural equation models (SEMs) for duty to obey the police, separately for each country.

Country	Procedural justice	Distributive justice	Lawfulness	Effectiveness	Fear of crime
Austria	.317	-.039	.033	.258	.050
Belgium	.313	.012	.038	.124	.026
Bulgaria	.288	.015	.013	.196	-.035
Croatia	.305	-.103	.191	.153	-.094
Cyprus	.358	-.106	.041	.157	.011
Czech Republic	.218	.023	-.030	.065	.007
Denmark	.303	.042	.039	.087	-.008
Estonia	.243	-.063	.130	.211	.007
Finland	.267	.020	.054	.147	.030
France	.202	.034	-.003	.161	.038
Germany	.260	.045	.038	.149	.072
Greece	.271	.041	.032	.145	.066
Hungary	.292	.010	.056	.104	.007
Ireland	.368	-.004	.112	.087	-.039
Israel	.299	.120	-.053	-.023	.004
Lithuania	.234	.084	.044	.189	.079
Netherlands	.229	.010	.009	.188	.026
Norway	.420	-.018	.026	.090	.055
Poland	.175	.053	.101	.141	.120
Portugal	.134	.035	.049	.077	-.029
Russia	.221	.059	.055	.203	-.012
Spain	.288	-.025	.068	.136	.028
Slovakia	.112	.011	.010	.152	.044
Slovenia	.444	-.093	-.039	.026	-.107
Sweden	.393	-.033	.052	.078	-.001
Switzerland	.178	-.012	.079	.044	.027
Ukraine	.258	.139	-.082	.101	-.074
United Kingdom	.216	.023	.054	.187	.050
South Africa	.046	.008	-.024	.300	.006
US	.274	-.009	.037	-.012	-.002

Table S4e: Estimated regression coefficients in the structural equation models (SEMs) for cooperation with the police, separately for each country.

Country	Procedural justice	Effectiveness	Fear of crime	Duty to obey	Normative alignment
Austria	.295	-.217	-.127	.035	.107
Belgium	.091	.021	-.006	.007	.049
Bulgaria	.140	-.010	-.001	.031	.110
Croatia	.095	.077	-.088	-.021	.003
Cyprus	.289	-.135	-.050	-.040	-.091
Czech Republic	.121	.014	-.033	.020	.090
Denmark	.094	-.086	-.076	.118	.054
Estonia	.255	-.088	-.027	.027	.148
Finland	.087	-.071	-.068	.150	.163
France	.133	-.097	-.006	.037	.094
Germany	.102	.026	-.016	-.001	.074
Greece	.043	.105	.015	-.031	.027
Hungary	.067	-.024	-.003	.112	.143
Ireland	.134	-.011	-.082	-.001	.151
Israel	.081	-.001	.101	.159	.009
Lithuania	-.093	.162	-.035	.037	.094
Netherlands	.127	-.032	.040	.092	.016
Norway	.088	-.065	-.096	.021	.166
Poland	.094	-.061	.048	.003	.092
Portugal	.016	-.048	.004	.027	.040
Russia	.071	.073	.009	.025	.030
Spain	.204	-.062	.017	.067	-.045
Slovakia	.289	-.017	-.005	.019	-.071
Slovenia	-.016	.129	-.010	-.066	-.027
Sweden	.180	.061	-.138	.064	.057
Switzerland	.132	-.062	-.056	-.002	.071
Ukraine	.107	-.050	.052	.004	.072
United Kingdom	.212	-.075	-.050	.115	-.030
South Africa	.210	-.046	.013	.013	.100
US	.146	-.186	-.017	.010	.208

Table S4f: Estimated conditional (residual) correlations between duty to obey the police and normative alignment with the police, obtained from the structural equation models (SEMs) separately for each country.

Country	Residual correlation
Austria	.200
Belgium	.155
Bulgaria	.238
Croatia	.240
Cyprus	.081
Czech Republic	.194
Denmark	.206
Estonia	.029
Finland	.300
France	.189
Germany	.207
Greece	.256
Hungary	.176
Ireland	.215
Israel	.117
Lithuania	.198
Netherlands	.149
Norway	.319
Poland	.085
Portugal	.339
Russia	.245
Spain	.099
Slovakia	.122
Slovenia	.236
Sweden	.244
Switzerland	-.022
Ukraine	.169
United Kingdom	.230
South Africa	.045
US	.332

Supplement S5. Adding police-initiated contact to the models

We examine whether (and how) accounting for police contact influences the conclusions drawn from our models of interest—specifically, the regression coefficients for perceived normative alignment, perceived duty to obey and cooperation with the police, as reported in Tables S4c–S4e of the paper. Given the large number of coefficients across different variables and countries, we present these comparisons in graphical form for clarity, as explained below. First, we adjust for police-initiated contact—the only available measure of personal experience with the police—to assess whether it alters key findings from the SEMs. Second, we conduct a split-group analysis to compare individuals who have had recent contact with the police to those who have not, examining whether the SEM results differ between these groups.

Controlling for contact with the police in the models

We incorporate information about police contact as additional control variables across all components of the structural model—specifically, in the regression models for all eight variables presented in Figure 1 of the paper. The key question is whether this adjustment affects the conclusions regarding the relationships between the other variables in the models. This could occur if police contact and satisfaction with such contact are strongly associated with key factors such as perceptions of normative alignment, obligation to obey and cooperation with the police.

These models include both the variables *nocontact* and *contactsat*, with *contactsat* now coded as 0 also for those for whom *nocontact* is 1. The contribution of these variables to the linear predictor of each model is of the form $b_0 + b_nocontact*nocontact + b_contactsat*contactsat$, where *b0* is the intercept term and *b_nocontact* and *b_contactsat* are coefficients of *nocontact* and *contactsat* respectively, for a given response variable. The contribution from this to the expected value of the response variable is then $b_0 + b_nocontact$ for those who have not had contact with the police, and $b_0 + b_contactsat*contactsat$ for those who have.

Values of the contact variables (*nocontact* and/or *contactsat* for those who have had contact) are missing for some respondents. Because of this, some of the respondents who were included in estimating the models in the paper are excluded here when the contact variables are added. This reduction is small, however: an additional 285 respondents, or around 0.5% of the samples that were used for the main analysis.

Comparisons of results from models with and without the contact variables are shown in Figures S5a-5c below. Each plot shows coefficients of one explanatory variable for one response variable, in the same order as in Figures 2-4 of the paper respectively. Each point in the plots shows the value of a coefficient for one country estimated from the model without the contact variables (i.e. the estimates reported in the paper; horizontal axis of the plots) against its estimated value from the model with the contact variables included (vertical axis).

The clear conclusion from these plots is that all the estimated coefficients are very similar whether or not the contact variables are included. A similar conclusion holds for the estimated standard errors (which are not shown here): on average over the 450 coefficients shown in these plots, the estimated standard errors are 0.4% higher from models which control for the contact variables than from ones which do not. Together, these results mean that the results shown in Figures 2-4 of the paper, and the conclusions drawn from them, would be essentially unchanged if measures of the respondents' contact with the police were included in the models as additional control variables.

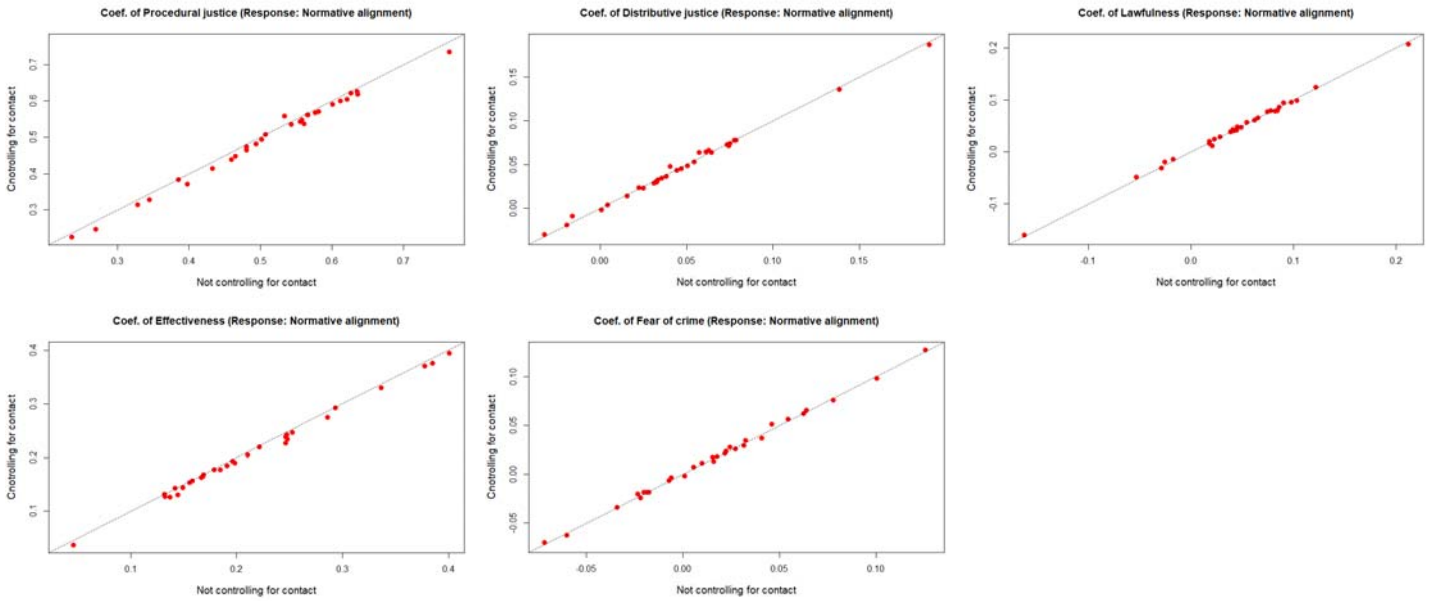

FIGURE S5a. Regression coefficients predicting normative alignment with the police in the 30 countries included in the analysis, from models which do not include predictor variables on contact with the police (horizontal axis; i.e. these are the coefficients that are shown also in Figure 2 of the paper) vs. models which do include predictor variables on contact with the police (vertical axis).

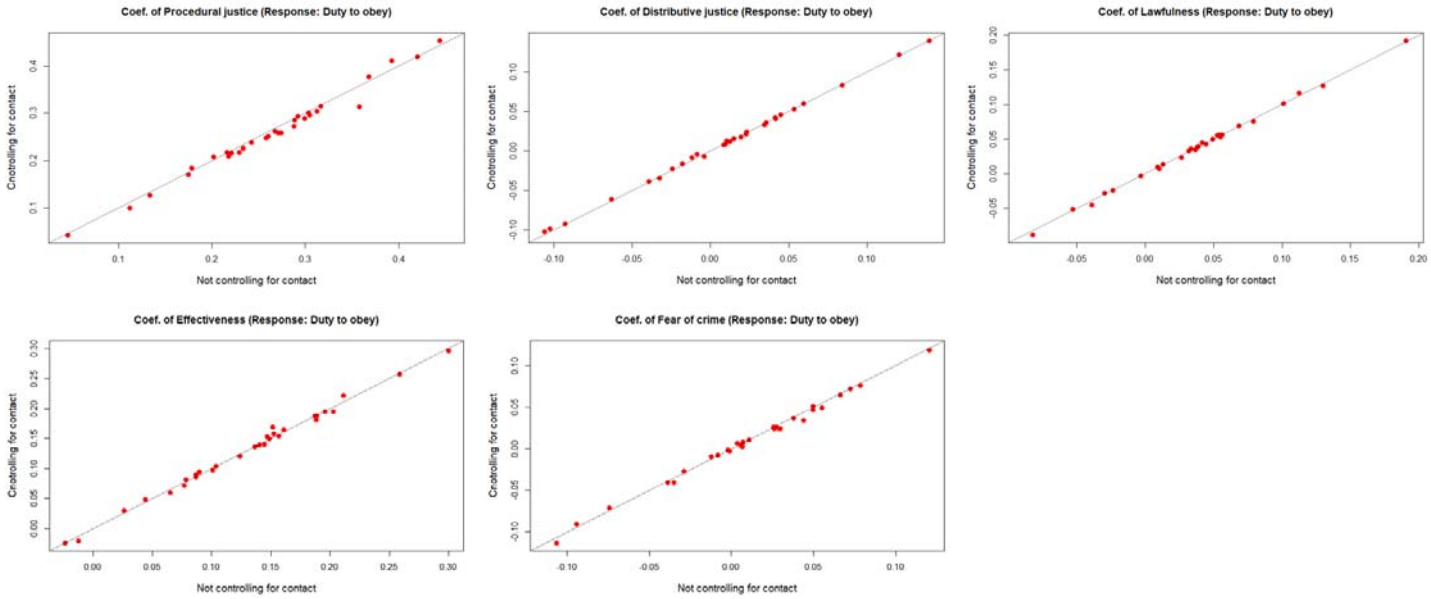

FIGURE S5b. Regression coefficients predicting duty to obey the police in the 30 countries included in the analysis, from models which do not include predictor variables on contact with the police (horizontal axis; i.e. these are the coefficients that are shown also in Figure 3 of the paper) vs. models which do include predictor variables on contact with the police (vertical axis).

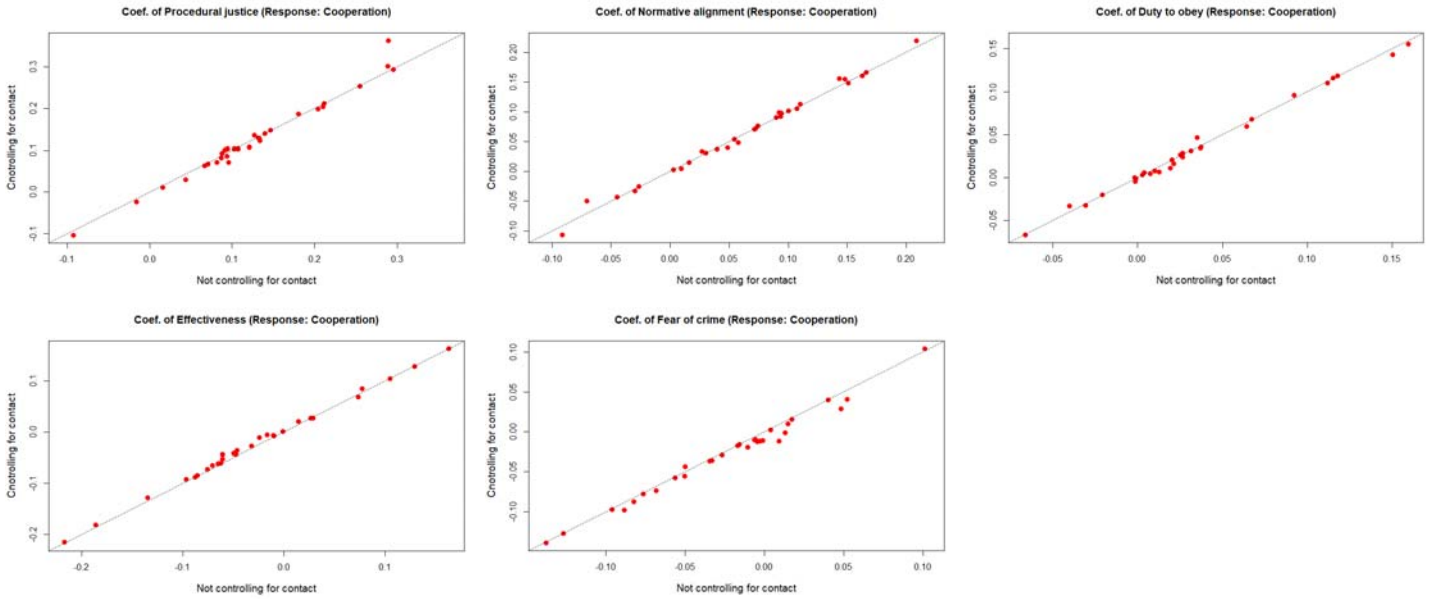

FIGURE S5c. Regression coefficients predicting cooperation with the police in the 30 countries included in the analysis, from models which do not include predictor variables on contact with the police (horizontal axis; i.e. these are the coefficients that are shown also in Figure 4 of the paper) vs. models which do include predictor variables on contact with the police (vertical axis).

Fitting the models separately for those who have had contact with the police and those who have not

We divide the data within each country into two groups: respondents who have had contact with the police in the last two years (*nocontact = 0*) and those who have not (*nocontact = 1*). We then estimate the models separately for each group to assess whether and how the coefficients differ in ways that might influence our conclusions. Differences between these groups could arise if the survey questions used in the analysis have varying salience or if respondents interpret them differently based on their personal experience with the police. Specifically, individuals with direct police contact may draw on firsthand experiences when answering, while those without such encounters may rely on broader societal perceptions. Identifying these differences helps evaluate whether and to what extent police contact shapes responses in ways that impact our findings.

Note that there is a reasonable amount of data for both groups in all the countries. The smallest sample size is 359, for respondents in Lithuania who have had contact with the police.

Our primary focus remains on the models for normative alignment, duty to obey and cooperation with the police. We first compare estimated models in which all coefficients for these variables are constrained to be the same across contact and non-contact groups with models where these coefficients are allowed to differ. In both cases, the coefficients for other key variables—procedural justice, distributive justice, lawfulness, effectiveness, and worry about crime—are permitted to vary freely between the groups.

Likelihood ratio tests conducted separately for each country indicate that these models differ significantly (at the 5% level) in 24 out of the 30 countries. This suggests that at least some parts of the models vary depending on whether respondents have had direct contact with the police. However, while these differences are statistically significant, they are not necessarily substantively meaningful for our analysis. The more important question is whether these differences follow a systematic pattern that could influence our main conclusions.

For instance, the results would be of greater concern if certain relationships—such as the link between procedural justice and cooperation—were consistently stronger for those with prior police contact, at least in some countries. This possibility is explored in Figures S5d–S5f. Each plot visualizes the coefficients for one explanatory variable in relation to one response variable, maintaining the same order as Figures 2–4 in the main paper. Each point in the plots represents the coefficient for a given country, with the horizontal axis showing estimates for respondents who have not had contact with the police and the vertical axis showing estimates for those who have had such contact. The points are drawn as blue circles when the absolute value of the coefficient is smaller for the contact group and as red triangles when it is larger.

The main message from these plots is the qualitative impression that red and blue markers are evenly distributed: approximately half of the estimated coefficients are further from zero for those who have had police contact, while the other half are further from zero for those who have not. Specifically, this pattern holds for 235 of the 450 coefficients (52.2%) in the contact group and 215 coefficients (47.8%) in the non-contact group.

While this even split suggests no systematic pattern overall, it is still possible that country-level variations exist—for example, that the associations might be consistently stronger for the contact group in some countries but systematically weaker in others. However, this is not the case. When we fit a binary logistic random intercepts (variance components) model to these data—where the response variable indicates whether a coefficient is larger in absolute value for the contact group—we find an estimated between-country variance of essentially zero. Additionally, a binomial test (assuming a probability of 0.5 for the pooled observations, i.e., 235 out of 450) yields $p = 0.370$, indicating that the distribution is statistically indistinguishable from a random 50/50 split.

Moreover, the estimated coefficients for the two groups closely align with those used in the main paper, which were obtained by pooling data from all respondents. The average absolute difference between these pooled estimates and the group-specific estimates is 0.035, and in only 5% of cases does this difference exceed 0.1—the same spacing as the axis ticks in Figures 2–4. In other words, if we were to replace the pooled estimates in these figures with those from either the contact or non-contact group separately, the shapes of the plots and the conclusions drawn from them would remain essentially unchanged.

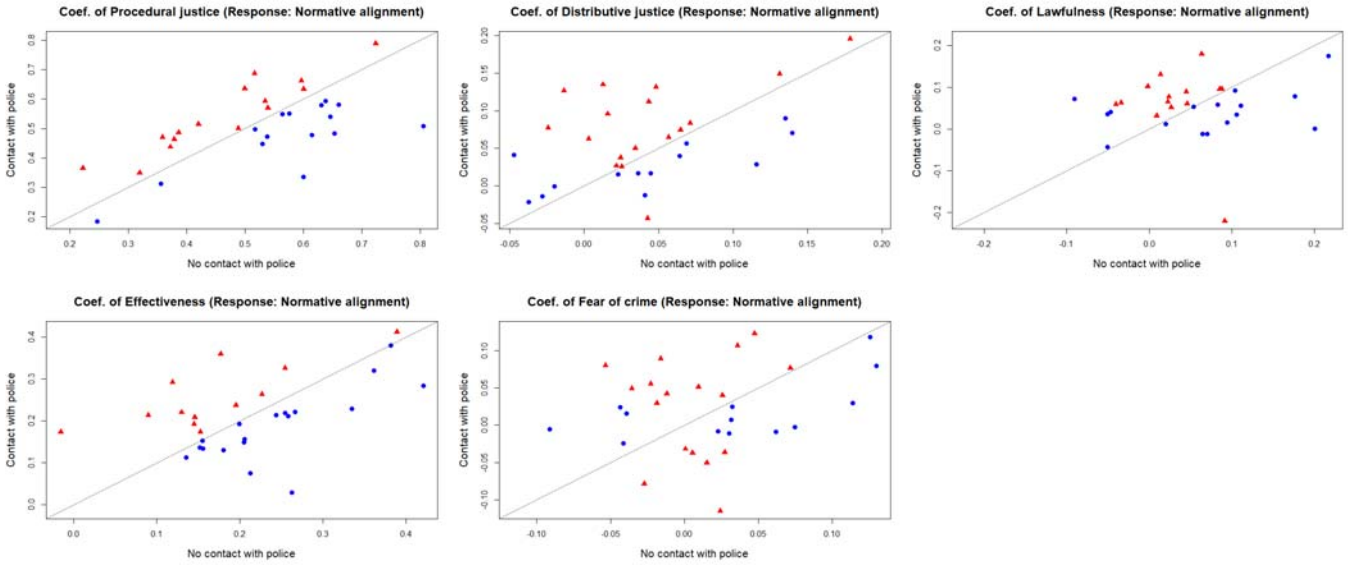

FIGURE S5d. Regression coefficients predicting normative alignment with the police in the 30 countries included in the analysis, from models fitted to data from respondents who have not had contact with the police in the last two years (horizontal axis) vs. data from respondents who have had such contact (vertical axis). The points are shown as blue circles when the estimate for the non-contact group is larger in absolute value, and as red triangles when the estimate for the contact group is larger in absolute value.

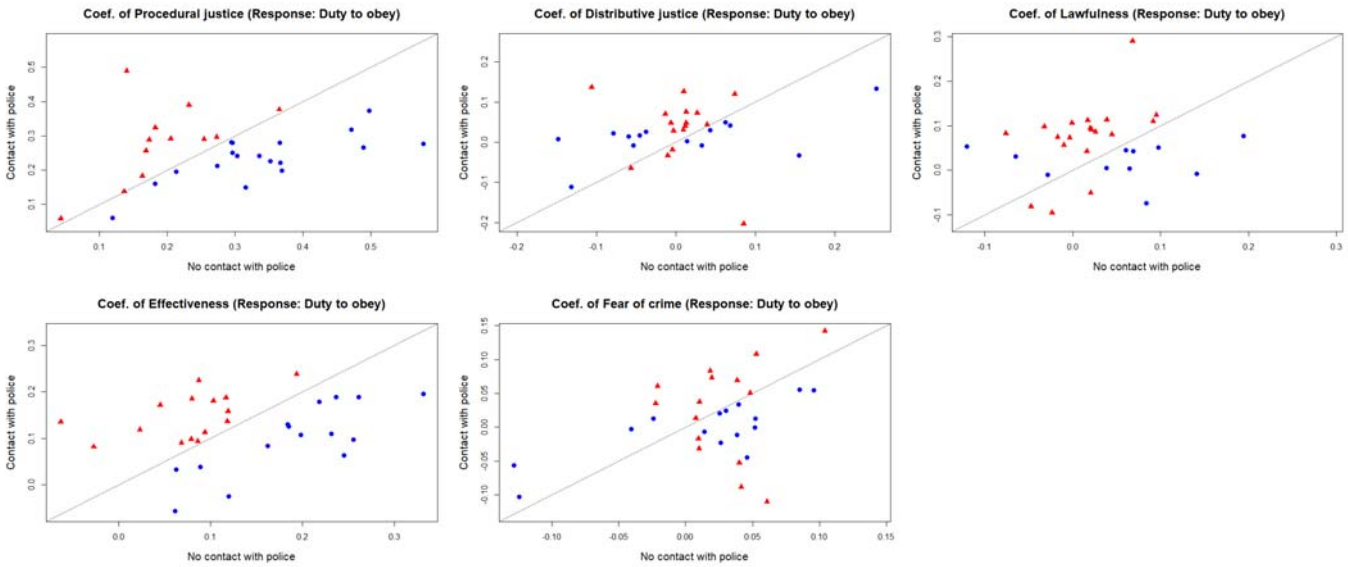

FIGURE S5e. Regression coefficients predicting duty to obey the police in the 30 countries included in the analysis, from models fitted to data from respondents who have not had contact with the police in the last two years (horizontal axis) vs. data from respondents who have had such contact (vertical axis). The points are shown as blue circles when the estimate for the non-contact group is larger in absolute value, and as red triangles when the estimate for the contact group is larger in absolute value.

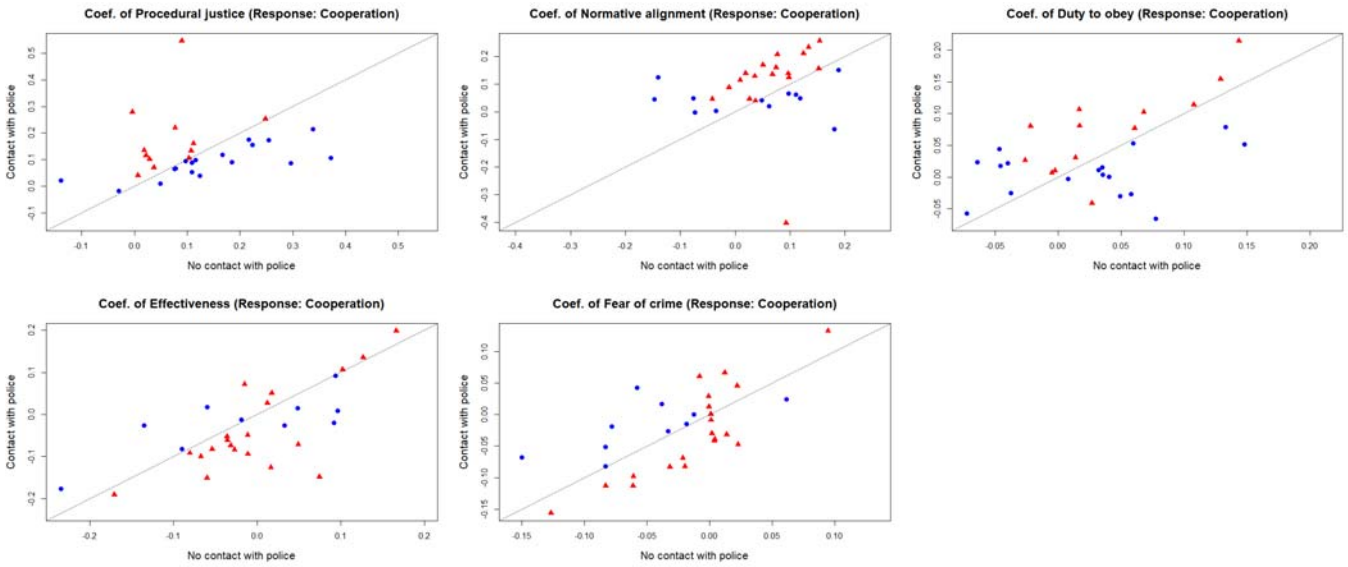

FIGURE S5f. Regression coefficients predicting cooperation with the police in the 30 countries included in the analysis, from models fitted to data from respondents who have not had contact with the police in the last two years (horizontal axis) vs. data from respondents who have had such contact (vertical axis). The points are shown as blue circles when the estimate for the non-contact group is larger in absolute value, and as red triangles when the estimate for the contact group is larger in absolute value.

Supplement S6. Further comments on police effectiveness as a predictor of cooperation

One aspect of our main results that may seem counterintuitive at first glance is that the coefficients for perceived police effectiveness as a predictor of cooperation with the police (as shown in Figure 4 of the paper) are negative (albeit mostly not statistically significant) in most countries. In other words, respondents who perceive the police as highly effective tend to report lower levels of cooperativeness, controlling for other predictor variables in the model. However, we caution against placing too much emphasis on this observation, given that effectiveness is not statistically significant (at the 5% level) in most countries. Nevertheless, since the majority of point estimates are negative, some commentary is warranted.

When we investigated the issue further, we found that the association between perceived effectiveness and cooperation turns negative only when controlling for other variables in the model—particularly perceived procedural justice. Figure S6 presents the effectiveness coefficients across the 30 countries, estimated under different model specifications that include varying sets of explanatory variables. The horizontal axis of each plot shows the estimates from the full model reported in the paper (as shown in Figure 4), which includes all explanatory variables. The vertical axis represents estimates from alternative models that retain demographic controls (age, gender and education) but omit some or all of the following: worry about crime, perceived procedural justice, normative alignment and obligation to obey the police.

- In the top-left plot, all of these variables are omitted, leaving only demographic controls and police effectiveness in the model. Here, the effectiveness coefficient is positive in all but one country, and almost all estimates are statistically significant (though standard errors are not shown).
- The top-right plot includes worry about crime and obligation to obey the police alongside demographic controls. The coefficients remain mostly positive.
- The bottom-left plot adds normative alignment as a control, leading to a shift in some estimates.
- The bottom-right plot, which controls for procedural justice (alongside demographic factors), most closely mirrors the pattern seen in the full model—where effectiveness is negative in most cases.

This pattern suggests that, among individuals with the same level of perceived procedural justice, those who view the police as more effective tend to be less cooperative with the police in many countries. Numerically, this result follows from the fact that procedural justice is both positively associated with willingness to cooperate (as seen in Figure 4) and positively correlated with perceived effectiveness (with correlations ranging from .4 to .8 across countries).

From a substantive perspective, the interpretation of these (mostly non-significant) partial associations between perceived effectiveness and cooperation remains unclear. Given the general lack of statistical significance and clear theoretical explanation, we refrain from offering speculative interpretations.

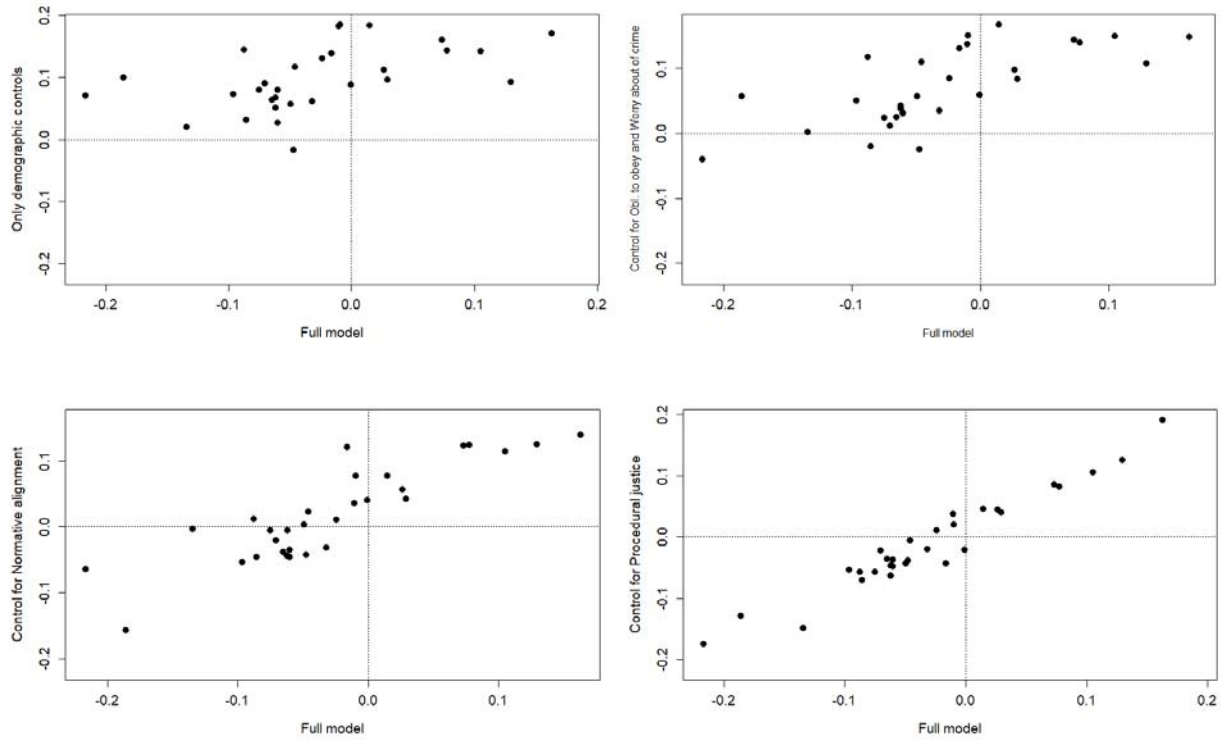

FIGURE S6. Regression coefficients of perceived police effectiveness in models for cooperation with the police in the 30 countries included in the analysis, from different models. In each plot the coefficients on the horizontal axis are from the full model used in the paper (i.e. the coefficients that are also shown also in Figure 4 of the paper). The coefficients on the vertical axis are from models which include different control variables (in addition to age, sex and education): none of them (top left), obligation to obey and worry about crime (top right), normative alignment (bottom left), or procedural justice (bottom right).

Supplement S7. Supplementary models for South Africa

Bradford et al. (2014) used data from the 2010 SASAS to examine whether procedural justice theory applies in post-Apartheid South Africa. To test whether procedural justice theory required adaptation to South Africa's unique context, Bradford et al. (2014) introduced four additional factors: perceived group threat, anti-immigrant sentiment, trust in government and satisfaction with service provision. Despite South Africa's high crime rates, deep social divisions and historically strained police-community relations, their findings indicated that procedural justice was a strong predictor of police legitimacy. However, unlike in wealthier, more stable democracies—where effectiveness is often a secondary concern—South Africans placed greater emphasis on police effectiveness in combating crime. Concerns about crime levels and state legitimacy also played significant roles, emphasizing the fragile social utility of public policing. Their results suggest that while procedural fairness remains central to police legitimacy, legitimacy is also deeply intertwined with broader socio-political conditions, particularly public confidence in the state's ability to govern effectively.

To assess whether the inclusion of perceived group threat, anti-immigrant sentiment, trust in government and satisfaction with service provision altered our main findings for South Africa, we conducted a supplementary analysis using the 2010 SASAS data (these measures were unavailable in the 2012 SASAS data). We fitted a model using only South African data, incorporating these additional control variables that were not used in our main cross-national analysis. This serves as an illustrative example of how country-specific variables—important in a particular national context—may not always be relevant or available for broader comparative cross-national analysis. The key focus here is whether including these additional variables modifies our main conclusions about procedural justice and legitimacy in South Africa.

Each of the four additional constructs—group threat, anti-immigrant sentiment, trust in government, and satisfaction with service provision—is measured using multiple indicators. For two-step estimation, measurement models were first fitted for each of the four additional constructs, and measurement parameters from this step were fixed in the second step, where structural models were estimated. Measurement models for the other latent constructs were fixed at the same values obtained from the 2012 SASAS data, ensuring consistency with our main analysis. To facilitate interpretation, the new variables were defined such that higher values indicate more *benign views*:

- Lower perceived group threat
- More pro-immigrant sentiment
- Higher trust in government
- Higher satisfaction with service provision

This approach allows us to examine whether controlling for these additional variables affects the previously reported conclusions for South Africa and whether their inclusion offers further insights into the relationship between procedural justice, police legitimacy and willingness to cooperate in the South African context.

We focus again on models for normative alignment, duty to obey and willingness to cooperate, and examine the estimated coefficients of their predictors, now including in one model also the four additional variables. Estimates of these coefficients are shown in Table S7, for three model fits. The first two of them are for the model considered in the paper, fitted to the 2012 data (as in the paper) and to the 2010 data. The third model adds the four additional variables. They are included in the same way as the variables in the left-hand box in Figure 1 of the paper, i.e. as predictors of some or all of normative alignment, duty to obey and cooperation (and themselves predicted by gender, age and education).

Comparing first the main model fitted in 2012 and 2010, the results are broadly similar. For example, the pattern that police effectiveness is a stronger predictor of normative alignment and duty to obey than is procedural justice, which is peculiar to South Africa, appears in both years. One difference appears in the model for cooperation, where in 2012 normative alignment is a significant predictor while duty to obey is not, but the opposite is the case in 2010.

The main comparison of interest in this supplementary analysis is between models for the 2010 data which control for the four additional variables. The conclusions are not affected by this, in that the magnitudes and levels of significance of the explanatory variables of interest remain essentially unchanged when we control for the additional variables. We refrain from speculating on a post-hoc explanation for this.

TABLE S7. Regression coefficients predicting normative alignment with the police, duty to obey the police and cooperation with the police in South Africa. Three models are shown: the model considered in the paper, fitted to data from 2012 (as in the paper), the same model fitted to data from 2010, and a model with four additional explanatory variables fitted to 2010 data.

Explanatory variable	Models reported in the paper, fitted to 2012 data	Same models as in the paper, but fitted to 2010 data	Models with four additional explanatory variables, fitted to 2010 data
Response variable: Normative alignment with the police			
Procedural justice	.236***	.246***	.245***
Distributive justice	.138***	.114***	.111***
Lawfulness	-.029	-.007	-.009
Effectiveness	.378***	.311***	.285***
Fear of crime	.021	-.021	-.031*
Group threat			-.037
Anti-immigrant sentiment			.026
Trust in government			.120***
Provision of services			.020
Response variable: Duty to obey the police			
Procedural justice	.046*	.042*	.038
Distributive justice	.008	.048*	.045*
Lawfulness	-.024	-.054*	-.046*
Effectiveness	.300***	.219***	.221***
Fear of crime	.006	-.021	-.021
Group threat			-.051*
Anti-immigrant sentiment			-.039*
Trust in government			.001
Provision of services			.087***
Response variable: Cooperation with the police			
Procedural justice	.210***	.226***	.222***
Normative alignment	.100***	-.013	-.005
Duty to obey	.013	.119***	.116***
Effectiveness	-.046	.051*	.062*
Fear of crime	.013	-.026	-.021
Trust in government			-.069**
Provision of services			.032

p-values: *** <.001; ** <.01; * <.05

Each model includes also respondent's age, sex and level of education as explanatory variables.

The four additional concepts and their survey measures in the SASAS 2010:

Group threat: taps into fears that other racial groups are trying to get ahead at the expense of the respondent's own, by, for example, excluding members of their group from positions of power and responsibility or by undermining their traditions and values.

- Q73: People of other race groups in South Africa are trying to get ahead economically at the expense of my group
- Q74: People of other race groups in South Africa try to exclude members of my group from positions of power and responsibility
- Q75: The traditions and values that are important to people of my race are under threat
- Q76: Other race groups in South Africa will never understand what members of my group are like

(response options: five-point Strongly agree – Strongly disagree scale)

Anti-immigrant sentiment: assessments of, on the one hand, the 'threat' immigrants pose to South African society, and, on the other, the contribution immigrants might make.

- Q92: Immigrants increase crime rates
- Q93: Immigrants are generally good for South Africa's economy
- Q94: Immigrants take jobs away from people who were born in South Africa
- Q95: Immigrants make South Africa more open to new ideas and cultures
- Q96: Immigrants bring disease to South Africa
- Q97: Immigrants bring skills that are needed in South Africa

(response options: five-point Strongly agree – Strongly disagree scale)

Trust in government: the extent individuals trust different institutions of government and politics.
Indicate the extent to which you trust or distrust the following institutions in South Africa at present:

- Q7: National government
- Q11: Parliament
- Q14: Local government
- Q17: Political parties
- Q18: Politicians

(response options: five-point Strongly trust – Strongly distrust scale)

Provision of services: satisfaction with basic services (water, electricity, refuse collection).
How satisfied are you with the way that the government is handling the following matters in your neighbourhood?

- Q20: Supply of water and sanitation
- Q21: Providing electricity
- Q22: Removal of refuse
- Q24: Access to health care

(response options: five-point Very satisfied – Very dissatisfied scale)

REFERENCES

Bradford, B., Huq, A., Jackson, J., & Roberts, B. (2014). What price fairness when security is at stake? Police legitimacy in South Africa. *Regulation and Governance, 8*(2), 246–268. https://doi.org/10.1111/rego.12012