



# “Truly free consent”? Clarifying the nature of police legitimacy using causal mediation analysis

Krisztián Pósch<sup>1</sup> · Jonathan Jackson<sup>2</sup> · Ben Bradford<sup>3</sup> · Sarah Macqueen<sup>4</sup>

Published online: 7 May 2020  
© The Author(s) 2020

## Abstract

**Objectives** To test whether normative and non-normative forms of obligation to obey the police are empirically distinct and to assess whether they exhibit different dynamics in terms of the downstream effects of police-citizen contact.

**Methods** Analysing data from the Scottish Community Engagement Trial of procedurally just policing, we use natural effect modelling for causally ordered mediators to assess causal pathways that include—but also extend beyond—the experimental treatment to procedural justice.

**Results** Normative and non-normative forms of obligation are empirically distinct. Normative obligation to obey the police is sensitive to procedurally just or unjust police behaviour, and influences cooperation with the police and traffic law compliance in a way that is consistent with procedural justice theory. Non-normative obligation to obey the police is ‘sticky’ and unresponsive.

**Conclusions** Legitimacy can reasonably be defined partly as normative obligation with its expected beneficial downstream effects, so long as it is measured properly.

**Keywords** Causal mediation analysis · Cooperation · Measurement · Obligation to obey the police · Police legitimacy · Procedural justice · Sense of power

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11292-020-09426-x>) contains supplementary material, which is available to authorized users.

---

✉ Jonathan Jackson  
j.p.jackson@lse.ac.uk

<sup>1</sup> Department of Methodology, London School of Economics, London, UK

<sup>2</sup> Department of Methodology, London School of Economics and University of Sydney Law School, Houghton Street, London WC2A 2AE, UK

<sup>3</sup> Department of Security and Crime Science, University College London, London, UK

<sup>4</sup> Scottish Government, Edinburgh, UK

## Introduction

‘Legitimacy is a psychological property of an authority, institution, or social arrangement that leads those connected to it to believe that it is appropriate, proper, and just. Because of legitimacy, people feel that they ought to defer to decisions and rules, following them voluntarily out of obligation rather than out of fear of punishment or anticipation of reward.’ (Tyler 2006a: 375).

As the right to power and the authority to govern, legitimacy is central to crime-control. On the one hand, legitimacy reduces the tension between power-holders and subordinates (Tyler and Jackson 2013, 2014). When people view the police as appropriate, proper and just—i.e. they acknowledge that the institution has a valid claim to power—they feel a corresponding obligation to obey officers (Tyler 2006a, b; Trinkner 2019). They are also more likely to comply with the law (Murphy et al. 2016; Sunshine and Tyler 2003; Trinkner et al. 2018) and cooperate with the police and criminal courts (Reisig and Lloyd 2008; Wolfe et al. 2016; Hagan et al. 2018; cf. Jackson et al. 2020). On the other hand, legitimacy constrains power in normatively appropriate ways. To be seen as legitimate, authority figures need to treat individuals with respect and dignity, make decisions in open, neutral and accountable ways and respect the limits of their rightful authority (Bradford et al. 2014a; Jonathan-Zamir and Harpaz 2018; Murphy and Cherney 2012; Trinkner et al. 2018). On this account, legitimacy forms part of a virtuous circle. By tilting the authority-citizen relationship from coercive to consensual, legitimacy reduces the need for costly and minimally effective forms of crime-control, opening up further space for policing strategies that prioritise consent over coercion (President’s Task Force on twenty-first Century Policing, 2015; Tyler et al. 2015a).

But is this portrayal of police legitimacy and power relations overly optimistic? A central proposition of procedural justice theory is that people feel a normatively grounded obligation to obey the rules and orders that emanate from a police institution that—in their eyes—wields its power in normatively appropriate ways (Tyler and Trinkner 2017). Measures of police legitimacy try to tap into the sense of felt obligation to obey the commands of officers and accept their decisions. To this end, research participants are asked to agree or disagree with attitudinal statements such as “you should accept police decisions because it is the right or proper thing to do” and “you should obey the orders of police officers even if you disagree with them”. The assumption underpinning this operationalization strategy is that positive answers to questions like these express the belief that the police are morally entitled to dictate appropriate behaviour—people feel an obligation to obey because it is the right thing to do, not because they fear punishment nor feel powerless to do otherwise (Jackson 2018; Jackson and Gau 2016; Tyler and Jackson 2013).

Yet, Tankebe and colleagues (Bottoms and Tankebe 2012; Tankebe 2009: 1279–1281; Tankebe 2013: 105–106; see also Johnson et al. 2014: 970) have recently argued that criminologists have failed to properly disentangle normative (consensual) and non-normative (prudential and/or instrumental) forms of felt obligation to obey the police. Research participants could report feeling an obligation to obey the police, not because they believe it is their duty to obey what they believe to be a normatively appropriate

institution (*it's my freely-chosen duty as a citizen to allow officers to dictate appropriate behaviour*) but because of pragmatism (*it's not worth risking non-compliance*) and dull compulsion (*it's not my place to question the orders of police*). It could even be that someone who experiences their relationship to the police as a 'power relationship, pure and simple, with no element of right' (Bottoms and Tankebe 2012: 126) could report feeling obligated to obey the police.

The argument that obligation to obey should be seen as downstream of legitimacy rather than a constituent part of legitimacy has proved to be influential. In a China-based study, Sun et al. (2018) specified obligation to obey as an outcome rather than indicator of police legitimacy, arguing that '...people in an authoritarian society may report an obligation to obey the law/police corresponding to a dull compulsion to do so' (Sun et al. 2018: 2). Analysing data from the Zurich Project on Social Development from Childhood to Adulthood, Nivette et al. (2020: 9) defined and measured police legitimacy as something other than obligation to obey because according to Tankebe (2013), "the concept of obligation is a much broader concept than legitimacy, which can be influenced in part by normative judgments, instrumental concerns (e.g., fear), as well as more fatalistic or pragmatic acquiescence". In the second of two US-based studies into officer perceptions of audience legitimacy, Nix et al. (2019: 4) bracketed off obligation to obey from the legitimacy construct, arguing "...while perceived legitimacy undoubtedly causes some people to feel obligated to obey authorities, others might obey due to fear of the consequences of disobedience, while still others might feel powerless and see no realistic alternative to obedience (i.e. "dull compulsion", see Carrabine 2004:180)".

In this paper, we respond to Bottoms and Tankebe's (2012) call for the disentanglement of motives in people's expressed obligation to obey the police. As they argue in relation to the law (although they also direct their point to the obligation to obey the police):

...there are several reasons other than true legitimacy why people might express feelings of obligation to obey the law: these include structurally-generated apathy and pragmatic acquiescence (dull compulsion) and instrumental calculations. To measure true legitimacy, these alternative motives need to be disentangled; however, most existing studies have not paid sufficient attention to the need for this disentanglement. (p. 165)

If it is true that we, as a community of criminologists, cannot properly distinguish between normative and non-normative forms of obligation to obey police commands, then we should reconsider (a) the conceptualization and measurement of police legitimacy, (b) the extant evidence base and (c) the policy prescription that flows from this body of research. Conversely, if normatively grounded obligation (properly measured) can be parsed from non-normative obligation to obey, and if normative obligation (properly measured) shows qualities consistent with *truly free consent*, then future studies could reasonably define and measure legitimacy partly as the conferred authority to dictate appropriate behaviour, so long as appropriate indicators are used.

By way of contribution, we address the open and empirical question as to whether a normatively grounded obligation to obey the police can be measured distinct from other

motives to obey. In order to draw conclusions from our findings as to whether police legitimacy can partly be defined as a normatively grounded form of obligation to obey the commands of officers, we call upon data from a randomised controlled trial (RCT) of procedurally fair traffic policing (MacQueen and Bradford 2015, 2017). We address two connected research questions:

1. Can normative (consensual) and non-normative (prudential) forms of obligation to obey the police be teased apart empirically?
2. If they can be teased apart empirically, what sort of dynamics do they show when placed at the heart of a procedural justice model of regulatory police-citizen encounters?

We answer the first question by drawing on data from the RCT's survey that fielded measures of normative obligation (e.g. 'I feel a moral obligation to obey the police') and non-normative obligation (e.g. 'I only obey the police because I am afraid of them') to obey the police. We use confirmatory factor analysis to test the empirical distinctiveness of the latent construct(s). To foreshadow the findings, we show that they are empirically distinct.

Yet, this does not necessarily mean that the first form of obligation actually reflects a consensual relationship with the police. The central goal of the second research question is, therefore, to test whether normative obligation (beyond the face validity of questions like 'I feel a moral obligation to obey the police') can reasonably be defined as a constituent component of legitimacy. Specifically, if what is being captured acts like *truly free consent*—i.e. a normatively grounded form of obligation to obey the commands of police officers based on willing consent—then it should (unlike non-normative obligation to obey) respond strongly and positively to behaviour from police officers that were seen to be respectful, trustworthy and unbiased, it should be connected to a sense of power and autonomy with respect to officers, and it should be a positive predictor of future intentions to comply with traffic laws and cooperate with the police.

To shed light on whether we are capturing the dynamics of consensual and coercive police-citizen relations with the two sets of measures, we fit a natural effect model for causally ordered mediators (Steen et al. 2017a; b). On the one hand, we assess whether the measures designed to capture normative obligation to obey the police reflect a latent construct that 'carries' the effects of previous contact with the police onto willingness to cooperate with the police and comply with traffic laws through procedural justice and personal sense of power. On the other hand, we evaluate the relationship between non-normative obligation and the treatment, procedural justice, personal sense of power, normative obligation, cooperation with the police and compliance with traffic laws. Does non-normative obligation to obey the police transmit some of the negative effects of the treatment through procedural justice and personal sense of power to cooperation and compliance?

In addition to finding that normative and non-normative obligation to obey the police are empirically distinct, normative obligation proves to be sensitive to procedurally just or unjust police behaviour in the ways expected by procedural justice theory (PJT). We argue, on this basis, that normative obligation to obey the police—measured using new indicators in the current context of road policing in Scotland—has qualities

consistent with Tyler's (2006a, b) conceptual definition of police legitimacy, i.e. the internalization of the moral value that one should obey the police because that is the right thing to do (Trinkner 2019). Strikingly, non-normative obligation to obey the police seems surprisingly 'sticky' and unresponsive. Non-normative obligation to obey the police appears to operate outside of the PJT; it does not transmit the impact of the contact on either cooperation or legal compliance and it is only weakly and negatively correlated with normative obligation to obey the police (although it has moderately strong negative correlations with procedural justice and sense of power). We conclude with the idea that police legitimacy can reasonably be defined partly as a normatively grounded obligation to obey officers in Scotland. We also call for more research into coercive forms of power relations between police and citizens (cf. Sherman 1993; Tyler et al. 2015; Murphy 2016).

Finally, the paper makes a methodological contribution in the area of causal mediation analysis. We utilise causal mediation analysis to test the dynamics of normative and non-normative obligation, but the challenge inherent in estimating the mechanism through which a causal effect is transmitted is often under-appreciated in criminology. Developments have emerged in other disciplines for testing direct and indirect effects that go beyond to the standard *product method* associated with Baron and Kenny (1986). We illustrate the use of a natural effect model for causally ordered mediators (Steen et al. 2017a; b) that has not yet (to our knowledge) been applied within the discipline. The technique better frames the problem and estimates mediation effects with potentially causal properties. Indeed, we argue that the current RCT is a particularly apposite application of this methodological tool. The RCT suffered from an unusual type of implementation failure (MacQueen & Bradford 2017): there was a treatment effect, but it was in the opposite direction expected, which only increases uncertainty about how the causal effect was transmitted. Building on Posch's (2020) assessment of selection bias and treatment consistency and heterogeneity, we argue that value can be extracted so long as sufficient methodological care is taken. Importantly, we show how that is done.

The paper proceeds as follows. First, we consider the concept of police legitimacy. Second, we motivate the research problem and current empirical goals. Third, we consider the challenge of causal mediation analysis. Fourth, we discuss the RCT's implementation failure. Fifth, we outline the study's design and findings. We close the paper with three main conclusions: (a) that police legitimacy can reasonably be partly defined as obligation to obey the police; (b) that normative and non-normative aspects of obligation to obey the police are unrelated, and that the latter does not seem to mediate the effect of the treatment, procedural justice and personal sense of power on cooperation with the police and compliance with traffic laws; and (c) that causal mediation analysis is a flexible and effective tool that should be more widely employed in criminology when causal mechanisms are under scrutiny.

## What is police legitimacy?

Scholars typically—but not universally—think of police legitimacy as having two constituent parts: (i) the acknowledgement of the rightfulness of power and (ii) the acceptance of the authority of the police to enforce the law and make demands on their

freedom in certain legally prescribed circumstances (Huq et al. 2017; Hamm et al. 2017; Tyler 2006a, b; Tyler and Jackson 2013, 2014; Jackson and Bradford 2019; Trinkner 2019). On this account, the police have legitimacy in the eyes of citizens when those citizens (i) believe that the institution exercises its authority appropriately and (ii) internalise the moral value that they should obey the orders of officers and accept their decisions because of the source not because of the content.

One way of operationalising the first component of legitimacy (the acknowledgement that the police have a valid claim to power) is institutional trust (Tyler 2006a, b). Research participants are asked whether they believe that officers wield their authority in ways that take into account the interests of citizens and society. Two indicative agree/disagree statements are ‘The police can be trusted to make decisions that are right for your community’ and ‘When the police deal with people they almost always behave according to the law’ (see Sunshine and Tyler 2003; Tyler et al. 2010; 2014). The rationale here is that, when people believe police can be trusted to wield their power in ways that encapsulate the interests of citizens, they also believe the institution has the moral right to power. Another way of operationalising the acknowledgement that the police have a valid claim to power is normative alignment, i.e. the sense of shared norms and salient values between the police and the policed (Tyler and Jackson 2014). Respondents are asked whether police officers act in ways that accord with values and norms about how authority should be exercised. Two example indicators are ‘The police usually act in ways that are consistent with my own ideas about what is right and wrong’ and ‘The police stand up for values that are important to you’ (see Jackson et al. 2012; Tyler and Jackson 2014; Tyler et al. 2015b). Positive answers to such attitudinal statements are assumed to reflect the belief that the power of the police is normatively justified because it is being exercised in ways that are consistent with normative expectations in a given political community.

The second part of the police legitimacy concept is the focus of the current paper. When citizens acknowledge the legitimacy of the police, they accept its authority. Echoing the Weberian insight that power is transformed into authority when it enjoys popular legitimacy (Tyler 2003, 2004), obligation to obey the police is assumed to be grounded in the perceived entitlement to dictate appropriate behaviour (Trinkner 2019) and enforce the law (including using reasonable levels of force, see Gerber and Jackson 2017; Bradford et al. 2017). When people believe that an institution is entitled to be obeyed, they treat orders and rules as superseding one’s own judgement (Tyler 2006a, b). Rooted in the normativity that grants officers the right to dictate appropriate behaviour in certain prescribed circumstances, obligation to obey is assumed to be a form of deference that is connected to the rights and responsibility of legal citizenship, not to a sense of powerlessness (a prudential “path of least resistance”) and/or fear of the consequences of non-compliance (Tyler and Jackson 2013).

Legitimacy is then posited to turn coercive modes of crime-control into consensual modes of crime-control through the generation of consent, with legitimacy motivating people to engage in societally desirable behavioural outcomes such as legal compliance or cooperation (Tyler et al. 2015). Importantly, studies tend to find (a) that procedural justice is key to legitimacy and (b) that procedural justice in police-citizen encounters predicts cooperation with the police and compliance with the law largely through citizens’ evaluation of the legitimacy of the authorities (for a review, see Jackson 2018). This means that self-regulation is encouraged through procedural justice—

based approaches to rule enforcement; when people are treated in a procedurally just manner, their propensity to cooperate with the police increases as a result of a strengthened sense of rightfulness, whether this is operationalised as institutional trust or normative alignment (Huq et al. 2011a; Huq et al. 2011b; cf. Jackson et al. 2019). Equally, when people are treated in a procedurally fair way, their propensity to comply with the law increases through a strengthened sense of obligation to obey the police and the law (Murphy et al. 2009; Slocum et al. 2016; Trinkner et al. 2018; Sampson and Bartusch 1998; Kirk and Papachristos 2011).

## Conceptual and theoretical contribution

That is the theory—but what about Bottoms and Tankebe's (2012) critique? In order to address the open and empirical question of whether it is possible to measure a normatively grounded sense of obligation to obey the police, the Scottish RCT of road policing (MacQueen and Bradford 2015, 2017) fielded measures of two forms of obligation to obey that were designed to have stronger face validity than prior indicators. Measures of normative obligation to obey the police were phrased to elicit truly free consent (e.g. 'I feel a moral obligation to obey the police' and 'I feel a moral duty to support the decisions of police officers, even if I disagree with them') while measures of non-normative obligation to obey the police were worded to elicit non-normative (instrumental) motives (e.g. 'People like me have no choice but to obey the police' and 'If you don't do what the police tell you they will treat you badly').

Our analysis has two stages. We first assess the open and empirical question of whether normative and non-normative forms of obligation to obey the police can be empirically disentangled using confirmatory factor analysis. If they cannot be separated empirically, then it would not make sense to define police legitimacy as obligation to obey the police alongside the perceived appropriateness of the institution. If instrumental and normative motives to obey cannot be disentangled, then it would be problematic to define police legitimacy partly as obligation to obey officer commands, because instrumental obligation is about coercive not consensual police-citizen relations—it does not 'sit well' with the belief that the police are a moral, just and proper institution.

To foreshadow the results of the first stage, our analysis indicates that they can be represented as two distinct latent constructs. The measures designed to tap into normative obligation loaded on one factor and the measures designed to tap into non-normative obligation loaded on another factor. So the second stage of analysis moves on to assess whether normative obligation to obey the police can reasonably be treated as a constituent component of police legitimacy. We do this by testing how it behaves in the procedural justice framework that the RCT tested. We also provide what is to our knowledge the first examination of the dynamics of non-normative obligation to obey the police. Might the experience of procedural justice decrease non-normative obligation to obey the police? If people feel relatively low level of power and autonomy in relation to the police, do they also tend to feel a dull compulsion towards the demands that officers might make on them? How does non-normative obligation relate to compliance and cooperation?

Figure 1 summarises the theoretical framework that we test. The treatment on the left-hand side represents whether the research participant was in the control group in the RCT (‘business as usual’) or the treatment group (the intended ‘procedurally just encounter’). Solid lines indicate a posited positive influence, while dashed lines indicate an uncertain predicted relationship where null or negative effects are possible given the lack of existing research.

**Is normative obligation to obey the police part of police legitimacy?**

If the form of obligation to obey the police that is measured by the three items—‘I feel a moral obligation to obey the police’, ‘I feel a moral duty to support the decisions of police officers, even if I disagree with them’ and ‘I feel a moral duty to obey the instructions of police officers, even when I don’t understand the reasons behind them’—can be reasonably claimed to be a normatively grounded sense of *truly free consent*, then we would expect a particular causal chain that goes from (a) the treatment, to (b) procedural justice and personal sense of power, to (c) normative and non-obligation obligation to obey the police, to (d) cooperation with the police and compliance with traffic laws.

The first factor relates to the procedural justice of the police in the police-citizen interaction. Procedural justice is a legitimating claim at the interpersonal level because procedural fairness is a key societal norm regarding how legal authorities should behave; when authority figures act in normatively appropriate ways, this encourages citizens to accept the rightfulness of an institution’s claim to power (Tyler 2006a, b). In the current context, if the officer conducting a vehicle stop acts in ways that are seen by research participants as respectful, neutral, accountable and trustworthy, then these

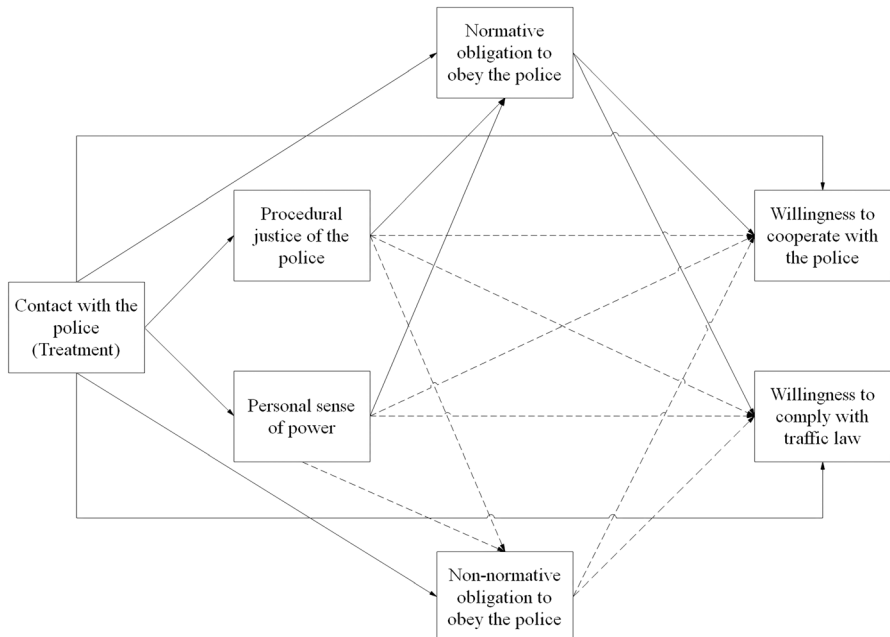


Fig. 1 Theoretical model for cooperation and compliance with two pairs of sequentially ordered mediators



citizens are expected to emerge from the encounter with a strengthened sense that police officers are entitled to enforce the law and have their decisions accepted and directives obeyed. Note, also, in Fig. 1, there is a direct relationship from treatment to normative obligation to obey the police that bypasses the indirect effect through perceived procedural justice. This represents any residual effect of the treatment on normative obligation that is not carried by a sense of fair treatment and neutral decision-making.

The second factor relates to people's sense of power and autonomy with respect to police. While this construct has received little attention in the literature, it is apposite in the current context given the *prime facie* nature of the two forms of obligation. People may be more likely to willingly allow police to dictate appropriate behaviour when they feel more autonomous and less bounded by police (cf. Tyler et al. 2015). For the sake of parsimony, we place personal sense of power as a mediator of the treatment, alongside procedural justice, rather than flowing out of procedural justice (although, see Appendix/D where we test that model). We predict that personal sense of power mediates some of the treatment effect on normative obligation. As with procedural justice, the expectation is that induced variation in the personal sense of power will be associated with higher average levels of normative obligation. Prior research has shown that procedural justice has an empowering and/or power equalising quality, with the experience of procedural justice decreasing people's sense of power distance regarding the police (Mentovich 2012). If normative obligation reflects a sense of active and willing consent, we expect some of the effect of the manipulation on normative obligation to run through personal sense of power.

The pathways from the treatment to compliance with traffic laws that flow through normative obligation reflect the idea that police behaviour can enhance or weaken voluntary deference to those who enforce traffic laws, which in turn can help to motivate voluntarily deference. The pathways from the treatment to willingness to cooperate with the police that flow through normative obligation reflect the idea that police behaviour can enhance or weaken deference to those who enforce the law (through procedural justice and/or personal sense of power) and deference may motivate a willingness to cooperate because (a) people are aware that the police want citizens to report crimes and provide information important to the investigation and (b) deference means allowing police to dictate appropriate behaviour in proactive as well as reactive ways.

### **What are the dynamics of non-normative obligation to obey the police?**

The dynamics exhibited by non-normative obligation may shed light on the potential effects of less positive police-citizen encounters. Is there a zero-sum game going on with normative and non-normative forms of obligation? Does procedural justice policing increase normative obligation and decrease non-normative obligation? If so, do cooperation and compliance respond in diametrically opposed ways to these two shifting motivations? As denoted by the dashed lines in Fig. 1, theoretical expectations are more tentative for the second form of obligation, but we posit a negative relationship between people's experience of procedural justice and non-normative obligation. Legitimacy transforms power into authority, and a positive encounter may help to turn a non-normative form of authority relations into something consensual. On the one hand, the experience of procedural justice

may shift concerns about the consequences of non-obedience to willing deference to police orders. On the other hand, a procedurally unjust encounter may increase prudential and/or instrumental obligation if procedural injustice serves to remind those experiencing disrespectful treatment and biased decision-making of the asymmetrical power relations that police officers have over citizens.

We also test whether some of the effect of the treatment on non-normative obligation goes through personal sense of power. Non-normative obligation reflects (among other things) a sense of pragmatism in the face of powerlessness, and we examine whether a negatively experienced encounter with the police is associated with lower levels of subjective power and a consequent increase in non-normative obligation (i.e. an increase in one's sense of dull compulsion and fear of the consequences of non-obedience). We also expect a set of negative pathways from the treatment to willingness to comply with traffic laws flowing through decreased procedural justice or sense of power and increased non-normative obligation. Finally, non-normative obligation may have a negative downstream effect on cooperation because people who feel powerless or afraid of the police may be unlikely to come forward voluntarily.

## The methodological contribution

In addition to the above conceptual and substantive goals, our paper makes a methodological contribution through the use of a recently developed form of causal mediation analysis. Mediation analysis has been typically carried out in the structural equation modelling (SEM) framework using the product method (Baron and Kenny 1986). This approach has several limitations including dangerous assumptions that are difficult to satisfy (e.g. effect homogeneity), modelling constraints (linearity) and difficulties handling multiple mediators and interactions (Pósch 2020). Recent advancements in mediation analysis have, among others, focused on mediation analysis in groups (Mackinnon et al. 2013), longitudinal mediation analysis (Preacher 2015) and most pertinent to this paper: causal mediation analysis. Causal mediation analysis is a challenge even with a single mediator; however, the nature of the current inquiry—where we hypothesise a chain of causal pathways where the causal effect is created by the (a) treatment (experimental condition) that is transmitted by (b) procedural justice and/or sense of power over the police through (c) normative and/or non-normative obligation to obey on (d) compliance and cooperation—makes the issue more complex.

Kenny (2008: 356) emphasised that mediation analysis is inevitably 'a form of causal analysis'. The critical consideration in the current study is to isolate what exactly seemed to change as a result of the manipulation and to estimate the relevant causal pathways leading from the treatment to the downstream constructs (Fig. 1). By identifying causally mediating effects, we can explain why and how the effect of the treatment (here, a kind of regulatory contact with the police that was judged by and large as more negative or more positive) has an effect on the various outcomes. For example, the effect of treatment (T) on normative obligation is posited to partly run through procedural justice as one mediator ( $M_1$ ) or personal sense of power as another mediator ( $M_2$ ). By focusing on the indirect effects, we can answer the question *how* and *why* the treatment influenced the outcome, i.e. to what extent the impact of the treatment is attributable to changes in other constructs such as procedural justice or personal sense of power.

### Causal mediation analysis with causally ordered mediators

Causal mediation analysis requires several identification assumptions that, if satisfied, lend the derived statistical estimates a causal interpretation (Manski 2007). This set of assumptions is usually referred to as sequential ignorability or conditional ignorability (Imai et al. 2011; Pearl 2001). For a single causal mediator, with treatment, *T*, a single mediator, *M*, an outcome, *Y*, and a vector of pre-treatment covariates, *C*, it requires that:

- i. The effect of *T* on *Y* is unconfounded controlling for *C*
- ii. The effect of *M* on *Y* is unconfounded controlling for *C* and *T*
- iii. The effect of *T* on *M* is unconfounded controlling for *C*
- iv. None of the confounders of *M* are affected by *T*

Importantly, randomisation of the treatment only accomplishes (i) and (iii). Assumption (ii) prescribes that there cannot be any unmeasured confounder which would affect the relationship between the mediator and the outcome. This can be achieved either by applying special design-based strategies (Pirlott and Mackinnon 2016) or relying on an adequate set of pre-treatment covariates *C* (which were unaffected by the treatment). Finally, (iv) demands that there cannot be any confounder of *M* which was affected by *T*. Crucially, such a post-treatment confounder would act as a second mediator (*L*), providing an alternative conduit transmitting the effect of the treatment both through itself and through *M*. This means that the current set of identification criteria only applies to models with a single mediator, which when satisfied allow researchers to derive generalised natural direct and indirect effects where these effects have assumed causal properties.

However, the current application (Fig. 1) posits four pathways with two causally ordered mediators, where procedural justice or sense of power are the first intermediate variables, followed by one of the obligation to obey constructs. Figure 2 represents this model, where *L* is the first mediator or post-treatment confounder, and *M* is the second mediator now transmitting the effects of both *T* and *L*. Importantly, the following four pathways emerge with two sequential (or ordered) mediators:

- (1) Treatment → cooperation with police/compliance with the law,
- (2) Treatment → procedural Justice/sense of power → cooperation with police/compliance with the law,
- (3) Treatment → normative/non-normative obligation to obey the police → cooperation with police/compliance with the law, and finally,
- (4) Treatment → procedural justice/sense of power → normative/non-normative obligation to obey the police → cooperation with police/compliance with the law.

According to the sequential ignorability assumption, however, there can be only a single mediator affected by the treatment (iv). If there are multiple mediators, this assumption is not violated, provided that these mediators are causally independent (orthogonal) of each other. This is assumed for procedural justice and personal sense of power, and normative and non-normative obligation respectively. However, the model used here also posits conditionality. It tests the mediated impact of procedural justice

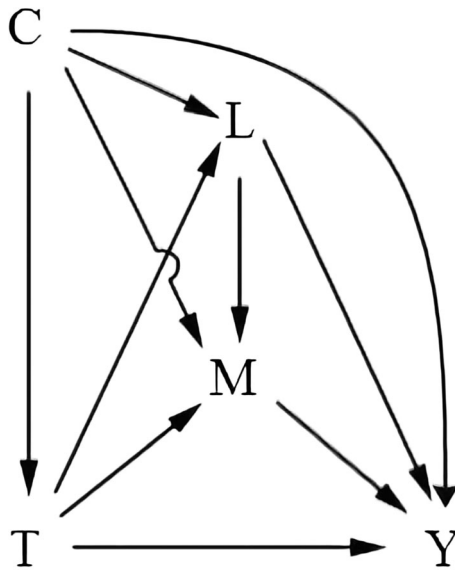
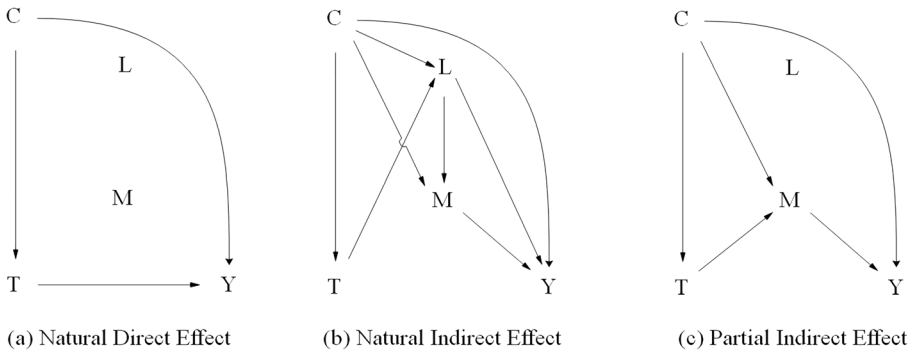


Fig. 2 Mediation analysis with two mediators

and sense of power on cooperation and compliance through normative and non-normative obligation to obey the police (4). Hence, in the presence of post-treatment confounders, mediated effects are not identifiable based on the previously outlined ignorability criteria.

An elegant resolution of this violation of a key assumption is to shift the focus from single mediators to a vector of mediators (Lange et al. 2014; Steen et al. 2017a; VanderWeele and Vansteelandt 2014). The difference between a mediator  $M$  and a post-treatment confounder  $L$  is only substantive; otherwise, they are statistically equivalent, which means that any variable affected by treatment  $T$  can be added to the vector that will then be robust to unmeasured common causes of various mediators. This approach is not sensitive to the initial ordering of the mediators. It also allows for interactions (moderated effects) to be taken into account. Thus, this approach partitions the different pathways to a natural direct effect (NDE) (pathway 1) and a (joint) natural indirect effect (NIE) (pathways 2–4). Estimating the joint indirect effect provides a robust test of the underlying causally mediated mechanism. Another advantage of this approach is that handling the mediators as a vector only requires a small change in the identification assumptions that need to apply to a vector of mediators instead of one mediator (e.g. procedural justice and normative obligation, or  $M$  and  $L$ ).

Even though these joint indirect effects may shed light on the grouped overall causally mediated effects of the studied constructs, this approach limits the scope of the analysis, because it does not allow for the assessment of specific pathways (pathways 2, 3 and 4 separately). It fails therefore to address the initial mediation hypotheses. A further issue is that this perspective does not permit the test of the order of the mediators either. Fortunately, some recent advancements in natural effect models (Steen, Moerkerke and Vansteelandt 2017) allow further partitioning of the joint



**Fig. 3** Mediation analysis with two mediators and three-way decomposition

indirect effect. Instead of relying on a two-way decomposition, this approach offers a three-way decomposition in the case of two sequential mediators<sup>1</sup>: the NDE remains the same (pathway 1), the NIE now incorporates both pathways (2) and (4) and pathway (3) will be the semi-natural/partial indirect effect (PIE). In other words, the joint natural indirect effect can be partitioned to a natural indirect effect that includes all pathways going through procedural justice/sense of power (L), including the ones that also go through the obligation constructs, and a semi-natural/partial indirect effect, which contains the indirect effect of normative/non-normative obligation (M) that does not go through L. Figure 3 depicts the decompositions in case of the separate pathways approach. The estimation of the joint indirect, natural indirect and partial indirect effects requires the imputation of potential outcomes and some further analytical steps which the interested reader can find in Appendix/A.

Because there are two Ls and Ms, it needs to be maintained that the pathways through procedural justice and sense of power and normative and non-normative obligation are none-intertwined, i.e. are independent of each other. Furthermore, to identify these effects and in addition to the extended sequential ignorability discussed for the joint indirect effect, two further assumptions need to be met:

- v. The effect of L on M is unconfounded controlling for T and C
- vi. None of the L and M confounders are affected by T

In addition to the above assumptions, researchers need to have a clear justification for the posited causal structure and order of their tested models. There is no statistical test which could establish the veracity of a causal order; every causal model must be informed by the available empirical evidence and theory (VanderWeele 2015). Although some scholars argue that causal structure should be established relying on temporal order and direction (Preacher 2015; Walters and Mandracchia 2017), these models also require a priori model specification and have challenges of their own, especially handling time-varying confounders (Clare et al. 2019; VanderWeele and Tchetgen Tchetgen 2017). Many causal mediation scholars, however, analyse RCTs in a similar manner to us (Imai et al. 2011; Imai and Yamamoto 2013; Keele et al. 2015).

<sup>1</sup> This method can be extended to more than two causally ordered mediators, as shown in Appendix/D.

The psychological literature has established that more fundamental psychological processes inform more complex ones, forming a unidirectional hierarchy (Von Hippel, Lakin and Shakarchi 2005; Kahneman 2011). Both the perception of fairness (procedural justice and the ‘fairness heuristic’, Lind 2001; Proudfoot and Lind 2015) and relational considerations (sense of power, Barclay et al. 2017) are basic psychological processes compared to legitimacy (which requires cognitive deliberation) or the behavioural outcomes of cooperation and compliance (which require taking action based on heuristics and the available evidence). Our model relies on this literature which provides qualified support to our models.

## Method

### Design

ScotCET was funded by the Scottish Government to inform the Justice Strategy for Scotland (MacQueen and Bradford 2015). Vehicle stops were conducted during the Scottish Festive Road Safety Campaign 2013/14, with police officers stopping citizens while driving to check the alcohol levels of drivers and conduct routine vehicle safety checks. The 20 road police units involved in the study were divided into 10 matched pairs, primarily according to shared geographical characteristics. Within each pair, one unit was randomly assigned to the control group and the other unit was randomly assigned to the treatment group (i.e. block-randomisation). On the one hand, the control group conducted ‘business as usual’ stops throughout the campaign. On the other hand, the treatment group, following a short ‘pre-period’ operating ‘business as usual’, received a combination of verbal and written instructions on how to successfully apply a procedurally just model of policing during routine encounters, aiming to communicate or enable the core aspects of procedural justice: dignity and respect, equality, trustworthy motives, neutrality of decision-making, clear explanation and the opportunity for citizen participation or ‘voice’.

ScotCET was designed as a partial replication—and extension—of QCET (Mazerolle et al. 2013). The main objective of QCET was to test the effect of introducing a procedural justice script to police activity in the context of traffic stops in Queensland. Mazerolle and colleagues found that procedural justice was higher, on average, among the treatment group compared to the control group. There was also a statistical effect of the manipulation on legitimacy, some of which seemed to be transmitted via procedural justice: path analysis (using the standard, non-causal approach to statistical mediation) indicated that procedural justice explained a good deal of the variation of legitimacy, that the treatment seemed to have an effect on legitimacy and that procedural justice explained some of the statistical effect of the treatment on legitimacy.

While QCET worked, ScotCET did not. There was a treatment effect, but it was in the opposite expected direction (MacQueen and Bradford 2015). A follow-up qualitative study suggested the possibility that there was a series of communication-based errors in the implementation process that occurred in a context of policing reform and a general perception of organisational injustice within the force (MacQueen and Bradford 2017). These factors seemed to have combined to produce a diffuse, negative effect on

the attitudes and behaviours of the officers involved. To give two examples, evidence from the focus groups suggested first, that officers felt the script impugned their professional integrity and represented an unwarranted intervention into their working lives. This may have affected their performance in an organisational context marked by significant, often unpopular and certainly poorly-communicated change (the formation of Police Scotland from the eight former regionally-based forces). Officers may have already been feeling relatively unhappy, and being asked to deliver the intervention may have made them unhappier still. Second, even when the script was implemented properly and by an officer in a positive frame of mind, it may have produced negative outcomes if, for example, it bureaucratized the encounters and/or made officer behaviour more formal and less natural. Relatedly, there is evidence from the QCET trial that longer encounters can be perceived by citizens as less procedurally fair (Mazerolle et al. 2015); the script may have increased the duration of the interview.

In short, what seems to have happened was not simply a case of implementation failure, which would have most likely led to 'nil' effects from the experiment, but rather that the design of the intervention, and the way it was communicated to the officers who delivered it, combined to produce a negative rather than a positive effect on driver perceptions. We can, however, extract significant value from the study: there was a treatment effect, it was just in the opposite expected direction. Recall, also, that Mazerolle et al.'s (2013) focus was not only on the link between treatment and procedural justice, but also on other, downstream constructs like legitimacy, cooperation and compliance. It is this secondary aspect of ScotCET that we focus on in the current paper, i.e. the impact of the treatment on constructs further down the potential causal chain. This can be done on the current dataset due to Pösch's (2020) reassessment of the implementation failure, which found treatment effect consistency and homogeneity and no selection bias. In other words, even though we do not possess additional information about what exactly transpired during the police contacts, we know that the differences caused by the treatment are attributable to the research design, i.e. that the treatment effects are reliable. Causal mediation analysis permits us to explain how and why the treatment affected the outcomes by relying on the well-defined mediators and focussing on the indirect effects.

## Survey and measures

All drivers stopped were issued with a self-completion questionnaire with a prepaid envelope (an online alternative was also offered). Excluding the baseline 'pre-period' of the trial (305 questionnaires returned), 510 questionnaires were returned. The response rate was 6.6%, which on the face of it is concerning. However, Groves and Peytcheva's (2008) assessment of response rates in the context of surveys and Hendra and Hill's (2018) assessment in the context of RCTs suggest that response rates are largely unrelated to nonresponse bias. Hendra and Hill (2018: 16) conclude that: '[our] findings are consistent with other research on the relationship between response rate survey generalizability, showing no substantive relationship between nonresponse bias and response rate.' A more pressing question is whether the average treatment effect (ATE) is consistent with the population average treatment effect (PATE). As discussed by Kohler et al. (2018), there are other factors, such as the effect of the treatment on self-selection and treatment effect heterogeneity that could have had a more serious

impact on the ATE and PATE equivalence than the response rate. Importantly, Pósch (2020) recently reassessed the ATE of ScotCET and found no sign of (1) self-selection bias (i.e. the same proportion and same kind of people returned the questionnaire pre- and post-treatment); (2) treatment effect heterogeneity (i.e. the ATE did not vary based on pre-treatment characteristics); and (3) treatment effect inconsistency (i.e. the magnitude of the ATE was approximately the same across the different matched pairs). These results imply that the ATEs are likely to be equivalent to the PATEs, making them representative to the study population, and the results transportable to the whole population of commuters in Scotland (Kohler et al. 2018; Pearl and Bareinboim 2012).

In terms of descriptive statistics, broad equivalence between treatment and control groups was achieved. Overall, 63% of respondents were male, the mean age was 50.7 (SD = 14.8, min = 17, max = 87), three quarters (77%) of respondents were homeowners. 41% had a university degree or higher (12% reported holding no qualifications) and the majority were employed (71%), and 73% were married or in a relationship.

Procedural justice was measured by asking research participants whether the officer seemed approachable and friendly, helpful, respectful, professional, fair and clear in explaining why they had been stopped, whether they trusted the intentions of the officer involved, whether they were confident that the officer was doing the right thing, whether the officer gave them the opportunity to express their views and whether the officer listened to what they had to say. Response alternatives ranged from 1 'no, not at all' to 4 'yes, completely'. Sense of power was measured by a single item, 'How much power do you think people like you have over the police?'. Response alternatives ranged from 1 'very little power' to 4 'a lot of power'.

Normative obligation was measured by asking research participants the extent to which they either agreed or disagreed to the following statements: 'I feel a moral obligation to obey the police', 'I feel a moral duty to support the decisions of police officers, even if I disagree with them' and 'I feel a moral duty to obey the instructions of police officers, even when I don't understand the reasons behind them'. Response alternatives were 'strongly disagree', 'disagree', 'neither agree nor disagree', 'agree' and 'strongly agree'.

Non-normative obligation was measured by asking respondents the extent to which they agreed or disagreed to the following statements: 'People like me have no choice but to obey the police', 'If you don't do what the police tell you they will treat you badly' and 'I only obey police because I am afraid of them. Response alternatives were 'strongly disagree', 'disagree', 'neither agree nor disagree', 'agree' and 'strongly agree'.

Willingness to cooperate with the police was captured by asking research participants 'If the situation arose, how likely would you be to ', 'call the police to report a crime you had witnessed', 'help police to find someone suspected of a crime by providing information' and 'report dangerous or suspicious activities to the police'. The response alternatives ranged from 1 'not likely at all' to 4 'very likely'.

Willingness to comply with traffic laws was measured 'All things considered, how likely are you in the future to break the speed limit while out driving' and 'All things considered, how likely are you in the future to jump a red light if you are in a hurry'. The response alternatives ranged from 1 'not likely at all' to 4 'very likely'.



## Results

### Scaling

Results from two fitted CFA models using MPlus 7.2 are shown in Table 1 (indicators were set as categorical and all latent constructs were allowed to covary). Included were multiple indicators of procedural justice, normative obligation, non-normative obligation, legal compliance and willingness to cooperate. Each model also had a single indicator of personal sense of power in its original form, set to be correlated with the latent variables. In the five-factor model, the indicator of non-normative obligation ‘People like me have no choice but to obey the police’ was allowed to cross-load, specifically onto both normative and non-normative obligation (where it loaded negatively on normative obligation and positively on non-normative obligation). This was motivated by Tankebe’s (2013) use of the indicator as a reverse-coded measure of obligation. The exact and approximate fit statistics suggest that the five-factor (M1) fits the data adequately, at least according to the approximate fit statistics, where one typically looks for CFI > .95; TLI > .95; RMSEA < .08 (see Kaplan 2008). The four-factor model, combining normative and non-normative obligation, has a relatively poor approximate fit, at least when judged on the basis of RMSEA and TLI. The chi-squared statistics are significant in both cases, indicating a bad fit, which is expected when there is a relatively large sample size.

We also estimated the reliability of each scale using Cronbach’s Alphas and MacDonald’s Omegas. It is always advisable to report multiple reliability coefficients as they may differ due to the way they are estimated (Hendrick et al. 2013). For instance, MacDonald’s Omega—unlike Cronbach’s Alpha—does not assume that each item contributes the same amount to the overall scale but considers the relative weight of each of them (Widaman et al. 2011). In our case, all scales performed well with all but one having reliability coefficients of 0.65 or higher (Table 2). The only exception was compliance with traffic laws, which still had Alpha and Omega values higher than 0.55. This scale only contained two items which usually produces lower scores than scales with more items.

As a robustness check, we split our sample in half at random and carried out an exploratory (EFA) and a confirmatory (CFA) factor analysis on each half of the data. The EFA was estimated using oblimin rotation (where the various factors were allowed to covary with each other) and a requirement of each emerging factor having an eigenvalue of at least one. Four factors emerged from the analysis: (1) procedural justice factor; (2) normative duty to obey and compliance with traffic laws factor; (3)

**Table 1** Fit statistics for two fitted CFA models

Confirmatory factor analysis models	Chi-square	df	<i>p</i>	RMSEA	RMSEA 90% CI	CFI	TLI
Five-factor model	268	104	< .005	.044	.038, .051	.992	.989
Four-factor model (combining normative obligation and non-normative obligation)	1801	110	< .005	.137	.132, .143	.915	.894

**Table 2** Reliability measures and factor loadings of the used scales

	Cronbach's Alpha ( $\alpha$ )	MacDonald's Omega ( $\omega$ )	Factor loadings ( $\lambda$ s) (EFA)	Factor loadings ( $\lambda$ s) (CFA)
Procedural justice	0,878	0.872	0.705–0.747	0.726–0.821
Normative duty to obey	0.800	0.768	0.539–0.600	0.675–0.850
Non-normative duty to obey	0.694	0.659	0.528–0.660	0.550–0.854
Willingness to cooperate	0.881	0.866	0.624–0.672	0.776–0.909
Compliance with traffic laws	0.588	0.569	0.316–0.355*	0.529–0.618

willingness to cooperate factor; and (4) non-normative duty to obey factor. The factor loadings on each scale were higher than 0.5, with the exception of compliance with traffic laws which had a weak loading on all factors (see Table 2). The only cross-loading that reached 0.3 was the third item of duty to obey ( $\lambda = 0.312$ , 'I feel a moral duty to obey the instructions of police officers, even when I don't understand the reasons behind them') on the factor of procedural justice. The Kaiser-Meyer-Olkin measure of sampling adequacy indicated the data was appropriate for factor analysis (KMO = 0.866).

For the CFA, we constrained the factors so they would represent the scales. The factor loadings of all factors were above 0.5 (see Table 2). The goodness of fit statistics also implied that the measurement model fit the data well ( $\chi^2(105) = 3561.06$ ,  $p < 0.001$ ; CFI = 0.964; TLI = 0.953; SRMR = 0.049; RMSEA = 0.056, RMSEA 90% = 0.046–0.065). Overall, both the tests of reliability and additional robustness test of the measurement models indicate that the scales fit the data well.

Without going into extensive detail regarding the correlational results (Table 3), we make three observations. First, the treatment has a significant negative correlation (on the 5% or 10% level) with all constructs apart from non-normative obligation to obey,

**Table 3** Correlational results

Variables	Treatment	Procedural justice	Sense of power	Normative obligation	Non-normative obligation	Cooperation
Procedural justice	-0.103*					
Sense of power	-0.113*	0.547**				
Normative obligation	-0.144**	0.463**	0.387**			
Non-normative obligation	-0.010	-0.411**	-0.248**	-0.061†		
Cooperation	-0.094*	0.543**	0.410**	0.365**	-0.279**	
Compliance	-0.085†	0.223**	0.216**	0.236**	-0.061†	0.172**

† $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$

which appears to be unaffected by it. Second, non-normative obligation, despite having moderately strong significant negative association with procedural justice, sense of power and cooperation, only has a weak negative correlation with normative obligation. This implies that, despite having the same correlates, prudential obligation is not the opposite of consensual obligation. Finally, procedural justice has a strong positive association with personal sense of power, which, as we discuss a bit later, might contradict our assumption that sense of power and procedural justice are independent of each other.

### Natural effect models

In each model, the treatment, mediators and outcome variables are regressed onto gender, age, housing status, marital status, education, employment and whether a breath test was conducted during the police encounter. These pre-treatment covariates are not included in either Table 4 or 5 for visual ease, but their coefficients can be found in Appendix/B.<sup>2</sup> The negative values in the tables might be counterintuitive, but they simply indicate the unexpected—and unintended—direction of the treatment effect, which shows that receiving the treatment had detrimental effects on cooperation and compliance. The standard errors shown in Tables 4 and 5 are bootstrapped with 1000 replications. Although generally speaking the sum of the partial and natural indirect effects should approximately coincide with the corresponding joint indirect effect of the mediators, some slight discrepancies can arise when modelling with continuous mediators (see details in the Appendix/A).

### Willingness to cooperate with the police

The first four columns in Table 4 present the indirect effect of the treatment on cooperation through procedural justice, sense of power, normative and non-normative obligation. Procedural justice, sense of power and normative obligation have weak negative natural indirect effects of different levels of significance ( $NIE_{coop\_pj} = -0.076$ ,  $p < 0.1$ ,  $NIE_{coop\_pow} = -0.062$ ,  $p < 0.1$ ,  $NIE_{coop\_ob} = -0.079$ ,  $p < 0.05$ ) with significant negative natural direct effects of the treatment ( $NDE_{coop\_pj} = -0.168$ ,  $p < 0.1$ ,  $NDE_{coop\_pj} = -0.184$ ,  $p < 0.05$ ,  $NDE_{coop\_ob} = -0.165$ ,  $p < 0.1$ ). The effects of procedural justice and normative obligation are largely identical, implying that when fitted separately the perceived fair treatment by the police and normative obligation mediate the effect of the treatment on people's willingness to cooperate to a very similar degree. Sense of power has a weaker indirect effect, indicating that the effect of the treatment on cooperation operates through increasing people's empowerment when judging future police-citizen encounters. In contrast, non-normative obligation to obey has a non-significant positive indirect effect with the effect size very close to zero ( $NIE_{coop\_pob} = 0.012$ ,  $p > 0.1$ ) with a moderately strong significant direct effect ( $NDE_{coop\_pob} = -0.255$ ,  $p < 0.01$ ). Thus, non-normative obligation does not seem to transmit the treatment's effect on cooperation.

The fifth column shows the joint, natural indirect and partial indirect effects of procedural justice and normative obligation. Their joint indirect effect is significant,

<sup>2</sup> As they are pre-treatment variables, they take on the same values in each model for the respective outcome.

Table 4 Natural effect models with two causally ordered mediators for cooperation with the police

Cooperation	1	2	3	4	5	6	7	8
Procedural justice (NIE)		-0.076† [0.042]						
Sense of power (NIE)		-0.062† [0.032]						
Normative obligation (NIE)			-0.079* [0.038]					
Non-normative obligation (NIE)				0.012 [0.017]				
Pj and ob joint indirect effect					-0.104* [0.051]			
Pj—natural indirect effect					-0.028 [0.043]			
Ob—partial indirect effect					-0.073* [0.036]			
Pj and pob joint indirect effect						-0.072 [0.048]		
Pj—natural indirect effect						-0.066† [0.040]		
POb—partial indirect effect						-0.010 [0.039]		
Pow and ob joint indirect effect							-0.091* [0.041]	
Pow—natural indirect effect							-0.041 [0.088]	
Ob—partial indirect effect							-0.047† [0.038]	
Pow and pob joint indirect effect								-0.033 [0.031]
Pow—natural indirect effect								-0.043† [0.040]
POb—partial indirect effect								0.011 [0.091]
Treatment (NDE)	-0.168† [0.089]	-0.184* [0.091]	-0.165† [0.089]	-0.255** [0.096]	-0.140† [0.087]	-0.171* [0.085]	-0.155† [0.086]	-0.214* [0.091]

pj procedural justice, pow sense of power, ob free/normative obligation to obey the police, pob prudential/non-normative obligation to obey the police  
† $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; the squared brackets straddle the bootstrapped standard errors

**Table 5** Natural effect models with two causally ordered mediators for compliance with traffic laws

Compliance	1	2	3	4	5	6	7	8
Procedural justice (NIE)	-0.061† [0.039]							
Sense of power (NIE)		-0.069† [0.040]						
Normative obligation to obey (NIE)			-0.086* [0.041]					
Non-normative obligation (NIE)				0.001 [0.008]				
Pj and ob joint indirect effect					-0.113* [0.055]			
Pj—natural indirect effect					-0.055† [0.036]			
Ob—partial indirect effect					-0.058* [0.017]			
Pj and pob joint indirect effect						-0.078† [0.048]		
Pj—natural indirect effect						-0.058† [0.035]		
POB—partial indirect effect						-0.016 [0.037]		
Pow and ob joint indirect effect							-0.099† [0.052]	
Pow—natural indirect effect							-0.039† [0.022]	
Ob—partial indirect effect							-0.060* [0.018]	
Pow and pob joint indirect effect								-0.058 [0.040]
Pow—natural indirect effect								-0.056† [0.040]
POB—partial indirect effect								-0.012 [0.042]
Treatment (NDE)	-0.096 [0.112]	-0.088 [0.123]	-0.071 [0.119]	-0.158 [0.126]	-0.043 [0.113]	-0.079 [0.117]	-0.076 [0.117]	-0.118 [0.125]

pj procedural justice, pow sense of power, ob free/normative obligation to obey the police, pob prudential/non-normative obligation to obey the police  
 †p < 0.1; \*p < 0.05; \*\*p < 0.01; the squared brackets straddle the bootstrapped standard errors

going slightly below  $-0.1$  ( $JIE_{\text{coop\_pjob}} = -0.104, p < 0.05$ ), with an also significant direct effect ( $NDE_{\text{coop\_pjob}} = -0.140, p < 0.1$ ). When the causal ordering is considered, normative obligation is capable of reserving most of its mediated effect ( $PIE_{\text{coop\_ob}} = -0.73, p < 0.05$ ), solely transmitting the impact of the treatment, as procedural justice does not have a significant natural indirect effect ( $NIE_{\text{coop\_pjob}} = -0.028, p > 0.1$ ). This provides further support to the original hypothesis, indicating that normative obligation on its own is capable of mediating the previous contact's (treatment) impact on willingness to cooperate, even without the jointly mediated effect of procedural justice and normative obligation.

The joint indirect effect of procedural justice and non-normative obligation to obey (shown in column six) is not significant ( $JIE_{\text{coop\_pjob}} = -0.072, p > 0.1$ ), with a significant direct effect of the treatment ( $NDE_{\text{coop\_pjob}} = -0.171, p < 0.05$ ). Normative obligation has a non-significant partial indirect effect on cooperation ( $PIE_{\text{coop\_ob}} = -0.010, p > 0.1$ ), which takes the opposite direction compared to its sole NIE. The significant natural indirect effect of procedural justice ( $NIE_{\text{coop\_pjob}} = -0.066, p < 0.1$ ) indicates that when police acts in fair and neutral ways, this increases people's willingness to cooperate with the police.

The seventh column shows that sense of power and normative obligation have a significant joint indirect effect on willingness to cooperate ( $JIE_{\text{coop\_powob}} = -0.091, p < 0.05$ ). Further decomposition indicates that the partial indirect effect of consensual obligation to obey ( $PIE_{\text{coop\_powob}} = -0.047, p < 0.05$ ) mediates the effect of the treatment, while sense of power has a non-significant natural indirect effect with a comparable effect size ( $NIE_{\text{coop\_powob}} = -0.041, p > 0.1$ ). This implies that, while the pathway that only goes through legitimacy does mediate the effect of the treatment, the pathways going through personal sense of power do not. The direct effect of previous contact remains significant here as well ( $NDE_{\text{coop\_powob}} = -0.155, p < 0.1$ ).

Finally, column eight shows the non-significant joint indirect effect of sense of power and non-normative obligation to obey the police ( $JIE_{\text{coop\_powob}} = -0.033, p > 0.1$ ). Non-normative obligation to obey's natural indirect effect is virtually unchanged compared to its sole NIE, with a weak non-significant positive effect ( $PIE_{\text{coop\_powob}} = 0.011, p > 0.1$ ), while sense of power has a smaller, but significant negative natural indirect effect ( $NIE_{\text{coop\_powob}} = -0.043, p < 0.1$ ). These findings mirror the ones found for procedural justice and non-normative obligation, insofar as prudential obligation does not appear to transmit the treatment's effect on willingness to cooperate with the police, while the pathways going through sense of power do. The effect of the treatment remains significant here as well ( $NDE_{\text{coop\_powob}} = -0.214, p < 0.05$ ).

In sum, the results suggest that the joint indirect effects of normative obligation and either of the two first mediators (procedural justice or sense of power) have the strongest indirect effect on cooperation. Further decomposition shows that—as theorised earlier—normative obligation almost fully transmits the effects of the treatment on cooperation, while the pathways going through either of the first two mediators are non-significant in the same models (models 5 and 7). Moreover, after the three-way decomposition, the mediated effect of normative obligation remains significant at the 5% level in only one of the models. In contrast, non-normative obligation to obey does not appear to be significant in any of the models. It follows that adding non-normative obligation to obey at best has no impact, at worst, marginally diminishes the mediated

impact of procedural justice and sense of power, producing non-significant joint, but significant natural indirect effects. All in all, non-normative obligation does not transmit the impact of contact with the police, while in the same models, the pathways going through procedural justice and sense of power do transmit the effect. Thus, the statistical evidence seems to support the idea that only normative obligation, not non-normative obligation, has a causally mediated effect on people's willingness to cooperate with the police.

Conspicuously, after considering the causal ordering, procedural justice and sense of power only remained significant in the models with prudential obligation (models 6 and 8). Yet, one would expect these variables to have a natural indirect effect on willingness to cooperate in all models, at least due the joint pathways going through consensual obligation to obey the police. Unfortunately, the decomposition pursued by the current paper does not allow us to determine whether the pathway going through both duty to obey and one of the first mediators would provide significant results.

Furthermore, it is important to acknowledge the effect sizes of the persistently significant natural direct effects, which consistently surpassed the strength of the indirect effects in all models. This indicates that the selected variables do not fully mediate the relationship between the treatment and cooperation with the police, or in other words, that a certain portion of the variation in experiences with the police remains unaccounted for by the mediators utilised here. The candidate mediators are only imperfect conduits of the effect of the treatment on willingness to cooperate.

### Compliance with traffic laws

Because the pattern of results for traffic law compliance closely resembles the ones for cooperation, and for the sake of brevity, we only discuss the main takeaway lessons of the models. For a detailed description of the results, please refer to Appendix/C.

From the models fitted for compliance with traffic law (Table 5), the joint indirect effect of procedural justice and normative obligation stands out as the strongest ( $JIE_{\text{compl\_pjob}} = -0.133, p < 0.05$ , column 5), followed by the joint indirect effect of sense of power and normative obligation to obey ( $JIE_{\text{compl\_powob}} = -0.099, p < 0.1$ , column 7), and the sole NIE of normative obligation ( $NIE_{\text{compl\_ob}} = -0.086, p < 0.05$ , column 3). Markedly, and in contrast with the models for cooperation, the sequential approach shows that normative obligation does not solely mediate the impact of the treatment and that the pathways through both procedural justice ( $NIE_{\text{compl\_pj}} = -0.055, p < 0.1$ ;  $PIE_{\text{compl\_ob}} = -0.058, p < 0.05$ , column 5) and sense of power ( $NIE_{\text{compl\_pjob}} = -0.039, p < 0.1$ ,  $PIE_{\text{compl\_pjob}} = -0.060, p < 0.05$ , column 7) retain a significant impact on traffic compliance. Hence, willingness to comply with traffic laws not only seems to be influenced by normative obligation but also by people's understanding of the fairness of the police and their personal sense of power. In line with the earlier findings for cooperation, non-normative obligation does not mediate the effects of earlier contact with the police (column 6 and column 8). This raises further doubt whether 'command and control' focussed policing (that strengthens instrumental forms of obligation) can influence either cooperation with the police or compliance with the law in either direction (cf. Motz et al. 2019). Finally, the natural direct effects of the treatment are non-significant which implies that the mediators managed to successfully transmit a substantial share of the influence of contact with the police.

## Conclusion

According to PJT, to acknowledge the legitimacy of the police is to accept its authority (Tyler 2006a, b; Trinkner 2019). To measure police legitimacy, surveys typically ask people whether they agree or disagree with attitudinal statements like “You should obey the police even when you think the police are wrong” (alongside measures of institutional trust and/or normative alignment to capture the perceived right to power), with agreement to such a sentiment taken to signify the acceptance by citizens that police have the moral right to tell them what to do in certain legally prescribed circumstances. For instance, the word ‘should’ in such an attitudinal statement is assumed to carry with it a normative sense of prescription to willingly comply with police orders and accept police decisions (Gur 2018: 190). Combined with the a priori definition of police legitimacy as partly the right to dictate appropriate behaviour, this forms the basis of the conceptual claim that obligation to obey is one constituent component of police legitimacy. Researchers testing PJT then go on to empirically assess which dimensions of police conduct legitimise the police in the eyes of citizens (is it procedural justice and/or distributive justice and/or effectiveness and/or police respect for boundaries?) and whether legitimacy predicts compliance and cooperation, net of instrumental motivations to comply and cooperate (Sunshine and Tyler 2003; Tyler 2006a, b; Tyler and Fagan 2008; Bradford et al. 2014c; Trinkner et al. 2018).

The main goal of this paper was to respond empirically to Tankebe and colleagues’ (Bottoms and Tankebe 2012; Tankebe 2013; Johnson et al. 2014) influential argument that normative and non-normative motives to obey the police have not yet been properly disentangled empirically in the criminological literature, and that, as a consequence of this failure, police legitimacy should not be defined and measured in part as obligation to obey the police. While prior studies have found that obligation to obey is highly correlated with (a) the belief that police officers are procedurally fair, (b) the judgement that the institution is appropriate, proper and just and (c) the willingness to proactively cooperate with the police and comply with the law (Jackson 2018; Jackson et al. 2012; Tyler and Huo 2002; Tyler and Fagan 2008; Tyler 2006b), the possibility remains that standard survey methods conflate two different forms of obligation. When people respond positively to statements like ‘I should obey police decisions because that is the proper or right thing to do’ and ‘You should do what the police tell you to do even if you disagree’, this could reflect the view that it is dangerous to defy the police, or that they have little choice but to be obedient and comply with officer instructions.

This is an issue that demands empirical attention. If we, as a community of criminologists, are to jettison the notion of obligation to obey from the legitimacy concept, we should do so based on evidence. Researchers need to be confident that the measurement tools are capturing only a normatively grounded form of obligation to obey (the duty to follow an authority’s directives and uphold the norms and rules associated with group membership, reflecting a *voluntary* deference to authorities that flows from the internalization of group norms and values) and not, in addition to, or instead of, pragmatic or strategic compliance from people who believe they lack the power to resist.

One aim of our study was, then, to address the open and empirical question of whether it is possible to capture normative obligation to obey distinct from instrumental obligation to obey the police. We sought to add to the available conceptual discussion



and methodological resources by assessing how two scales of obligation operate in a procedural justice framework tested in a field experiment of police-citizen contact among road-users in Scotland. We assessed whether measures designed to capture normatively grounded sense of obligation (the items stress a moral duty to obey the commands of officers) and measures designed to capture instrumental-grounded sense of obligation (the items stress compliance through fear of reprisal and/or dull compulsion) formed empirically distinct constructs.<sup>3</sup> Having found that they were empirically distinct in the Scottish road-user context, we then tested whether measures designed to capture normative obligation exhibited the expected requisite empirical properties in terms of connection to procedural justice, social identity, personal sense of power, instrumental obligation, compliance with traffic laws and cooperation with the police.

What, then, have we learnt from the study? Exploiting the fact that the RCT induced variation in the experience of procedural fairness and personal sense of power, albeit in small and expectantly negative ways, we showed that normative obligation operated differently to non-normative obligation, at least in the context of road policing in Scotland. First, normative obligation was positively (and strongly) correlated with key theoretical variables, including experienced procedural justice, personal sense of power, willingness to cooperate and traffic law compliance. Second, non-normative obligation was negatively correlated with procedural justice, sense of power and cooperation. Third, and despite their similar correlates, normative and non-normative forms of obligation to obey showed a surprisingly weak and tenuous association, implying that the two constructs are largely independent of each other. Fourth, normative obligation to obey emerged as the most important causal mediator for willingness to cooperate, and one of the important mediators for compliance with traffic law. Fifth, and by contrast, non-normative obligation did not mediate the impact of the treatment on either cooperation or compliance, implying that non-normative considerations do not channel the influence of the previous contact towards either of the outcomes. Finally, procedural justice and sense of power only mediated the effect of the treatment on cooperation in the absence of normative obligation in the model, whilst they always mediated the impact on legal compliance in addition to normative obligation.

Our findings should be situated in ongoing debate about the nature of legitimacy (Tyler 2006a, b; Reisig et al. 2007; Bottoms and Tankebe 2012; Jackson et al. 2013; Johnson et al. 2014; Tyler and Jackson 2013; Hamm et al. 2017; Jackson and Bradford 2019; Trinkner 2019). On the one hand, we have presented evidence that normative obligation is a sense of active consent rooted in the experience of fair treatment, fair decision-making, the provision of voice, the belief that the officer had trustworthy intentions and a relative sense of power and autonomy with respect to the police. On the other hand, non-normative obligation seemed to be a resistant sense of dull compulsion to the restriction of freedom that police officers can represent. Researchers need to be confident that the survey questions they use tap solely into a sense of obligation that centres upon rights and responsibilities in the context of legal authority. If one defines legitimacy as the right to power and the authority to govern, then we

<sup>3</sup> In this, we built upon Akinlabi and Murphy's (2018) Nigeria-based study. They fielded measures of police legitimacy like 'Even if police are doing the wrong thing, I still feel a moral obligation to obey police' and measures of dull compulsion, e.g. 'It is better to obey police than argue with them' and 'If you do not cooperate with police, they will get tough with you'. They found that the concepts were empirically distinct and barely associated with each other.

think that it makes sense to include normatively grounded obligation to obey the police in the conceptual and operational definition of police legitimacy. On the basis of our findings, we recommend that scholars define and measure police legitimacy partly as obligation to obey office commands and, importantly, use indicators of obligation to obey the police that stress the phrase *moral duty*.

Lastly, causal mediation analysis appeared to be an effective tool to distinguish between the treatment's direct and indirect effects. Applied to a real-world setting of road policing in Scotland, the RCT tested the effect of altering the dynamics of police-citizen encounters on the outcomes of legal compliance and willingness to cooperate with effects mediated through (a) immediate outcomes (i.e. experienced fairness and sense of power over the police) and (b) distal outcomes (i.e. normative and non-normative obligation to obey). We recommend that in future studies where several sequential mediators are present, researchers carry out the analytical steps outlined earlier. They should examine the mediated effect with a single mediator, then the joint indirect effects, and only after that turn to further decomposition. Decomposing the effects to natural and partial indirect effects is crucial, because this is often the only way to identify which mediator is actually influential in the model.

## Limitations of the analysis

Although the three-way decomposition presented here can help to unpack the underlying causal mechanisms, some difficulties still prevail when interpreting the various effects. The partial indirect effects are straightforward, as they represent the pathway going through only the second mediator towards the outcome. By contrast, the natural indirect effects incorporate both the pathway going through the first and second mediators (i.e. their jointly mediated effect). This means that the interpretation of the natural indirect effects for procedural justice and sense of power is murkier than it would be desirable, and future studies might want to seek finer decomposition to elucidate the effects (Daniel et al. 2015; Pósch 2019).

Ultimately, causal mediation analysis models always need to be informed by the existing literature and judged against the causal identification criteria. The assumptions (i–iv) posit that there cannot be any unmeasured pre- or post-treatment confounders which could affect the emerging results. Although we controlled for a long list of covariates, capturing the most important demographic characteristics of the individuals and one aspect of the police stop (i.e. breath analysis), some potentially influential covariates (such as previous contact with the police or victimisation) were not measured. Unfortunately, due to the multiply nested counterfactuals, sensitivity analysis to unmeasured confounding can only be carried out if the sequential mediators do not affect each other (Daniel et al. 2015; Pósch 2019), so additional tests could only run for a subset of the models and even then, only for pre-treatment confounding. Future studies should include more covariates and compare their results to the benchmark of controls set in this paper.

Possibly the most serious limitation of the current study is the strong assumption of causal independence made for the pairs of normative and non-normative obligation to obey, and procedural justice and sense of power respectively. This assumption is fundamentally untestable, yet it determines the viability of the presented effect

decomposition, and the causal claims made throughout the article. For normative and non-normative obligation, we believe that there are strong reasons to assume that these constructs are independent of each other. The correlational evidence shows that they have a very weak barely significant relationship. Moreover, they appear to be functionally different. While consensual obligation seems to channel the effect of the treatment to the outcome variables, prudential obligation does not seem to transmit the same effect. Yet, other results make us more cautious regarding this claim, as both obligation items have fairly similar bivariate relationships with procedural justice and sense of power. Further studies are needed to establish whether this assumption of independence can be justified.

As far as procedural justice and sense of power are concerned, our assumption stands on an even shakier ground. Correlational results imply that these constructs are strongly related to each other. Moreover, there are other competing theoretical models from the one presented here which might be equally plausible. It is possible for instance that procedural justice informs how people evaluate their personal sense of power, which in turn influences their ideas about the legitimacy of the police. Such a model would require three causally ordered mediators (i.e. procedural justice → sense of power → normative/non-normative obligation to obey). Thus, to check the robustness of our results, we also ran natural effects models to test this proposition, which produced very similar results to the ones presented here (see Appendix/D), suggesting that loosening this assumption would not substantially change the conclusions drawn here.

## Thoughts on future research

Our findings support the idea that when police officers act in ways that accord with normative expectations regarding fair interpersonal treatment and decision-making, this helps to create a sense among citizens that the institution is normatively appropriate and therefore entitled to be obeyed (see also Cheng 2015; White, Mulvey and Dario 2016). Scotland is a country with relatively low crime rates and little history of the sort of tense and fraught police-citizen relations that one can find in some other parts of the world. In a country like this, people may tend to interpret the measures of normative obligation in the way that is intended by researchers. However, this may not be the case in a country like South Africa (Bradford et al. 2014b), Ghana (Tankebe 2009) or Brazil (Jackson et al. 2019), or indeed in certain communities in large Metropolitan cities in the USA. We encourage research in other parts of the world to see if similar findings emerge. Does, in other contexts, the entitlement to be obeyed aspect of legitimacy accord with the normative obligation, given its grounding in fair and legitimate authority relations?

One of the biggest remaining questions is about the dynamics of non-normative obligation to obey. The current findings imply that normative and non-normative considerations might have different downstream effects (Tyler et al. 2015). These findings chime with studies on the perception of procedural injustice that appear to influence outcomes very differently than procedural justice (Augustyn 2016; Reisig et al. 2018). They also contribute to our understanding of the potentially contrasting nature of normative and instrumental authority relations (Anderson et al. 2012; Mentovich 2012).

Finally, the natural effect models used here are only one approach in the big family of methods of causal mediation analysis. There are semi-parametric alternatives which allow for post-treatment confounding (Imai and Yamamoto 2013), g-computation solutions which can be used for sequentially ordered mediators and post-treatment confounding alike (Daniel et al. 2015; De Stavola et al. 2015; Pósch 2019) and so on. We hope that the current example will encourage other researchers who want to estimate causal indirect effects to immerse themselves in similar methods to the one applied here.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akinlabi, O. M., & Murphy, K. (2018). Dull compulsion or perceived legitimacy? Assessing why people comply with the law in Nigeria. *Police Practice and Research*, 19(2), 186–201.
- Anderson, C., John, O. P., & Keltner, D. (2012). The personal sense of power. *Journal of Personality*, 80(2), 313–344. <https://doi.org/10.1111/j.1467-6494.2011.00734.x>.
- Augustyn, M. B. (2016). Updating perceptions of (in)justice. *Journal of Research in Crime and Delinquency*, 53(2), 255–286. <https://doi.org/10.1177/0022427815616991>.
- Barclay, L. J., Bashshur, M. R., & Fortin, M. (2017). Motivated cognition and fairness: Insights, integration, and creating a path forward. *Journal of Applied Psychology*, 102(6), 867.
- Baron, R. M., & Kenny, D. A. (1986). Moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 173–182.
- 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.
- Bradford, B., Murphy, K., & Jackson, J. (2014a). Officers as mirrors: policing, procedural justice and the (re) production of social identity. *British Journal of Criminology*, 54(4), 527–550. <https://doi.org/10.1093/bjc/azu021>.
- Bradford, B., Huq, A., Jackson, J., & Roberts, B. (2014b). 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>.
- Bradford, B., Jackson, J., & Hough, M. (2014c). Police legitimacy in action: lessons for theory and practice. In M. Reisig & R. Kane (Eds.), *The Oxford handbook of police and policing* (pp. 551–570). Oxford: Oxford University Press.
- Bradford, B., Milani, J., & Jackson, J. (2017). Identity, legitimacy and “making sense” of police use of force. *Policing: An International Journal of Police Strategies and Management*, 40(3), 614–627. <https://doi.org/10.1108/PIJPSM-06-2016-0085>.
- Carrabine, E. (2004). *Power, discourse, and resistance: A genealogy of the Strangeways prison riot*. Farnham: Ashgate.
- Cheng, K. K. Y. (2015). Prosecutorial procedural justice and public legitimacy in Hong Kong. *British Journal of Criminology*, 57(1), 94–111.
- Clare, P. J., Dobbins, T. A., & Mattick, R. P. (2019). Causal models adjusting for time-varying confounding—a systematic review of the literature. *International journal of epidemiology*, 48(1), 254–265.
- Daniel, R. M., De Stavola, B. L., Cousens, S. N., & Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, 71(1), 1–14. <https://doi.org/10.1111/biom.12248>.

- De Stavola, B. L., Daniel, R. M., Ploubidis, G. B., & Micali, N. (2015). Mediation analysis with intermediate confounding: Structural equation modeling viewed through the causal inference lens. *American Journal of Epidemiology*, *181*(1), 64–80. <https://doi.org/10.1093/aje/kwu239>.
- Gerber, M., & Jackson, J. (2017). Justifying violence: legitimacy, ideology and public support for police use of force. *Psychology, Crime and Law*, *23*(1), 79–95. <https://doi.org/10.1080/1068316X.2016.1220556>.
- Gur, N. (2018). *Legal directives and practical reasons*. Oxford: Oxford University Press.
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias : a meta-analysis. *Public Opinion Quarterly*, *72*(2), 167–189. <https://doi.org/10.1093/poq/nfn01>.
- Hagan, J., McCarthy, B., Herda, D., & Chandrasekher, A. C. (2018). Dual-process theory of racial isolation, legal cynicism, and reported crime. *Proceedings of the National Academy of Sciences*, *115*(28), 7190–7199.
- Hamm, J. A., Trinkner, R., & Carr, J. D. (2017). Fair process, trust, and cooperation: moving toward an integrated framework of police legitimacy. *Criminal Justice and Behavior*, *44*(9), 1183–1212. <https://doi.org/10.1177/0093854817710058>.
- Hendra, R., & Hill, A. (2018). Rethinking response rates: new evidence of little relationship between survey response rates and nonresponse bias. *Evaluation Review*. <https://doi.org/10.1177/0193841X18807719>.
- Hendrick, T. A. M., Fischer, A. R. H., Tobi, H., & Frewer, L. J. (2013). Self-reported attitude scales: current practice in adequate assessment of reliability, validity, and dimensionality. *Journal of Applied Social Psychology*, *43*(7), 1538–1552. <https://doi.org/10.1111/jasp.12147>.
- Huq, A., Tyler, T., & Schulhofer, S. (2011a). Mechanisms for eliciting cooperation in counter terrorism policing: evidence from the United Kingdom. *Journal of Empirical Legal Studies*, *8*(4), 728–761.
- Huq, A. Z., Tyler, T. R., & Schulhofer, S. J. (2011b). Why does the public cooperate with law enforcement? The influence of the purposes and targets of policing. *Psychology, Public Policy, and Law*, *17*, 419–450.
- Huq, A., Jackson, J., & Trinkner, R. (2017). Legitimizing practices: revisiting the predicates of police legitimacy. *British Journal of Criminology*, *57*(5), 1101–1122.
- Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, *105*(4), 765–789. <https://doi.org/10.1017/S0003055411000414>.
- Jackson, J. (2018). Norms, normativity and the legitimacy of legal authorities: international perspectives. *Annual Review of Law and Social Science*, *14*, 145–165. <https://doi.org/10.1146/annurev-lawsocsci-110316-113734>.
- Jackson, J., & Bradford, B. (2019). Blurring the distinction between empirical and normative legitimacy? A methodological commentary on “police legitimacy and citizen cooperation in China”. *Asian Journal of Criminology*, *14*(4), 265–289.
- Jackson, J., & Gau, J. M. (2016). Carving up concepts? Differentiating between trust and legitimacy in public attitudes towards legal authority. In E. Shockley, T. M. S. Neal, L. PytlikZillig, & B. Bornstein (Eds.), *Interdisciplinary perspectives on trust: towards theoretical and methodological integration* (pp. 49–69). Berlin: Springer.
- Jackson, J., Bradford, B., Hough, M., Myhill, A., Quinton, P., & Tyler, T. R. (2012). Why do people comply with the law? *British Journal of Criminology*, *52*(6), 1051–1071.
- Jackson, J., Bradford, B., Stanko, E. A., & Hohl, K. (2013). *Just authority? Trust in the police in England and Wales*. Oxon: Routledge.
- Jackson, J., Brunton-Smith, I., Bradford, B., Oliveira, T. R., Posch, K., & Sturgis, P. (2019). Police legitimacy and the norm to cooperate: using a mixed effects location-scale model to estimate the strength of social norms at a small spatial scale. <https://doi.org/10.31235/osf.io/jb74u>.
- Jackson, J., Posch, K., Rodriguez-Oliveira, T., Bradford, B., Natal, A. L., Mendes, S. M. and Zanetic, A. (2020). ‘Fear and legitimacy in Sao Paulo, Brazil: police-citizen relations in a high violence, high fear context’. LSE working paper.
- 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.
- Jonathan-Zamir, T., & Harpaz, A. (2018). Predicting support for procedurally just treatment: the case of the Israel National Police. *Criminal Justice and Behavior*, *45*(6), 840–862. <https://doi.org/10.1177/0093854818763230>.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
- Kaplan, D. (2008). *Structural equation modeling - foundations and extensions* (2nd ed.). Thousand Oaks: SAGE.
- Keele, L., Tingley, D., & Yamamoto, T. (2015). Identifying mechanisms behind policy interventions via causal mediation analysis. *Journal of Policy Analysis and Management*, *34*(4), 937–963. <https://doi.org/10.1002/pam.21853>.

- Kenny, D. A. (2008). Reflections on mediation. *Organizational Research Methods*, *11*(2), 353–358.
- Kirk, D., & Papachristos, A. (2011). Cultural mechanisms and the persistence of violence. *American Journal of Sociology*, *116*(4), 1190–1233.
- Kohler, U., Kreuter, F., & Stuart, E. A. (2018). Nonprobability sampling and causal analysis. *Annual Review of Statistics and Its Application*, *6*(1), 149–172. <https://doi.org/10.1146/annurev-statistics-030718-104951>.
- Lange, T., Rasmussen, M., & Thygesen, L. C. (2014). Assessing natural direct and indirect effects through multiple pathways. *American Journal of Epidemiology*, *179*(4), 513–518. <https://doi.org/10.1093/aje/kwt270>.
- Lind, E. A. (2001). Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. In J. Greenberg & R. Cropanzano (Eds.), *Advances in organization justice* (p. 56–88). Redwood: Stanford University Press.
- Mackinnon, D. P., Kisbu-Sakarya, Y., & Gottschall, A. C. (2013). Developments in mediation analysis. In T. D. Little (Ed.), *Oxford handbook of quantitative methods* (Vol. 2, pp. 1–28). New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199934898.013.0016>.
- MacQueen, S., & Bradford, B. (2015). Enhancing public trust and police legitimacy during road traffic encounters: results from a randomised controlled trial in Scotland. *Journal of Experimental Criminology*, *11*(3), 419–443. <https://doi.org/10.1007/s11292-015-9240-0>.
- MacQueen, S., & Bradford, B. (2017). Where did it all go wrong? Implementation failure—and more—in a field experiment of procedural justice policing. *Journal of Experimental Criminology*, *13*(3), 321–345. <https://doi.org/10.1007/s11292-016-9278-7>.
- Manski, C. F. (2007). *Identification for prediction and decision*. Cambridge: Harvard University Press.
- Mazerolle, L., Antrobus, E., Bennett, S., & Tyler, T. R. (2013). Shaping citizen perceptions of police legitimacy: a randomized field trial of procedural justice. *Criminology*, *51*(1), 33–63.
- Mazerolle, L., Bates, L., Bennett, S., White, G., Ferris, J., & Antrobus, E. (2015). Optimising the length of random breath tests: results from the Queensland community engagement trial. *Australian & New Zealand Journal of Criminology*, *48*, 256–276.
- Mentovich, A. (2012). *The power of fair procedures - the effect of procedural justice on perceptions of power and hierarchy*. New York: New York University.
- Motz, R. T., Barnes, J. C., Caspi, A., Arseneault, L., Cullen, F. T., Houts, R., ... & Moffitt, T. E. (2019). Does contact with the justice system deter or promote future delinquency? Results from a longitudinal study of British adolescent twins. *Criminology*.
- Murphy, K. (2016). Turning defiance into compliance with procedural justice: understanding reactions to regulatory encounters through motivational posturing. *Regulation & Governance*, *10*(1), 93–109.
- Murphy, K., & Cherney, A. (2012). Understanding cooperation with police in a diverse society. *British Journal of Criminology*, *52*(1), 181–201. <https://doi.org/10.1093/bjc/azr065>.
- Murphy, K., Tyler, T. R., & Curtis, A. (2009). Nurturing regulatory compliance: is procedural justice effective when people question the legitimacy of the law? *Regulation and Governance*, *3*(1), 1–26. <https://doi.org/10.1111/j.1748-5991.2009.01043.x>.
- Murphy, K., Bradford, B., & Jackson, J. (2016). Motivating compliance behavior among offenders: procedural justice or deterrence? *Criminal Justice and Behavior*, *43*(1), 102–118.
- Nivette, A., Eisner, M., & Ribeaud, D. (2020). Evaluating the shared and unique predictors of legal cynicism and police legitimacy from adolescence into early adulthood. *Criminology*, *58*(1), 70–100.
- Nix, J., Pickett, J. T., & Wolfe, S. E. (2019). Testing a Theoretical Model of Perceived Audience Legitimacy: The Neglected Linkage in the Dialogic Model of Police–community Relations. *Journal of Research in Crime and Delinquency*. <https://doi.org/10.1177/0022427819873957>.
- Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence UAI'01* (pp. 411–420).
- Pearl, J., & Bareinboim, E. (2012). External validity : from do-calculus to transportability across populations. *Statistical Science*, *29*(4), 579–595. <https://doi.org/10.1214/14-STS486>.
- Pirlott, A. G., & Mackinnon, D. P. (2016). Design approaches to experimental mediation ☆. *Journal of Experimental Social Psychology*, *66*, 29–38. <https://doi.org/10.1016/j.jesp.2015.09.012>.
- Pösch, K. (2019). Testing complex social theories with causal mediation analysis and G-computation: toward a better way to do causal structural equation modeling. *Sociological Methods & Research*. <https://doi.org/10.1177/0049124119826159>.
- Pösch, K. (2020). Prying open the black box of causality: a causal mediation analysis test of procedural justice policing. *Journal of Quantitative Criminology*, In Press.
- Preacher, K. J. (2015). Advances in mediation analysis: a survey and synthesis of new developments. *Annual Review of Psychology*, *66*, 825–852. <https://doi.org/10.1146/annurev-psych-010814-015258>.

- President's Task Force on 21st Century Policing. (2015). Final report of the President's task force on 21st century policing.
- Proudfoot, D., & Lind, E. A. (2015). Fairness heuristic theory, the uncertainty management model, and fairness at work. *The Oxford handbook of justice in the workplace*, 371, 385.
- Reisig, M. D., & Lloyd, C. (2008). Procedural justice, police legitimacy, and helping the police fight crime: results from a survey of Jamaican adolescents. *Police Quarterly*, 12(1), 42–62. <https://doi.org/10.1177/1098611108327311>.
- Reisig, M. D., Bratton, J., & Gertz, M. G. (2007). The construct validity and refinement of process-based policing measures. *Criminal Justice and Behavior*, 34, 1005–1027.
- Reisig, M. D., Mays, R. D., & Telep, C. W. (2018). The effects of procedural injustice during police–citizen encounters: a factorial vignette study. *Journal of Experimental Criminology*, 14(1), 49–58. <https://doi.org/10.1007/s11292-017-9307-1>.
- Sampson, R. J., & Bartusch, D. J. (1998). Legal cynicism and (subcultural) tolerance of deviance: the neighborhood context of racial difference. *Law & Soc'y Rev.*, 32, 777.
- Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of research in Crime and Delinquency*, 30(4), 445–473.
- Slocum, L. A., Ann Wiley, S., & Esbensen, F.-A. (2016). The importance of being satisfied. *Criminal Justice and Behavior*, 43(1), 7–26. <https://doi.org/10.1177/0093854815609069>.
- Steen, J., Loeys, T., Moerkerke, B., & Vansteelandt, S. (2017). Flexible mediation analysis with multiple mediators. *American journal of epidemiology*, 186(2), 184–193.
- Steen, J., Loeys, T., Moerkerke, B., & Steen, J. (2017a). Flexible mediation analysis with multiple mediators. *American Journal of Epidemiology*, 186(2), 184–193.
- Steen, J., Loeys, T., Moerkerke, B., & Vansteelandt, S. (2017b). Medflex : an R package for flexible mediation analysis using natural effect models. *Journal of Statistical Software*, 76(11), 1–45.
- Sun, I. Y., Wu, Y., Hu, R., & Farmer, A. K. (2017). Procedural justice, legitimacy, and public cooperation with police: does Western wisdom hold in China? *Journal of Research in Crime and Delinquency*, 54(4), 454–478. <https://doi.org/10.1177/00224278166638705>.
- Sun, I. Y., Li, L., Wu, Y., & Hu, R. (2018). Police legitimacy and citizen cooperation in China: testing an alternative model. *Asian Journal of Criminology*, 13(4), 275–291.
- 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), 513–548.
- 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>.
- Trinkner, R. (2019). Clarifying the contours of the police legitimacy measurement debate: a response to Cao and Graham. *Asian Journal of Criminology*, 14(4), 309–335.
- Trinkner, R., Jackson, J., & Tyler, T. R. (2018). Bounded authority: Expanding “appropriate” police behavior beyond procedural justice. *Law & Human Behavior*, 42(3), 280–293. <https://doi.org/10.1037/lhb0000285>.
- Trinkner, R., Mays, R. D., Cohn, E. S., Van Gundy, K. T., & Rebellon, C. J. (2019). Turning the corner on procedural justice theory: exploring reverse causality with an experimental vignette in a longitudinal survey. *Journal of Experimental Criminology*, 1–11.
- Tyler, T. R. (2003). Procedural justice, legitimacy, and the effective rule of law. *Crime and Justice*, 30, 283–357.
- Tyler, T. R. (2004). Enhancing police legitimacy. *The Annals of the American Academy of Political and Social Science*, 593, 84–99.
- Tyler, T. R. (2006a). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology*, 57, 375–400.
- Tyler, T. R. (2006b). *Why people obey the law*. Princeton: Princeton University Press.
- Tyler, T. R., & Fagan, J. (2008). Legitimacy and cooperation: why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law*, 6, 231–275. <https://doi.org/10.2139/ssrn.887737>.
- Tyler, T. R., & Huo, Y. J. (2002). *Trust in the law - encouraging public cooperation with the police and the courts*. New York: Russell Sage Foundation.
- Tyler, T. R., & Jackson, J. (2013). Future challenges in the study of legitimacy and criminal justice. In J. Tankebe & A. Liebling (Eds.), *Legitimacy and criminal justice - an international exploration* (pp. 83–104). Hoboken: Wiley.
- Tyler, T. R., & Jackson, J. (2014). Popular legitimacy and the exercise of legal authority: motivating compliance, cooperation, and engagement. *Psychology, Public Policy, and Law*, 20(1), 78–95.

- Tyler, T. R., & Trinker, R. (2017). *Why children follow rules: legal socialization and the development of legitimacy*. Oxford: Oxford University Press.
- Tyler, T. R., Schulhofer, S. J., & Huq, A. Z. (2010). Legitimacy and deterrence effects in counterterrorism policing: a study of Muslim Americans. *Law and Society Review*, 44, 365–402.
- Tyler, T. R., Fagan, J., & Geller, A. (2014). Street stops police legitimacy: teachable moments in young urban men's legal socialization. *Journal of Empirical Legal Studies*, 11(14), 751–785.
- Tyler, T. R., Jackson, J., & Mentovich, A. (2015). The consequences of being an object of suspicion: Potential pitfalls of proactive police contact. *Journal of Empirical Legal Studies*, 12(4), 602–636.
- Tyler, T. R., Goff, P. A., & MacCoun, R. J. (2015a). The impact of psychological science on policing in the United States: procedural justice, legitimacy, and effective law enforcement. *Psychological Science in the Public Interest*, 16(3), 75–109.
- Tyler, T. R., Jackson, J., & Mentovich, A. (2015b). The consequences of being an object of suspicion - potential pitfalls of proactive police contact. *Journal of Empirical Legal Studies*, 12(4), 602–636.
- VanderWeele, T. J., & Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiologic Methods*, 2(1), 95–115. <https://doi.org/10.1515/em-2012-0010>. **Mediation**.
- VanderWeele, T. J., & Tchetgen Tchetgen, E. J. (2017). Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3), 917–938.
- Vander Weele, T. J. (2015). *Explanation in Causal Inference - Methods for Mediation and Interaction*. Oxford: Oxford University Press.
- Von Hippel, W., Lakin, J. L., & Shakarchi, R. J. (2005). Individual differences in motivated social cognition: The case of self-serving information processing. *Personality and Social Psychology Bulletin*, 31(10), 1347–1357.
- Walters, G. D., & Mandracchia, J. T. (2017). Testing criminological theory through causal mediation analysis: Current status and future directions. *Journal of Criminal Justice*, 49, 53–64. <https://doi.org/10.1016/j.jcrimjus.2017.02.002>.
- White, M. D., Mulvey, P., & Dario, L. M. (2016). Arrestees' perceptions of the police: Exploring procedural justice, legitimacy, and willingness to cooperate with police across offender types. *Criminal Justice and Behavior*, 43, 343–364.
- Widaman K, Little T, Preacher, K, Sawalani G. (2011). On creating and using short forms of scales in secondary research. In: K. Trzesniewski, T. Little, R. Lucas (Eds.), *Secondary data analysis: an introduction for psychologists* (pp. 33-61). Washington, DC: American Psychological Association.
- 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(2), 253–282. <https://doi.org/10.1007/s10940-015-9263-8>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Krisztián Pósch** is an ESRC Postdoctoral Research Officer at the LSE's Department of Methodology. Chris's research interests include public perception of the police, especially procedural justice policing and quantitative methods with special focus on survey methods and casual inference.

**Jonathan Jackson** is a Professor of Research Methodology at the LSE's Department of Methodology. He is also an Honorary Professor at University of Sydney Law School. Jon's research interests centre around various issues in criminology and psychology and law.

**Ben Bradford** is Professor of Global City Policing at the Department of Security and Crime Science, UCL. He is also Director of the Institute for Global City Policing, an initiative joint funded by UCL, the Metropolitan Police Service and the Mayor's Office for Policing and Crime to promote policing research in London. Ben's research interests include public trust, police legitimacy, cooperation and compliance in justice settings, and social identity as a factor in all these processes.

**Sarah Macquoen** was formally employed by the University of Edinburgh, and was a member of the Scottish Centre for Crime and Justice Studies. She now works for the Scottish Government.



## Affiliations

Krisztián Pósch<sup>1</sup> · Jonathan Jackson<sup>2</sup> · Ben Bradford<sup>3</sup> · Sarah Macqueen<sup>4</sup>

✉ Jonathan Jackson  
j.p.jackson@lse.ac.uk

<sup>1</sup> Department of Methodology, London School of Economics, London, UK

<sup>2</sup> Department of Methodology, London School of Economics and University of Sydney Law School, Houghton Street, London WC2A 2AE, UK

<sup>3</sup> Department of Security and Crime Science, University College London, London, UK

<sup>4</sup> Scottish Government, Edinburgh, UK