THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Enhancing public trust and police legitimacy during road traffic encounters

OPEN ACCESS

Enhancing public trust and police legitimacy during road traffic encounters: Results from a randomized controlled trial in Scotland

Abstract

Objectives: This paper reports results from the Scottish Community Engagement Trial (ScotCET), devised to replicate the Queensland Community Trial (QCET). ScotCET was an RCT that tested the effects of 'procedurally just' policing on public trust and police legitimacy

Methods: A block-randomized (matched pairs) design, with pretest and posttest measures, was implemented in the context of road policing in Scotland. Participants were drivers stopped by police in December and January 2013/14 as part of Police Scotland's 'Festive Road Safety Campaign'. The experimental intervention comprised a checklist of key messages to include in routine roadside vehicle stops, and a leaflet for officers to give to drivers. Analysis proceeds via random effects regression models predicting latent variable measures of trust, satisfaction and legitimacy

Results: Contrary to expectations the intervention did not improve trust and legitimacy; rather, trust in the officers who made the stop, and satisfaction with their conduct fell in the test sites, relative to the controls, after implementation of the intervention. The intervention had no significant effect on general trust in the police, or on police legitimacy

Conclusions: Results demonstrate the difficulty in translating experimental interventions across policing contexts, and challenge the notion that public perceptions may be improved through a simple, additive approach to the delivery and communication of procedural justice.

Key words: legitimacy; procedural justice; road policing; trust.

Introduction

Attitudes towards the police and the legitimacy of police authority comprise an area of persistent and, at the present time, growing policy and academic interest. Over the last decade there has been an explosion of academic research in the field, and an increasing emphasis in policing policy and practice on enhancing public trust and confidence in, and conferment of legitimacy on, the police. In Scotland, the context for the present study, interest in these goals is evidenced by the development of the national 'Reassuring the Public' programme within the Scottish Government's Justice Strategy for Scotland, which aims to both reduce fear of crime and build confidence in Scotland's justice system, advocating for a system that '*treats people fairly and with respect*' (Scottish Government 2012:51). In line with the emerging and growing research evidence, the Justice Strategy rationalises this ambition by stating that improving public confidence is:

> "… likely to lead to better compliance and co-operation with the law and improved crime reporting and engagement with justice authorities. This approach will support community engagement and cohesion leading to people feeling safer in their homes and communities." (ibid.)

The underlying assumption (that fair and respectful treatment of people by the various components or interfaces of the justice system will lead to positive judgements and outcomes) is consistent with the literature on procedural justice theory, which posits that procedurally fair or just treatment directly influences satisfaction with, and confidence in, criminal justice agents and thereby enhances legitimacy. This paper presents the findings of a Randomized, Controlled Trial (RCT) designed to test this

assumption. Replicating the Queensland Community Engagement Trial (Mazerolle et al. 2011; 2012), the study tested whether written and verbal communication of procedural justice during routine encounters between members of the public and the police can indeed positively influence public opinion and enhance the legitimacy of police authority.

The RCT took place during a targeted road safety campaign run by Police Scotland in December 2013. The experimental intervention comprised road police officers delivering key messages of procedural justice to drivers during routine roadside vehicle stops, and distributing supporting leaflets following each encounter. It was hypothesized that this intervention would enhance perceptions of procedural justice, satisfaction with the officers conducting the encounter, general trust in the police and police legitimacy. However, to anticipate the findings described below, results suggest that the intervention in fact undermined drivers' sense that procedural justice had been adhered to during the stop, and their overall satisfaction with encounters. There was no statistically significant effect on general trust in the police, or on driver's legitimacy judgments. Results from the ScotCET experiment therefore suggest that operationalization of the procedural justice model is not straightforward. If police are to positively influence, or indeed simply maintain, existing levels of public trust a more nuanced consideration of the context, content and style of communication during encounters is required.

The paper proceeds in six parts. Part 1 outlines the procedural justice model, while part 2 lays out the hypotheses tested in the experiment. Part 3 discusses the design and

methods used in ScotCET and part 4 describes the data and approach to analysis. Part 5 provides the results while part 6 offers concluding comments.

Procedural justice and police legitimacy

Procedural justice theory, which seeks to understand individuals' reactions to the use of power and authority within group settings, was first developed by social psychologist Tom Tyler (2006; Tyler and Huo 2002), building on earlier work by Thibaut and Walker (1975) and others. Increasingly well evidenced by a growing body of international work (e.g. Bradford et al. 2014a; Hinds and Murphy 2007; Hough et al. 2013a, 2013b; Jackson et al. 2013; Jonathan-Zamir and Weisburd 2013; Murphy et al. 2008), procedural justice theory suggests that when people are interacting with others who have power over them, and who represent social groups they feel affiliation to or membership of, they are intensely attuned to the fairness of the *process* through which the interaction takes place. Fairness in this context means being treated with dignity and respect; allowed a voice in the interaction; and given clear communication of what is going to happen. The experience of fair treatment encourages a mutual sense of trust between the parties involved. When people experience procedural fairness and trust during an interaction with a power-holder they are more likely to accept final decisions or conclusions reached during a given encounter, more likely to be confident in the authority and to support and cooperate with it in the future, and more likely to grant it legitimacy. Studies usually find that, while people do care about the instrumental effectiveness of authorities, this is generally a less important predictor of important outcomes, such as legitimacy, than procedural fairness. An important exception here may be some

developing countries – here, some research suggests effectiveness may be relatively more important than procedural fairness (Bradford et al. 2014b; Tankebe 2009).

The relevance of procedural justice in policing contexts is obvious. Police officers rely on the support and cooperation of those they police, and, in the absence of relatively high levels of public trust and legitimacy, will be forced to rely on increasing levels of force, or at least the threat of force, to achieve desired aims. Should large numbers of people withdraw cooperation, policing would become more difficult, if not impossible. Moreover, research further suggests that individuals granting the police legitimacy are more likely to abide by the law (Jackson et al. 2012; Papachristos et al. 2012; Tyler 2006). Such normative commitment to the rule of law, produced and secured by legitimate legal authorities, is considered to be more stable and long lasting than law-abidingness generated by deterrent threat.

Procedural justice theory therefore envisages that an important 'two-step' process occurs when people interact with police officers (Jackson et al. 2013). First, dignity, respect, voice and trust during contact with officers will enhance people's sense that the encounter was procedurally just, increase levels of satisfaction with the officers involved, and generate higher levels of compliance in the immediate context. Second, however, a sense of trust and procedural justice generated during encounters with individual officers will have important downstream effects on more general opinions of the police, enhancing trust and confidence, generating legitimacy, and increasing propensities to cooperate with police and comply with the law in the future. It is important, then, that police and policy makers understand what generates, or

undermines, a sense among the public that the police operate in a procedurally just fashion, and how procedural justice is related to outcomes such as legitimacy.

According to process based models of policing (Myhill and Bradford 2012; Sunshine and Tyler 2003) personal interactions between officers and citizens are vital moments when trust and legitimacy are generated or undermined (Bradford et al. 2009; Skogan 2006). During personal contact with citizens police can communicate important messages to the public about their fairness, trust-worthiness and legitimacy. However, research in this area has, before now, been limited primarily to cross-sectional or, on occasion, panel designs (e.g. Bradford et al. 2014a; Tyler 2006). While there is much evidence concerning the *correlation* between a sense of procedural justice and, for example, legitimacy, it is much less certain that there is a causal link between these variables*.* It is not yet known to any degree of certainty whether increasing the procedural fairness of police activity has an immediate effect on the trust and legitimacy judgements of those exposed to that activity. Equally, studies seeking to consider how the principles of procedural justice can be operationalized at the level of everyday policing have also been rare (Mazerolle et al. 2013).

To address such questions, the Queensland Community Engagement Trial (QCET) adopted randomized field trial methodology to test the effect of police using the principles of procedural justice during routine encounters with citizens. Mass roadside random breath tests were the key focus and over 20,000 encounters were included in the trial, split broadly equally between control and experiment conditions. The baseline encounter here was very short, even abrupt, averaging just 30 seconds. Motorists would be pulled over, usually into a well-marked area with police cars and several officers in

obvious attendance. The officers conducting the breath tests would approach the car, ask the driver to blow into the breathalyser and, on return of a negative test, allow the driver on their way. Officers in the experimental condition of QCET, by contrast, followed a 'script' during their encounters which was designed specifically to communicate the messages and principles of procedural justice within each encounter. Drivers were greeted by police officers and had the random breath test operation explained to them. They were invited to ask questions. Encounters were closed with drivers being issued with a community newsletter and being thanked by the officers for complying with some element of road safety guidelines, e.g. wearing their seatbelt. Together the script and these gestures were intended to convey dignity and respect, neutrality of decision making, and trustworthy motives, as well as facilitate citizen participation. The trial found that the enhanced quality of interaction between public and police during the experiment encounters did indeed have a direct positive effect on: satisfaction of members of the public with the process and outcome of the encounter; perceptions of police fairness; respect for the police; trust and confidence in the police; and self-reported willingness to comply with police directives (Murphy et al 2014; Mazerolle et al, 2012; Mazerolle et al, 2011).

The key limitation of QCET has been the specificity of its context and the lack of replication to confirm the results achieved: the generalizability and reliability of the QCET findings must be tested elsewhere to determine whether the methods used to achieve the positive results may be successfully adopted in different policing contexts. It is this gap in existing knowledge that the current study seeks to address. ScotCET replicated, as far as possible, the methodology employed in QCET to test whether improved communication during routine encounters can positively influence:

perceptions of police officers and the belief that they adhere to principles of procedural justice; trust in the officers involved; satisfaction with police conduct during these routine encounters and their final outcomes; trust and confidence in the police more broadly; and police legitimacy.

Hypotheses

It was anticipated that the positive findings of the original QCET could be replicated in the Scottish context, such that an experimental intervention in routine encounters would shift levels of procedural justice, satisfaction, trust and confidence and legitimacy in a positive direction. Four specific hypotheses guide the analysis presented here (Figure 1 summarises these hypotheses, providing a conceptual map for the study):

H1 is a dual hypothesis. First, H1a proposes the experimental intervention will encourage a sense amongst drivers being stopped by police that they have been treated in a procedurally fair manner. Second, H1b proposes the intervention will increase driver's trust in the officers who stopped them.

H2 is that the intervention will increase overall satisfaction with the encounter.

H3 is that the intervention will result in higher levels of trust and confidence in the police in general.

 H4 is that the intervention will enhance the legitimacy of the police in general.

**Insert figure one here

The dotted lines in Figure 1 above indicate the underlying idea that satisfaction, trust and confidence and legitimacy will be enhanced in as much as a sense of procedural justice and trust *during the encounter* is enhanced – however, these relationships are not tested directly in the current paper. Rather, the focus here is whether the *intervention* itself had an effect on these different aspects of public opinion.

Design and methods

In this section we describe the design of ScotCET and the methods used in the experiment and analysis of the data. We start, however, with a brief description of the wider context of policing in Scotland, not least because this is required to understand the organization of road traffic policing and the design of the experiment.

Policing in Scotland

Although part of the United Kingdom, Scotland has its own systems of law and policing. Scottish police look like their counterparts south of the border, but operate according to a different set of laws and regulations. Moreover, following a significant reorganization in April 2013 Scotland has a single national police force, Police Scotland, replacing the previous arrangement of eight regional forces. Coming just before the development of ScotCET this reorganization had significant implications for the experiment that we describe below. In short, while nominally one organization, the 'legacy forces' within Police Scotland retained significant affective meaning for police and, perhaps, public. Officers we spoke to in the course of ScotCET were very aware of having 'come from'

Strathclyde, Lothian and Borders, Fife or other forces and constabularies, and it was a widespread assumption that part of the legacy of these forces was that there were different ways of doing policing within the new unitary organization.

Like QCET, ScotCET was designed in a road policing context, with experiment and control conditions encompassing all road police in Scotland. Police Scotland operates 20 road policing units, comprising 14 divisional (area-based) units, four trunk roads units and two motorcycle units, and these comprised the basic unit of randomization in the RCT (see below). As in the rest of the organization, staff in these units retained significant links, in terms of location and team membership, with the legacy forces.

Experimental design

ScotCET was implemented during the Festive Road Safety Campaign, an annual campaign aiming to prevent drink driving and encourage safe driving in winter conditions. Under the campaign a large volume of roadside stops with a shared focus were to be conducted, allowing a relatively high level of uniformity across the encounters to be included in the trial. The campaign ran for a five-week period over December 2013 and January 2014 and, prior to commencement, Police Scotland estimated that 20,000 roadside stops would be conducted over these five weeks. Very few of these stops would concern possible criminal offences; like QCET, the experiment was based around a high volume, relatively mundane encounter between police and public.

Initial scoping work soon suggested, however, that a direct replication of the QCET design and methodology would not be possible within the Scottish context. Random

breath testing is not permitted in Scotland, and police roadside stops are conducted on the basis of much broader issues of driver and vehicle safety. Encounters were therefore inevitably much more varied in terms of nature, focus and length than was the case in Queensland. Preliminary qualitative fieldwork also revealed a high level of interaction between drivers and police during roadside stops, with core elements of procedural justice already incorporated into the practice of many officers as a matter of course. Road police officers reported, and were observed: relying on verbal communication to explain the purpose of each stop and reassure drivers; adapting their delivery according to the situation and person encountered; and placing strong emphasis on achieving rapport with drivers. Thus, ScotCET retains a broadly similar design and method to the original QCET but with key contextual differences taken into account.

Business as usual

Routine encounters during the Festive Road Safety campaign involved a combination of 'mass vehicle stops', where sections of roads were used by a team of police officers to stop several vehicles at once, and individual stops, where pairs of officers stopped vehicles having seen signs of poor driving or vehicle condition. Individual stops could occur 'ad hoc' (i.e. as a one off occurrence due to officers having observed key signals while passing a driver) or be part of a strategy on the part of the officers and their immediate supervisors (i.e. a decision is taken that part of the shift must involve the officers standing at the roadside to pull over a series of vehicles). The nature of the encounters was largely the same in all such instances. One officer spoke to the driver of the vehicle to explain why they have been stopped and to ascertain whether a breath test might be appropriately requested (which can only be done in Scotland if the officer has reasonable suspicion the driver has been drinking). Both officers would then run

through a series of safety checks on the vehicle, asking the driver to demonstrate signalling, lights and washers, and inspecting tyres and the car body as appropriate. Where drink driving was suspected, the driver of the vehicle would be requested to give a breath test, and where safety defects were found tickets were issued to compel the driver to address them.

The experimental intervention

The intention of the ScotCET intervention was to provide road police officers with a tool for enhancing driver's perceptions of the procedural fairness of the encounters. As in the original QCET the focus was on whether changing or improving methods of communicating with drivers could achieve this. However, given concerns expressed by officers during the preliminary stages of the project about the imposition of a script for encounters in the experimental group – they felt this would be overly prescriptive, and would not mesh well with practice 'on the ground' – a different approach to that used in QCET was employed. Following discussions with Police Scotland, the research team undertook a series of meetings with road police officers of all ranks to develop an alternative intervention. The consensus was that the best way to proceed would be to devise a series of 'key messages' for officers to include in encounters. These messages were designed to communicate the core elements of procedural justice, with the intention that officers in the experiment group would be requested to ensure incorporation of *all* of these 'key messages' communicating procedural justice indicators across *all* encounters they conducted. That is, the experiment group were asked to adhere to a level of consistency in communication through all encounters, but without having to follow a rigid 'script'. Officers could retain flexibility and adapt their style of delivery according to the needs of the driver and the situation at hand. In

addition to a detailed instruction sheet explaining the key messages and the rationale behind them, card 'aide memoires' summarising the messages were also prepared for each of the officers in the experiment group to carry on shift (see Appendix 1).

From the outset, however, it was apparent that the script might not be a strong enough intervention. Recall that preliminary qualitative fieldwork suggested many officers were already using key principles of procedural justice in their practice. This meant that the experimental intervention might be weaker (because it was not so different from business as usual) and more diffuse (because it was applied in a wider variety of encounters) than was the case in the original QCET experiment, where random breath test operations were highly uniform and, in their baseline state, conducted with almost no recourse to principles of procedural justice at all (Mazerolle et al, 2012; Mazerolle et al, 2011). A weaker, more diffuse intervention is, of course, less likely to have an observable effect.

Moreover, it was recognised that some encounters would present particular challenges to achieving the incorporation of all of the key messages described above. For example where issues with drivers or vehicles are uncovered encounters might take a considerable amount of time. This, coupled with legal requirements placed upon officers to communicate several pieces of information during encounters, implied that it would not always be realistic to expect them to remember to check they had delivered all of a series of additional messages. Preliminary work with Police Scotland further suggested that drivers often reacted to being stopped by the police in very different ways than seems to have been the case for their Australian counterparts, at least as described in the QCET literature. For some at least the experience of being stopped was far less

'routine' in Scotland than in Queensland, and drivers were often observed in a worried state when interacting with officers, explaining in large part the emphasis on reassurance and rapport building during encounters expressed by many officers interviewed and observed during the preliminary stages of the project. Relying solely on relatively subtle verbal differences in communication in such a context might not be effective in differentiating control and experimental conditions.

In light of these issues a leaflet was introduced for distribution to all drivers stopped by officers in the experiment group (a leaflet was also used in the original QCET experiment, although it was rather different to the one described here). It was anticipated that the leaflet would both strengthen and standardize the experimental intervention. A high quality print leaflet was produced in collaboration between marketing and communications colleagues at the University of Edinburgh and Police Scotland, and, drawing on evidence on effective police communication (Wünsch and Hohl 2009), was intended to reinforce the verbally delivered key messages described above. In particular, the emphasis was on communicating to drivers the reasons behind the Road Safety Campaign, and thus why they had been stopped, emphasising the need to minimise the risk of harm to all those using Scotland's roads. The leaflet, that is, was intended to clearly communicate that police had the right motives in conducting the stops, and were not acting capriciously or in ways that unnecessarily took up driver's time (something they may find disrespectful). The leaflet opened by thanking drivers for their time and inviting them to contact police to share their views or seek further

information, reiterating other core components of the procedural justice model (see

Appendix 2)[1].

Implementing the study

For implementation, random assignment of encounters to experiment and control

conditions was undertaken at the unit level, such that all officers within a particular unit

were assigned to a single condition, and all stops conducted by that unit therefore fell

under the same condition. Accounting for the low 'n' of units (20), and potential bias

arising through variance in unit size, activity, historical practice, and local baseline

perception, a block randomised (matched pairs) design was employed to assign units to

experiment and control conditions (Ariel and Farrington 2010; Weisburd and Gill

2014). See the Technical Appendix for details of the pairing process. To provide a

robust experiment, such that the equivalence achieved between experiment and control

groups could be tested and baseline measures of the key constructs gathered, a pre-post

design was applied. All drivers stopped during the trial period were presented with a

four sided self-completion questionnaire and instruction sheet to capture attitudes

post-encounter. Return postage was provided and an online alternative offered in an

attempt to further encourage response.

Data and analysis

Survey responses

---

[1] Only 20 drivers receiving the experimental intervention could not recall receiving the leaflet.

Over the course of the trial 12,431 questionnaires were issued to drivers. In total 816 questionnaires were returned by the cut-off point in April 2014: 305 in the baseline (pretest) period, comprising 122 responses from the units assigned to the experimental condition and 183 from those assigned to the control condition; and 511 in the 'post' period, comprising 176 responses from the experiment condition and 335 responses from the control condition. Thus the overall response rate is 6.6 per cent. At the baseline the response rate was 7.2 per cent (6.2 per cent within the experiment group and 8.2 within the control group), and during the post period it was 6.2 per cent (5.2 per cent within the experimental group and 6.9 per cent within the control group).

This is substantially lower than the response rate achieved in the original QCET (13 per cent overall)[2]. This may be due to the timing of ScotCET. Basing the experiment around the festive road safety campaign allowed the high volume and uniformity of encounters required for a robust experiment design, but it may be that implementing the trial during a busy holiday period compromised the response rate, with time pressures combining with the relatively lengthy questionnaire to put individuals off. Moreover, the response rate dropped over the course of the trial period, again indicating that the timing of the trial and the pressures placed on people over the Christmas period may have had an impact. However, there are a number of possible reasons for surveys achieving low response rates, such as anger or disgruntlement with stakeholder institutions (like Police Scotland or the Scottish Government) or a general lack of trust in institutions of all kinds, and these potential biases must be borne in mind.

---

[2] Although in QCET, too, a lower response rate was achieved from the experimental group (Mazerolle et al, 2011, 2012).

Table 1 below shows the distribution of survey responses by road police unit. These were not evenly distributed, with 39 per cent of responses coming from just two units, A and D, both in the control condition. Moreover, the response rate varied considerably across the different units, and, again, within units across the pre and post periods of the trial. Unfortunately, with no information available on who the 'non-responders' were (see Technical Appendix for discussion on this), it is unclear why this happened.

**Insert table one here

Implications of achieved response rates

ScotCET therefore presents with a low response rate, which appears to vary over time and by road police unit. Low and varied response rates are of course problematic. Systematic non-response, whereby particular groups or sub-groups of potential respondents have consistently lower levels of participation, can lead to the introduction of bias, particularly if in an experimental context non-response is concentrated within one condition and not the other/s. This is a particular issue if it is believed the intervention itself might have an effect on response rates; a concern in the original QCET project, where a lower response rate amongst those experiencing the experimental intervention was considered a possible effect of 'irritation' felt by drivers who were stopped for longer than they might otherwise have anticipated in order to 'receive' the intervention. Drivers irritated by the intervention, and therefore more likely to express negative opinions, may have been less likely to respond to the survey, thereby inflating the positive scores in the experimental group. However subsequent analyses of the QCET data (Antrobus et al 2014) suggest that results were in fact robust against moderate distortions caused by differential non-response.

Similar calculations have not been undertaken with the ScotCET data. However its block randomized design, which should have ensured all 'types' of potential respondents and non-respondents were evenly distributed between experiment and control groups, should guard against bias arising from demographic and other factors (i.e. young people, who may be less likely to respond to surveys, should be evenly distributed between experiment and control groups – see below). Moreover, as noted above, the response rates for the experiment group were 6.2 per cent in the 'pre' period and 5.2 per cent in the 'post' period, a decline verging on statistical significance ($z=1.5$, $p=.12$). However the response rate also fell in the control group, from 8.2 per cent to 6.9 per cent, with the difference statistically significant at the 10 per cent level ($z=-1.95$; $p=.1$). The fact that response rates fell in *both* experiment *and* control sites, by around 1 percentage point, suggests that the intervention did not, in and of itself, affect response rates in the experimental group.

In sum, the low response rate seems unlikely to have a major impact on the internal validity of the experiment, which, given the design of ScotCET, depends most importantly on random assignment of the treatment (Shadish *et al.* 2002). Low response rates may provide a threat to external validity, but probably more important here is that the population itself – road users in Scotland driving over the Christmas period who were stopped by police – is distinct and quite specific. It is far from certain that other populations would respond in the same way to similar interventions, and this would be true whatever response rate was achieved.

The ScotCET sample

The ScotCET sample had the following characteristics. The majority of respondents were male (63%) and the mean age was 50.7 (SD=14.8, min=17 years, max=87 years). Three quarters (75%) were owner-occupiers, and 40 per cent had a first degree or higher, while 12 per cent reported they had no qualification. Seventy one per cent were in employment, 21 per cent were retired; and 73 per cent were married or in a de facto married relationship.

Table 2 below shows the demographic breakdown of the sample of drivers responding to the survey. Crucially, there was no significant difference pre and post, or between experimental or control groups, on any of these measures, suggesting that the matched pairs approach was successful in producing balanced experiment and control groups.

**Insert table two here

Outcome measures

As per the hypotheses outlined above, the key latent concepts the questionnaire sought to capture are: perceptions of officer adherence to procedural justice; trust in the police during the encounter; satisfaction with encounter and police ; general trust in the fairness of the police; general trust in the effectiveness of the police; and police legitimacy. In view of ongoing conceptual debate around the constituent 'elements' of legitimacy, the survey was designed to capture a dual component concept of legitimacy: respondents' sense of duty toward the police, as legitimate authority commands obedience from those subject to it, and respondents' assessment of the extent to which police operate according to a general, shared moral framework, as authority is

legitimate when it is applied in a manner congruent with shared norms and values

(Jackson et al 2012, 2013).

In this paper we present results in relation to these latent variables, rather than individual indicators, since they provide better, more robust, measures of the underlying constructs of interest. Full details of the confirmatory factor analysis used to assess the scaling properties of the individual indicators, their ability to capture the underlying latent constructs, and their empirical 'distinctness' from one another are provided in the technical appendix to this paper. Results from this analysis suggested a seven-factor solution fitted the data well, and we proceeded with these seven as our response variables.

An important initial finding with respect to each of the key encounter outcome measures described above is that the responses to all were overwhelmingly positive. For each indicator of the latent variables derived above, at least 80% of respondents were located on the positive end of the scales. Appendix Table 3 shows the sample distributions for each observed indictor used in the three encounter outcome measures. The positive skew observed is not a surprising finding. The Scottish Crime and Justice Survey fields a series of questions in each sweep on public confidence and experience of the police and consistently records high levels of positive responses (Scottish Government 2014).

Analytical technique

To compare the difference between reported perceptions amongst drivers in the experimental group and those in the control group, a 'difference in differences' approach was used with the unit level randomization. Use of a matched pairs design in

allocation also had to be taken into account in case organizational and geographical factors had influenced the outcomes observed (Boruch et al 2010). Accordingly, random effects linear regression models predicting outcomes on each of the key constructs were estimated in Stata 12.1 (Rabe-Hesketh and Skrondal 2008; Snijders and Bosker 2012).

Four coefficients are shown in each model below.[3] Coefficients in the rows marked 'Baseline period' (1) show the difference between the experiment and control groups at the baseline (i.e. during the 'pre' period before the experimental intervention was implemented). This will, ideally, be non-significant, since a significant coefficient here would indicate that there was a systematic difference between experiment and control sites in relation to the indicator in question. The second coefficient, in the rows marked 'Control areas' (2), shows the difference between the pre- and post-periods in the *control* areas, that is, the pretest-posttest change in driver assessments of each aspect of police behaviour in the areas that *did not* receive the experimental intervention. The third coefficient, in the rows marked 'Experiment areas' (3) shows the pretest-posttest change in the *experiment* areas; the change in driver assessments about each aspect of police behaviour in the areas that *did* receive the experimental interventions. The final coefficient presented in the rows marked 'Difference in differences' (4), indicates the change in perceptions (from pre to post period) within the experiment areas *relative to* the control areas. It is this coefficient that provides the test of the hypothesis that the experimental intervention enhanced perceptions of police. A positive, significant, coefficient here would mean that the experimental intervention resulted in improved assessments of police behaviour the experiment sites compared to the control sites. The

---

[3] Coefficients 1, 2 and 4 were taken directly from each model. Coefficient 3, along with its standard error and significance level, was calculated using the lincom function in Stata.

value for coefficient (4) is simply (3) minus (2) – it represents the difference between change in the experiment areas and change in the control areas.

The tables also show the Intra-class correlation (ICC), which indicates how much variance in the response variable is explained at the level of the matched pair; that is, at a geographical or an administrative level separate from that of the individual encounters represented in the data. Relatively high ICCs, which in this context might mean greater than about .05, could indicate that perceptions of encounters within the pairs were being systematically affected by, for example, the nature of the driving or traffic in those areas, underlying levels of public confidence in police, or different 'ways' of doing policing. Because this potential level 2 variation is taken into account in the models, however, it will not bias the results shown.

Results

Table 3 presents three models assessing the constructs relating to respondents judgements about the police involved in the Festive Road Safety Campaign encounter itself. We find that in the control areas there was a consistent pattern of improvement from pretest to posttest. Yet, this pattern was not repeated in the experiment areas, and, most importantly, the difference in differences coefficients indicate that judgements of the procedural justice of the encounter and overall satisfaction *fell* in the experiment areas relative to the control areas (effects significant at the 10 per cent level).

**Insert table three here

A brief examination of the contextual information gathered in the questionnaire that may help to explain this finding reveals no significant difference between experiment and control groups. People in the experiment areas were no more likely to think they had been stopped for the 'wrong' reason, nor where they more likely to have been breathalyzed. It therefore seems that something in the experimental intervention *damaged* the perceived procedural justice of officer's actions, and this effect dampened, in the experiment areas, the improvement in opinions that occurred in the control areas. That is, since the only systematic difference between the experiment and control areas was the experimental intervention it would be expected that, if the intervention had no effect, change observed in the control areas would also have been observed in the experiment areas. Since this is not what is observed, it can be concluded with some certainty that something in the experimental intervention *stopped* it from occurring.

Turning to general perceptions of police, Table 4 shows results from random effects linear regression models predicting trust in police fairness and effectiveness. The key finding here is that none of the coefficients achieve statistical significance, suggesting experimental intervention had little effect on trust in the fairness or effectiveness of the police in a general sense (note, however, that both difference in differences coefficients are negative).

**Insert table four here

Finally, Table 5 presents the models predicting the separate components of police legitimacy. Once again, the experimental intervention seems to have had little effect on perception, with no significant differences occurring between experiment and control

groups, or in the pre and post periods; although, again, both difference in differences coefficients are negative.

**Insert table five here

Additional analysis

Recall that the individual survey items relating to respondents' assessments of the officers conducting the stops were generally very positive. This resulted in significant skew to the latent variables representing these assessments. To check that heteroscedasticity was not affecting our results we estimated a second set of random effects models, this time with the individual survey items used to measure the latent constructs described above as the response variables. In each case these were dichotomized into positive/negative responses. Results closely matched those described above, with negative 'difference in differences' regression coefficients found for many of the individual items relating to respondents' assessments of the encounters (although, when considered individually, none achieved statistical significance at conventional levels). Results from this analysis are available from the lead author.

Discussion and conclusion

To return to the hypotheses raised above, H1a, first, stipulated that the experimental intervention would increase feelings of procedural justice during the stop among those in the experiment group. No evidence in support of this hypothesis was found; instead, the evidence suggests the intervention lead to *diminished* feelings of procedural justice.

24

No evidence was found that the intervention increased trust in the officers conducting the stop (H1b). Second, H2 stipulated that the experimental intervention would increase overall satisfaction with the encounter. Again, no evidence in support of this hypothesis was found; instead, again, the intervention actually appeared to *decrease* overall satisfaction with the stop.

Turning to H3, which stated that the experimental intervention would result in higher levels of trust and confidence in the police among those who experienced it, no evidence in favour of this hypothesis was found. Finally, H4 addressed the issue of legitimacy, suggesting that the experimental intervention would enhance the legitimacy of the police among those who experienced it. Again, no strong evidence in favour of this hypothesis was found. It is noteworthy that the 'difference in differences' coefficients in *all seven models* were negative, adding some further credence to the idea that the intervention had, on average, a consistently negative effect.

The experimental intervention therefore seems to have had a significant and detrimental effect on drivers' impressions of the officers they encountered and their satisfaction with that particular experience. At the same time there were apparently 'naturally' occurring *improvements* observed within the control group. These results are, it is fair to say, unexpected. The experimental intervention was designed in line with existing evidence on procedurally just modes of policing and effective police-public communication, incorporating the fundamental elements of the procedural justice model: treating drivers with dignity and respect; demonstrating the neutrality of decision making and the trustworthy motives of the officers; and presenting drivers with the opportunity to voice their concerns, questions or otherwise and be an active

participant both during and after the encounter (Tyler, 2006; Tyler and Huo, 2002; Murphy, 2008). The design was led by previous successful experimental intervention in the field (Mazerolle et al, 2011; 2012), and, moreover, those police officers responsible for implementing the experimental intervention here were key contributors. Drawing on extensive collective experience of policing and interaction with the public, officers devised key messages and shaped the ways in which these ought to be communicated. For this to have had a detrimental effect on perceptions of procedural justice and satisfaction is surprising.

Considering the results achieved in the original QCET, and comparing the difference between the outcomes achieved there and here, it is prudent to bear in mind the very different starting points of each study. While it was anticipated that the positive results achieved in QCET would be replicated in ScotCET, albeit under quite different field conditions, the baseline and contexts of the two studies may have had important effects. Although road policing-specific public opinion data was not available to provide a Scottish baseline, sufficient evidence existed to suggest that perceptions of the police in Scotland, including roads police, are broadly favourable. The baseline data gathered in the ScotCET pre-period conform to this idea, with univariate analyses of key items demonstrating judgements of the police officers encountered and the police in general were overwhelmingly positive. Observations conducted during the study suggested police officers were attuned to, and delivering, the procedural justice model, or at least elements of it, and as such these positive driver assessments were to have been anticipated.

This contrasts with the experience in Queensland, where in the context of RBT operations interaction between police and drivers was limited and public opinion of the police considerably less positive (Murphy et al, 2014; Mazerolle et al, 2012). In the Australian context, the 'small dose' of communication on procedural justice or fairness represented a real shift from business as usual and was sufficient to shift the judgement of drivers in a significantly positive direction. Within Scotland, it is arguably more difficult to achieve such an effect. We are faced with a very different policing context, necessitating the utilization of a more diffuse experimental intervention design, albeit distinct in its adoption of a more 'complete' procedural justice oriented model of policing. Moreover, much previous research suggests it is difficult for police 'on the ground' to achieve a positive impact on public perception (i.e. to increase positivity of attitudes or improve on judgements previously made) through interaction and contact (Skogan, 2006; Bradford et al 2009; although see Myhill and Bradford 2012). When people are already generally positive about police, improving opinions further may be even more challenging, and, arguably, less meaningful as an endeavor. Nevertheless, the ScotCET control units appear to have positively influenced some elements of public perception over the course of the trial period, demonstrating a pattern that, we anticipate, would have occurred within the experiment group in the absence of the experimental intervention.

So why did the experimental intervention lead to a negative effect on perceptions of procedural justice and driver satisfaction? Further research is needed to answer this question. There are a number of possible interpretations of the findings observed but, as yet, nothing in the data and analysis can explain why the experimental intervention led to the effect it did. That said, it seems reasonable to speculate that the intervention may

have had a detrimental impact on the way officers conducted their vehicle stops, leading to more negative public perception. Surmising from previous experiences and the existing literature, there are a range of potential reasons for such an impact having occurred that merit further exploration. Perhaps, as in QCET, the delivery of key messages and distribution of the leaflet had the effect of lengthening encounters. Mazerolle et al (2014a) suggest there is a 'sweet spot' for encounter length. Too short an encounter precludes the incorporation of elements of the procedural justice model, but if drivers are engaged for too long any positive impact of the procedural justice model began to diminish. It appears that, perhaps unsurprisingly, delaying drivers unnecessarily results in negative reaction or 'backfire' effect. It may be that drivers in our experiment group were simply held up too long by officers delivering the intervention. The varying length of encounters and the fact that participants completed the ScotCET survey after the event made it problematic to ask them how long the encounter took. In retrospect, though, it might have been better to include a survey item on this issue despite the potential issues with respondent recall. Arguably it is how long the respondent felt or perceived they had been held up, rather than an objective measure of encounter length, that is important here and future studies may usefully include a combination of objective and subjective measures on this to assess any potential effect.

The intervention may also have resulted in more scripted and 'bureaucratized' interaction between police officers and drivers. While strong emphasis was placed on officers retaining their natural flow and flexibility of communication during the stops, it may have been that the pressures of remembering the additional messages and tasks required by the experiment led to officers reverting to the aide memoires and scripting their delivery to ensure nothing was forgotten. Based on the results obtained in the

original QCET, much of the related literature has stressed the potential value of scripts in enhancing both the adoption of procedurally just policing practice and citizen perception (Mazerolle et al 2013a; Mazerolle et al 2014b). It is argued that the addition of dialogue communicating procedural justice, no matter how brief or 'complete' in terms of addressing the 'full' procedural justice model, will have a positive effect. However, if this were the case, we would expect that the introduction of our key messages, whether delivered as a scripted encounter or in the intended more natural, responsive manner, would have at best enhanced perceptions within the experiment group and at worst had no effect at all. That we have achieved a negative effect suggests there is more to consider here. A script may contain the all of the appropriate 'ingredients' for a procedurally just encounter, but it appears that encounters where this was not provided for, where key ingredients would have been missed and excluded, could and did fare better. The implication is that dialogue alone is not enough.

At this stage it seems reasonable to suggest that judgements of police officers' fairness, helpfulness, professional capability, respectfulness and personal demeanour may hinge on the qualities of interaction not captured or provided for by verbal scripts. The manner in which a script is delivered is likely to be key in informing the judgement process, such that it is not just the mere presence or inclusion of dialogue, but the *quality* of the dialogue and the skill of the individual delivering it that are important. Perhaps if the post-test period had lasted longer, delivery of key messages and materials would have been more 'naturalized' in practice and scores across the experiment group may have improved. Future research could helpfully examine, ideally on a longitudinal basis, the impact of scripts being imposed on officers in contexts where verbal interaction is already a core

element of policing practice, both in terms of their adoption of the values the script embodies and their delivery style.

Similarly, there may have been something in the leaflet, or something about the way in which the leaflet was presented, that led to the differences observed, albeit that this notion is tempered by the fact that views among the 11 per cent (n=20) of respondents in the experiment group who could not recall being given a leaflet tended to be even more negative than those who could recall the leaflet. Nevertheless, the leaflet represents another form of procedural justice 'script' and should be considered as a potential factor in the results achieved. The effectiveness of the leaflet as such a script is open to question. Similar experimental studies could helpfully explore this issue, building in the opportunity to capture respondent assessment of written messages and the ways these may or may not shape judgements on policing encounters (see Hohl et al. 2010 for one earlier study investigating these issues).

The results obtained may also be attributable to officer effects. There may have been some unintended effect of assignment to the experiment group that influenced the ways in which officers conducted stops (they may have felt pressure to 'perform' in some way, for example). Research exploring the impact of participation in experimental policing studies would provide useful methodological and substantive learning. It might also be that the 'mix' of police conducting the stops may have changed over the course of the trial period, with officers coming on roster or going on leave who may have been more or less inclined to use procedurally just practice. However this, at least, seems unlikely, as annual leave was not allowed over most of the period in question (although this does not preclude illness or other factors having some small effect).

This discussion points to some of the key limitations of ScotCET. Crucially, the experimental design and survey instrument did not include measures needed to address the questions raised above. Moreover, the survey is unlikely to be representative of all drivers in Scotland; while ScotCET has a relatively high level of internal validity and provides a robust set of experimental results, we cannot claim that the results described are generalizable to other groups of drivers, or to encounters occurring in other police settings. The intervention fielded in ScotCET was also itself problematic, since it comprised two separate elements, the 'checklist' provided to officers and the leaflet. The design of the experiment did not allow determination of whether it was one or other of these, or both together that caused the observed effects. Future research might profitably decompose these or similar elements into separate interventions and examine their independent effects on the outcomes of interest.

Notwithstanding these shortcomings, ScotCET has delivered some interesting and challenging findings, and points to the importance of, as well as some of the difficulties with, the replication of experimental studies across different policing contexts. Procedural justice theory, as detailed at the outset, provides us with a framework through which to understand the process of perception and reaction by citizens to authority exercised by legal actors. In conducting this experimental study, we have contributed to a growing field of literature exploring how the police might utilize such a framework to maintain or enhance their legitimacy. This is critical at a time when, in Scotland and beyond, procedural justice theory is rapidly developing into a model of policing and practice. Crucially, we highlight that the implementation of a procedural justice model of policing is not a straightforward matter. In as much as QCET demonstrated the apparent

ease with which such a model could be incorporated into practice, leading its authors espouse that even the shortest encounters between citizens and police can provide 'gain' (Mazerolle et al 2013a, 2013b), our data suggest that, at least in policing contexts where interaction and satisfaction are already high, it is not enough to verbally 'add in' the various components of the procedural justice model and up the 'dosage' administered. Life is not quite that simple. Indeed, such an approach in the present context has in fact led to losses, albeit small and, for the time being, confined to perceptions of specific encounters rather than 'global' perceptions of the police (Brandl et al 1994).

It appears that subtleties and nuances of communication context, content and style may be important but, as yet, under-developed elements of delivering policing that both is, and is perceived to be, procedurally just. Failure to acknowledge and provide for these in attempting to operationalise the procedural justice model appears to have led to unintended detrimental effects on public perception that, if adopted on a broader basis, could undermine public trust and police legitimacy. As such, the work begun here must continue. Procedural justice theory does not in and of itself provide a guide to effective and appropriate policing practice; nor does the literature available to date. Simplistic approaches focusing on the content of dialogue and written messages may create losses rather than gains. Further empirical or experimental research must seek to establish exactly *why* and *how* these observed effects might come about, and what further critical elements of communication and interpersonal skill might be required to implement procedurally just policing.

References

Antrobus, E., Elffers, H., White, G. and Mazerolle, L. (2014) 'Non-response bias in randomized control experiments: Putting the Queensland Community Engagement Trial (QCET) under a microscope' *Evaluation Review* 37(3-4) 197-212

Ariel, B. and Farrington, D.P. (2010) 'Randomised block designs' in Piquero, A.R. and Weisburd, D. (eds) (2010) The Handbook of Quantitative Criminology, New York: Springer

Boruch, R., Weisburd, D. and Berk, R. (2010) '*Place Randomised Trials'* in Piquero, A.R. and Weisburd, D. (eds) (2010) The Handbook of Quantitative Criminology, New York: Springer

Bradford, B., Jackson, J. and Stanko, E.A. (2009) 'Contact and Confidence: Revisiting the Impact of Public Encounters with the Police' in *Policing and Society* 19(1) pp.20-46

Bradford, B., Huq, A., Jackson, J. and Roberts, B. (2014a) 'What price fairness when security is at stake? Police legitimacy in South Africa' in *Regulation and Governance* 8(2) pp.246-268

Bradford, B., Murphy, K. and Jackson, J. (2014b) 'Officers as mirrors: Policing, procedural justice and the (re)production of social identity' in *British Journal of Criminology* Advance Access.

Brandl, S., Frank, J., Worden, R. and Bynum, T. (1994) 'Global and specific attitudes toward the police: Disentangling the relationship' *Justice Quarterly* 11(1) pp.119-134

Brown, T.A. (2006) *Confirmatory Factor Analysis for Applied Research*, New York: The Guilford Press

Hinds, L. and Murphy, K. (2007), 'Public Satisfaction with Police: Using Procedural Justice to Improve Police Legitimacy' in Australian and New Zealand Journal of Criminology 40 pp.27–42.

Hohl, K., Bradford, B. and Stanko, E.A. (2010). Influencing trust and confidence in the London Metropolitan Police: results from an experiment testing the effect of leaflet-drops on public opinion, *British Journal of Criminology* 50(3): 491-513.

Hough, M., Jackson, J. and Bradford, B. (2013) 'The governance of criminal justice, legitimacy and trust' in Body-Gendrot, S., Lévy, R., Hough, M., Snacken, S. and Kerezsi, K. (eds) *The Routledge Handbook of European Criminology*, Oxford: Routledge

Jackson, J., Bradford, B., Hough, M., Myhill, A., Quinton, P., & Tyler, T. R. (2012) 'Why do People Comply with the Law?: Legitimacy and the Influence of Legal Institutions' in *British Journal of Criminology 52*(6) pp.1051–1071

Jackson, J., Bradford, B., Stanko, B. and Hohl, K. (2013) *Just Authority? Trust in the police in England and Wales*, London: Routledge. Lawrence

Jonathan-Zamir, T. and Weisburd, D. 2013 'The effects of security threats on antecedents of police legitimacy: Findings from a quasi-experiment in Israel', *Journal of Research in Crime and Delinquency* 50(1): 3-32 .

Mazerolle, L., Bates, L., Bennett, S., White, G., Ferris, J. and Antrobus, E. (2014a) 'Optimising the length of random breath tests: Results from the Queensland Community Engagement Trial' in *Australian and New Zealand Journal of Criminology* DOI: 10.1177/0004865814532661

Mazerolle, L., Sargeant, E., Cherney, A., Bennett, S., Murphy, K., Antrobus, E. and Martin, P. (2014b) *Procedural Justice and Legitimacy in Policing*, Switzerland: Springer International Publishing

Mazerolle, L., Bennett, S., Davis, J., Sargeant, E. and Manning, M. (2013a) 'Procedural justice and police legitimacy: a systematic review of the evidence' in *Journal of Experimental Criminology* 9(3) pp.245-274

Mazerolle, L., Antrobus, E., Bennett, S. and Tyler, T. (2013b) 'Shaping citizen perceptions of police legitimacy: A randomized field trial of procedural justice' *Criminology* 51(1) pp.33-64

Mazerolle, L., Bennett, S., Antrobus, E. and Eggins, E. (2012) 'Procedural justice, routine encounters and citizen perceptions of the police: Main findings from the Queensland Community Engagement Trial (QCET)' *Journal of Experimental Criminology* 8(4) pp.343-366

Mazerolle, L., Bennett, S., Eggins, E., Antrobus, E., White, G. and Davis, J. (2011) *Testing police legitimacy…one breath at a time: The Queensland Community Engagement Trial Technical Report*, Brisbane: ARC Centre of Excellence in Policing and Security

Myhill, A., & Bradford, B. (2012) 'Can police enhance public confidence by improving quality of service? Results from two surveys in England and Wales' in *Policing and Society 22*(4) pp.397–425

Murphy, K., Mazerolle, L. and Bennett, S. (2014) 'Promoting trust in police: findings from a randomised experimental field trial of procedural justice policing' in *Policing and Society* 24(4) pp.405-424

Murphy, K., Hinds, L. and Fleming, J. (2008), 'Encouraging Cooperation and Public Support for Police' in *Policing and Society* 18 pp.138–57

Papachristos, A., Meares, T. and Fagan, J. (2012), 'Why do Criminals Obey the Law? The Influence of Legitimacy and Social Networks on Active Gun Offenders' in *Journal of Criminal Law and Criminology* 102 pp.397–439

Rabe-Hesketh, S. and and Skrondal A. (2008) *Multi-level Modelling Using Stata* (2nd edition), College Station, TX: Stata Press

Skogan, W.G. (2006) 'Asymmetry in the Impact of Encounters with the Police' in *Policing and Society* 16(2) pp.99–126.

Scottish Government (2014) *Scottish Crime and Justice Survey 2012-13: Main Findings*, Edinburgh: Scottish Government.

Scottish Government (2012) *The Strategy for Justice in Scotland*, Edinburgh: Scottish Government.

Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002) *Experimental and Quasi-experimental Designs for Generalized Causal Inference,* Belmont: Wadsworth Cengage Learning

Snijders, T. and Bosker, R. (2012) *Multilevel Analysis, 2nd edition*, London: Sage

Sunshine, J. and Tyler,T. (2003b) 'The Role of Procedural Justice and Legitimacy in Public Support for Policing' in *Law and Society Review* 37(3) pp.513–48

Tankebe, J (2008) 'Police effectiveness and police trustworthiness in Ghana: An empirical appraisal' in *Criminology and Criminal Justice: An International Journal* 8 pp.185–202

Thibaut, J., & Walker, L. (1975) *Procedural justice: A psychological analysis*, Hillsdale: Erlbaum

Tyler, T.R. (2006) *Why People Obey the Law*, Princeton: Princeton University Press

Tyler, T. R. and Huo, Y. J. (2002) *Trust in the Law: Encouraging Public Cooperation with the Police and Courts,* New York: Russell Sage Foundation

Tyler, T.R and Wakslak, C.J (2004) 'Profiling and police legitimacy: Procedural justice, attributions of motive and acceptance of police authority' in *Criminology* 42(2) pp.253-281

Weisburd, D.L. and Gill, C. (2014) 'Block randomised trials at places: Rethinking the limitations of small N experiments' in *Journal of Quantitative Criminology* 30(1) pp.97-112

Wünsch, D. and Hohl, K. (2009) 'Evidencing a "good practice model" of police communication: The impact of local policing newsletters on public confidence' *Policing: A Journal of Policy and Practice* 3(4) pp.331-339

Figure 1: ScotCET conceptual map
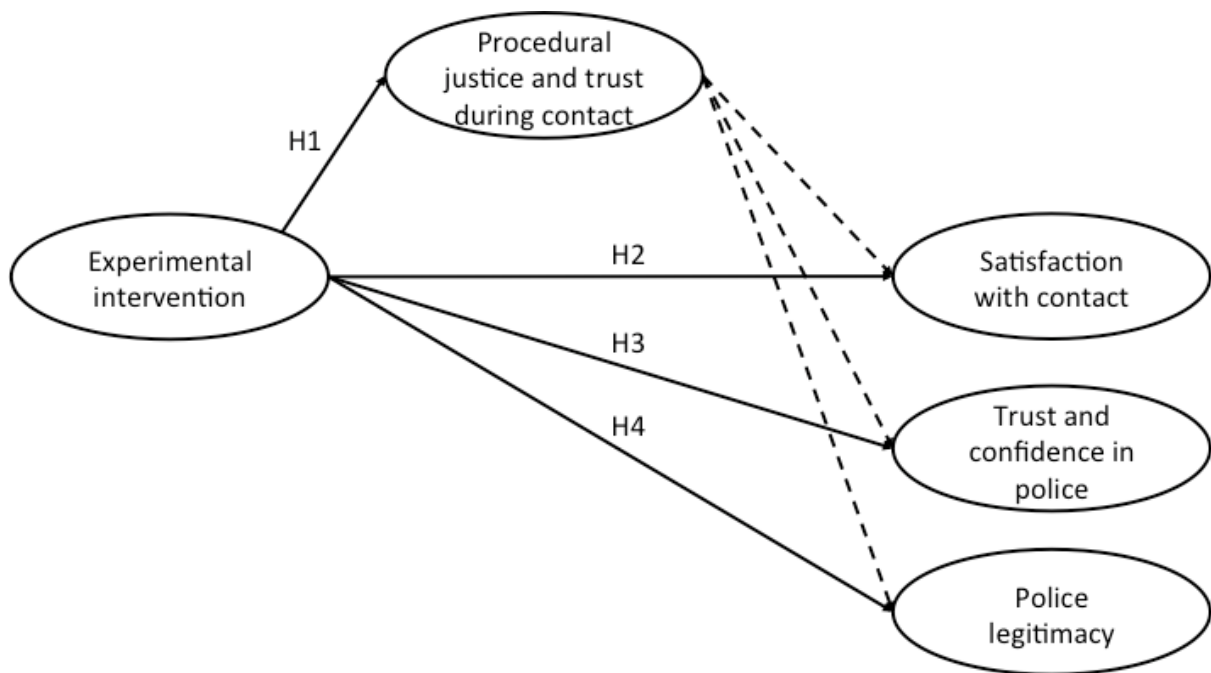
Table 1: Response rates across the RPUs[4]

	Pre-test	Test	Pre-control	Control	% of total responses
A/B DRPU			15.5%	13.2%	20%
D DRPU			13.4%	12.7%	19%
M/C West			10.9%	11.4%	7%
C DRPU			9.3%	6.2%	8%
J DRPU			4.0%	1.1%	1%
U DRPU			2.9%	3.7%	2%
E DRPU			2.8%	-	< 1%
Glasgow/ Irvine TRPG			8.6%	2.5%	3%
K DRPU			-	1.2%	1%
V DRPU	41.7%	11.8%			6%
Perth/ Stonehaven TRPG	15.6%	5.6%			9%
L DRPU	9.6%	6.8%			4%
Edinburgh TRPG	3.9%	6.5%			1%
G DRPU	2.9%	5.0%			4%
N DRPU	2.6%	7.7%			3%
Q DRPU	2.2%	7.4%			4%
P DRPU	1.4%	1.6%			2%
M/C East	-	2.0%			<1%

---

[4] Separate response rates for the Motherwell and Lockerbie TRPG units could not be calculated due to administrative error. The leftover questionnaires were combined into a single bundle before a count was undertaken for each individual unit. The combined Motherwell and Lockerbie TRPG response rates are 9.8% for the 'pre' period and 2.3% for the 'post' period.

Table 2: Demographic breakdown of sample (% of respondents)

	'Pre' survey		'Post' survey		
	Experiment	*Control*	*Experiment*	*Control*	**Total**
Age (years)					
17-24	*4*	*4*	*6*	*6*	*5*
25-34	*11*	*9*	*8*	*10*	*10*
35-44	*16*	*16*	*19*	*15*	*16*
45-54	*33*	*28*	*28*	*22*	*26*
55-64	*14*	*23*	*21*	*21*	*21*
65-74	*15*	*14*	*10*	*16*	*14*
75+	*7*	*6*	*7*	*9*	*8*
Gender					
Male	*68*	*64*	*62*	*60*	*63*
Female	*32*	*36*	*38*	*40*	*37*
Housing tenure					
Owner occupier	*76*	*78*	*74*	*78*	*77*
Renter	*18*	*15*	*20*	*15*	*17*
Other	*6*	*8*	*6*	*7*	*7*
Marital status					
Married or de facto	*77*	*77*	*71*	*71*	*73*
Single	*23*	*23*	*29*	*29*	*27*
Employment					
Employed	*71*	*70*	*74*	*70*	*71*
Retired	*20*	*21*	*20*	*22*	*21*
Other	*9*	*9*	*6*	*8*	*8*
Educational attainment					
Degree or higher	*38*	*36*	*44*	*44*	*41*
HNC/HND/SVQ/A level	*29*	*38*	*32*	*28*	*31*
O grade/Standard grade	*19*	*15*	*11*	*16*	*15*
No qualifications	*13*	*11*	*13*	*12*	*12*
Total (n)	*122*	*183*	*176*	*335*	*816*

Table 3: Results from linear random effects models predicting assessments of the encounters

	Procedural justice		Trust		Satisfaction	
	ß	se(ß)	ß	se(ß)	ß	se(ß)
Baseline period						
Test area (ref: control area)	-0.04	-0.09	-0.05	-0.09	-0.04	-0.09
Control areas						
Post period (ref: pre period)	0.14*	-0.07	0.13+	-0.07	0.12+	-0.07
Test areas						
Post period (ref: pre period)	-0.04	0.09	-0.02	0.09	-0.09	0.09
Difference in differences	-0.19+	-0.11	-0.15	-0.12	-0.21+	-0.11
ICC	0		0		0	
n	814		814		814	

* p<.05; + p<.1

Table 4: Results from linear random effects models predicting general trust in the police

	Trust in police fairness		Trust in police effectiveness	
	ß	se(ß)	ß	se(ß)
Baseline periond				
Test area (ref: control area)	-0.09	-0.09	-0.02	-0.09
Control areas				
Post period (ref: pre period)	0.06	-0.07	0.06	-0.07
Test areas				
Post period (ref: pre period)	-0.01	0.09	-0.11	0.09
Difference in differences	-0.01	0.09	-0.09	0.09
ICC	0		0	
N	814		814	

Table 5: Results from linear random effects models predicting police legitimacy

	Duty to obey		Moral alignment	
	ß	se(ß)	ß	se(ß)
Baseline periond				
Test area (ref: control area)	-0.10	-0.09	-0.05	-0.09
Control areas				
Post period (ref: pre period)	0.01	-0.07	0.06	-0.07
Test areas				
Post period (ref: pre period)	-0.09	0.09	-0.04	0.09
Difference in differences	-0.10	-0.11	-0.10	-0.11
ICC	0.05		0.03	
N	814		814	

Appendix 1 – ScotCET Aide Memoire

ScotCET - requirements during Festive Road Safety Campaign
Introduce yourself by name and/or home location
Explain why you have stopped the person: • Make sure to mention the festive road safety campaign • Tell them over 20,000 drivers are stopped on Scotland's roads every month throughout the year • Explain <u>specifically</u> why you stopped them and what you are about to do.
Tests run as normal
If safety checks reveal no problems: • Say thanks, and sorry for taking up their time • Re-iterate importance of campaign for maintaining road safety.
If safety check reveals a vehicle safety defect: • Still say thanks • Introduce the legal wording – don't just launch into it. • Make sure to close with "Do you understand everything I/we have told you just now? "
If it is deemed necessary to breathalyse motorist: • [*Provide reason for requiring breath test*]. • Mention that drink driving is a major problem, and that average 170 people killed or seriously injured by drunk drivers in Scotland each year. • Ensure breath test requirement and subsequent processes are as per Police Scotland SOP
• Hand out leaflet and explain what it is. • Hand out survey.
• Check if driver has any questions/concerns
• Close by thanking driver for their time.

Appendix 2 – ScotCET Leaflet

POLICE SCOTLAND
Keeping people safe

Thank you for your time today. We welcome your views on all aspects of policing in Scotland. Details of our latest initiatives, local policing teams and how to contact us can be found at: www.scotland.police.uk

FESTIVE ROAD SAFETY CAMPAIGN 2013/14

Police Scotland is committed to keeping people safe on the roads. We seek to influence driver behaviour in positive ways and reduce the number of casualties by helping to prevent road accidents. While reported casualty figures are the lowest for over 60 years there is much more work to be done.

Last year, 174 people lost their lives in traffic collisions in Scotland, and a further 1,974 were seriously injured. Road stops and vehicle checks, such as the one you experienced today, are vital in ensuring that the roads are kept as safe as possible for everyone. Most people use Scotland's roads safely and legally but there are still too many who choose not to.

Every week an average of 1,700 drivers are caught speeding and a further 120 arrested for driving under the influence of drink or drugs. Highly visible patrolling is a key element of our road safety strategy. One of the most effective ways to influence road user behaviour is to make clear that the risk of being caught is high and that the consequences can be serious. Yet we also rely upon the public's support. We need you to report poor driving behaviour or illegal drivers such as those who are disqualified, drunk/ drugged, or uninsured.

You can help us keep people safe by calling our national non-emergency telephone number 101 or Crimestoppers Scotland on 0800 555 111.

Appendix 3

Procedural justice during the encounters

Thinking specifically about the Festive Road Safety Encounter, do you think the police were:

Percentages

	No, not at all	No, not really	Yes, to some extent	Yes, completely	Total (=100%)
Approachable and friendly	1	3	12	85	797
Helpful	1	4	15	80	747
Respectful	1	1	10	87	757
Professional	1	1	9	89	758
Fair	1	2	11	86	747
Clear in explaining why you had been stopped	1	3	9	87	764

Trust during the encounters

Please indicate how much you agree or disagree with the following statements about your experience:

Percentages

	Strong disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Total (=100%)
I trusted the intentions of the police officer involved	1	2	6	30	62	786
I was confident the police officer was doing the right thing	1	2	7	30	60	768
The police officer gave me the opportunity to express my views	2	4	22	29	44	738
The police officer listened to what I had to say	1	3	16	32	48	740

Satisfaction with the encounter

Please indicate how satisfied or dissatisfied you were with the following aspects of your experience:

Percentages

	Very dissatisfied	Quite dissatisfied	Neither	Quite satisfied	Very satisfied	Total (=100%)
The way the police officer conducted the stop	1	2	8	23	66	797
The way you were treated	1	2	5	19	73	789
The final action taken by the police officer(s)	2	2	6	17	73	786

Enhancing public trust and police legitimacy during road traffic encounters: Results from a randomized controlled trial in Scotland – Technical Appendix

The Scottish Community Engagement Trial (ScotCET) was originally conceived as a replication of the Queensland Community Engagement Trial (QCET). However, key contextual differences (see main article) meant that a direct replication of the design and method applied in QCET was neither feasible nor desirable. This technical appendix is intended to support an article discussing the main findings from the final study (MacQueen and Bradford 2015, this issue) by providing a detailed and transparent account of the final design, method and analytical approaches applied in ScotCET. We outline our approach to assignment to experiment and control conditions; the adoption of a pre-post design; the administration of our survey instrument; the key outcome indicators utilised within the survey, as well as the results of confirmatory factor analyses assessing the scaling properties of these indicators and the extent to which they captured the latent constructs in question.

Assignment to experiment and control conditions

As per the main article, random assignment of encounters to experiment and control conditions was undertaken at the unit level, such that all officers within a particular unit were assigned to a single condition, and all stops conducted by that unit therefore fell under the same condition. This reduced the risk of contamination or treatment migration that assignment at the level of operational pair, individual officer or individual encounter might incur, and lessened the burden on officers and the Inspectors responsible for managing each of the road policing units. However, using simple or 'naïve' random assignment (Weisburd and Gill 2014) to allocate units to conditions would have been problematic for two reasons. The different sizes of the individual units, in terms of officer numbers and geographical coverage meant that the volume of stop activity varied from unit to unit. Uneven distribution of stop activity across the different

units might introduce the risk of an unbalanced volume of encounters between the experiment and control groups. Furthermore, the recent creation of the Police Scotland, and the feedback from officers at the study outset, suggested that practice in local units would likely be influenced by 'historical' practice within legacy forces. There is also evidence from the Scottish Crime and Justice Survey (SCJS) that opinion of the police varies significantly across Scotland, suggesting that baseline levels of trust in police may vary significantly from area to area; this, possibly, could have influenced the way the intervention was received.

To address the risk that simple random assignment would result in an imbalanced distribution of these characteristics across the experiment and control groups, a block randomised (matched pairs) design was employed to assign units to experimental and control conditions. As in other examples of experimental policing research (Braga et al. 1999; Weisburd and Green 1995) the roads policing units were 'matched' into ten pairs on the basis of four key characteristics: geographical proximity and unit relation to legacy forces; volume of encounters; similarity of procedures and policing focus; and public perceptions (using the most up-to-date, reliable data from the SCJS). The primary factor was the location of each unit geographically in relation to the legacy police forces, but other important considerations included the unique focus of particular units, for example specialist motorcycle units or those sharing a focus on main arterial roads (which were all paired with another unit with the same specialism). Within each pair, units were randomly assigned to experimental or control conditions. By ensuring greater equivalence of the experimental and control groups, this process minimises the risk of bias resulting from chance assignment of units with similar characteristics to one or other condition (Ariel and Farrington 2010). Moreover, this technique has been demonstrated to improve the statistical power of studies where the overall 'n' of units is comparatively low (Weisburd and Gill 2014).

Pre-post design

In the first stage of the experiment, before administration of the intervention, a baseline measure of the key constructs or concepts of interest was achieved across both experiment and control groups. This allowed the initial equivalence (or non-equivalence) of the experiment and control groups to be established. During the first week of the trial (the pretest period) all units operated 'business as usual' during encounters, with the only difference from normal being the distribution of a questionnaire to drivers at the end of each encounter. A break of approximately one week followed, during which the experimental materials were distributed to units in the experimental condition and verbal and written briefings were delivered to Inspectors and officers. In the final weeks of the campaign (the 'post' period), half of the units operated under experimental conditions while the others continued business as usual. All units continued to distribute the study questionnaire.

Ensuring experiment fidelity

The nature of road policing in Scotland (for example its geographical spread), and the limited resources of the small research team, meant that observation of all of the officers delivering the experimental intervention was not possible. To ensure fidelity the research team did undertake observations within some of the experiment units, although this relied heavily on achieving 'buy-in' from senior and operational officers in the run up to the experiment implementation. Buy-in was sought through the research team attending meetings with Inspectors heading up the individual units, undertaking focus groups with operational officers, and visiting each unit within the experiment group during the break period to engage in dialogue with officers about the nature and purpose of the study. A question was also added to the questionnaire distributed by officers in the experiment group, asking respondents whether or not they had received a leaflet during their encounter. Only 20 respondents (11% of the experiment group) could not recall receiving one.

The survey instrument

Questionnaires were handed to drivers by police officers in an A4 envelope at the end of each encounter. A Business Reply envelope was enclosed to allow participants to return the questionnaire free of charge to the University of Edinburgh and the instruction sheet also directed participants to an alternative online questionnaire administered via Survey Monkey (in the event only 16 respondents utilised this option and, since all surveys were administered in a visual mode, these 16 were simply analysed together with the postal returns). The voluntary nature of participation in the survey was explained, and contact details for the research team were included.

Given the high volume of encounters included in the trial, this method (self-completion postal survey) was the only viable option for capturing driver opinion. However, this method typically yields low response rates, as had been the case in the original QCET (Antrobus et al, 2014). Postal surveys generally fare better when the option for follow up of non-responders is present. However, ethical and administrative concerns precluded follow up in this instance. When drivers are stopped by the police in Scotland, as in Australia, their personal details are not requested nor recorded by the police. To require officers to request personal details for the purposes of the study would have greatly increased the burden on officers and drivers, substantially lengthening the duration of stops and necessitating that officers carry and securely store additional 'tools' (notepads, pens, tablets) for recording purposes. This raises a number of ethical issues as well, not least of all how officers would be expected to store and transfer personal details to the research team. Crucially, though, having a police officer ask for personal details may be misunderstood by drivers as a function of the stop rather than the research, potentially leading to misinformed consent, not freely given. On the flip side, assuming the request was well communicated and understood, there is also a question of whether people would be prepared to give up their personal details. It seems unlikely that those who chose not to return their questionnaire would have consented to having their details taken for research purposes. It may also have raised suspicion amongst drivers, potentially leading to a reduced response rate overall.

A small incentive was added to the questionnaire in an attempt to boost response rates. By way of thanking drivers, the research team offered to make a donation to a charity of the driver's choice following completion of the study. Three charities were included for drivers to choose between: Macmillan Cancer Support; Barnardo's Scotland; and Brake (a road accident charity). In the event, some 43% (of 816) of respondents ignored this element of the questionnaire, failing to specify which charity they wished to donate to. The reasons for this are unclear. It may be that respondents simply did not notice the options (the question was posed on the cover page at the end of the instructions), or it may be that they did not wish to donate to any of the available charities. In either case this pattern, coupled with the overall low response rate, suggests that adding such an incentive had limited impact on our response rate. Future similar studies could perhaps fruitfully explore other forms of incentive.

Outcome measures

The content of the survey instrument was developed to facilitate comparison of findings and outcomes with the original QCET, but was altered to reflect the Scottish context and provide a 'test bed' for emerging theory in this particular criminological field. Tried and tested questions (drawn largely from the Scottish Crime and Justice Survey and the European Social Survey) and new questions intended to develop arguments and theories regarding the key constructs under consideration were utilized. All attitudinal questions employed Likert-type response scales.

The questionnaire assessed respondents' perceptions of their encounter with police officers, asking about: judgements of the police officer during the road stop and her or his adherence to principles of procedural justice or fairness; their trust in the police officers conducting the stop (note that this aspect of peoples' experience is likely to be highly correlated with, and is in a sense part of, procedural justice); and their satisfaction with both their treatment during the encounter and its final outcome. Contextual questions about why respondents thought the police had stopped them; the number of officers present; and whether a breath test was administered were also included. Further sections of the questionnaire examined attitudes to

police more generally, for example general trust in the fairness and effectiveness of the police, legitimacy, and demographic information was requested in order that the experimental and control groups could be compared for equivalence.

As per the hypotheses outlined above, the key latent concepts the questionnaire sought to capture are: perceptions of officer adherence to procedural justice; trust in the police during the encounter; satisfaction with encounter and police; general trust in the fairness and effectiveness of the police, and police legitimacy. In the main article (MacQueen and Bradford 2015, this issue) we present results in relation to these latent variables, since they provide better, more robust, measures of the underlying constructs of interest. Confirmatory factor analysis (CFA) using the statistical package Mplus 7.1 was employed to assess the scaling properties of the individual indicators and their ability to capture the underlying latent constructs, and also whether or not these constructs were themselves empirically distinct from one another.

Taking first those constructs relating to the encounters, six well-established items were used to capture perceptions of procedural justice during the stop, asking to what extent respondents believed that officer(s) who stopped them were approachable and friendly, helpful, respectful, professional, fair and clear in explaining why they had been stopped. Four further items probed respondents' views of the extent to which they trusted the officer(s) they encountered, covering trust in the officer(s) motives, actions, and the quality of the relationship these generated. Trust in these terms is closely related to, and generated by, procedural fairness: when someone treats us with dignity and respect we are encouraged to believe they have the right motives toward us. Finally, four items were employed to capture satisfaction with the encounter, including satisfaction with the behaviour of the officers involved, the way the stop was conducted, the way the respondent was treated by the officers, and the final action taken by the police.

 Turning to attitudes toward the police more generally, the literature suggests that 'trust and confidence' in the police comprises (at least) two distinct components: trust in police fairness and trust in police effectiveness (Jackson and Bradford, 2010). In assessing general trust in the

police in Scotland, the questionnaire attempted to capture both elements by including four items on fairness (in decision making and decision making processes, belief in equality, and the treatment of individuals with dignity and respect) and six on effectiveness (including a series of tasks that comprise policing practice, such as dealing with incidents and solving crime, and maintaining police values).

Finally six survey items tapped perceptions of police legitimacy. Again, these were intended to tap into two sub-components of the concept: 'duty to obey' and 'moral alignment' (Jackson et al 2012, 2013). The first component captured respondents' assessment of their sense of duty toward the police, as legitimate authority commands obedience from those subject to it. The second component captured respondents' assessment of the extent to which police operate according to a general moral framework that they themselves share, as authority is legitimate when it is applied in a manner congruent with shared norms and values.

**Technical Appendix Table 1 near here.

Appendix Table 1 shows results from a seven-factor CFA model, with the above items as observed indicators. [5] Indicators were set as categorical, and no cross loadings were allowed. The approximate fit statistics indicate a relatively good fit to the data (Byrne 2012); note also that factor loadings and the R^2 of the observed indicators are almost uniformly very high. However, as Appendix Table 2 shows the three latent variables representing respondents' assessments of the encounter itself were very strongly inter-correlated – discriminant validity here is poor. In practical terms, this means that from the perspective of the respondents, issues of procedural fairness and trust during the stop, and satisfaction with the officers who conducted it, were very strongly associated with each other (indeed, this is hardly surprising). However we proceed with three separate variables, both to maintain comparability with the

---

[5] Note that full information maximum likelihood estimation was used in this and all models, meaning that cases (respondents) with missing values on some variables were not excluded from analysis (cases with all missing values are excluded). This had the effect of boosting the effective sample size considerably.

QCET outputs and to investigate whether, despite their high correlations, assessments across the three factors were affected differently by the experimental intervention. A similar pattern is found across the two trust variables and the two measures of legitimacy (see Appendix table 2); discriminant validity was again poor, but we proceed with separate measures to maintain comparability with other studies and to explore whether the intervention had a differential impact across different constructs (with trust in police fairness and effectiveness being the most obvious case where some variation might be expected).

**Technical Appendix Table 2 near here

References

Ariel, B. and Farrington, D.P. (2010) 'Randomised block designs' in Piquero, A.R. and Weisburd, D. (eds) (2010) The Handbook of Quantitative Criminology, New York: Springer

Braga, A A., Weisburd, D.L., Waring, E.J., Mazerolle, L., Spelman, W. and Gajewski, F. (1999) 'Problem-oriented policing in violent crime places: a randomised controlled experiment' in *Criminology* 37(3) pp.541-580

Byrne, B. (2012) *Structural Equation Modeling with MPlus*, New York: Routledge

Jackson, J., & Bradford, B. (2010) 'What is trust and confidence in the police?' in *Policing: An International Journal of Police Strategies & Management*, *4*(3), 241–248. Retrieved from http://eprints.lse.ac.uk/27716/

Jackson, J., Bradford, B., Hough, M., Myhill, A., Quinton, P., & Tyler, T. R. (2012) 'Why do People Comply with the Law?: Legitimacy and the Influence of Legal Institutions' in *British Journal of Criminology 52*(6) pp.1051–1071

Jackson, J., Bradford, B., Stanko, B. and Hohl, K. (2013) *Just Authority? Trust in the police in England and Wales*, London: Routledge. Lawrence

MacQueen, S. and Bradford, B. (2015) 'Enhancing public trust and police legitimacy during road traffic encounters: Results from a randomized controlled trial in Scotland' in *Journal of Experimental Criminology*

Weisburd, D.L. and Gill, C. (2014) 'Block randomised trials at places: Rethinking the limitations of small N experiments' in *Journal of Quantitative Criminology* 30(1) pp.97-112

Weisburd, D., and Green, L. (1995) 'Policing drug hot spots: The Jersey City Drug Market Analysis Experiment' in *Justice Quarterly* 17 pp.11-736

Technical Appendix Table 1: Results from 7-factor Confirmatory Factor Analysis model

	Factor loading	R^2
Stop procedural justice *(Do you think the police were)*		
Approachable and friendly	0.93	0.87
Helpful	0.91	0.83
Respectful	0.96	0.93
Professional	0.95	0.89
Fair	0.96	0.93
Clear in explaining why you had been stopped	0.73	0.53
Trust in officers conducting stop *(Please indicate how much you agree or disagree with the following statements)*		
I trusted the intentions of the police officer involved	0.95	0.90
I was confident the police officer was doing the right thing	0.98	0.96
The police officer gave me the opportunity to express my views	0.92	0.85
The police officer listened to what I had to say	0.95	0.90
Overall Satisfaction with encounter *(Please indicate how satisfied or dissatisfied you were with the following)*		
The way the police officer conducted the stop	0.93	0.86
The way you were treated	0.99	0.98
The final action taken by the police officer(s)	0.90	0.82
Trust in police fairness *(How often do you think the police in Scotland …)*		
Make fair decisions	0.90	0.81
Listen to people before making decisions	0.87	0.76
Treat people with dignity and respect	0.93	0.87
Treat everyone equally	0.88	0.78
Trust in police effectiveness *(how confident are you in the ability of the police in Scotland to)*		
Prevent crime	0.87	0.76
Respond quickly to appropriate calls from the public	0.84	0.71
Deal with incidents as they occur	0.89	0.79
Solve crimes	0.94	0.88
Catch criminals	0.95	0.90
Keep people safe	0.91	0.82

Legitimacy - duty to obey *(Please indicate how much you agree or disagree with the following statements)*

Please indicate how much you agree or disagree with the following statements

I feel a moral obligation to obey the police	0.85	0.72
I feel a moral duty to support the decisions of police officers, even if I disagree with them	0.80	0.63
I feel a moral duty to obey the instructions of police officers, even when I don't understand the reasons behind them	0.82	0.67

Legitimacy - moral alignment *(Please indicate how much you agree or disagree with the following statements)*

The police have the same sense of right and wrong as me	0.81	0.62
The police stand up for values that are important for people like me	0.93	0.87
I support the way the police usually act	0.91	0.83

Fit statistics

Chi^2	1409.9
df	356
p	<.00005
RMSEA	0.06
CFI	0.98
TLI	0.98

Results from a seven factor solution with no factor loadings

Technical Appendix Table 2: Correlation matrix of latent variables

	1	2	3	4	5	6	7
Stop procedural justice (1)	1						
Trust in officers conducting stop (2)	0.83	1					
Overall satisfaction with encounter (3)	0.95	0.89	1				
Trust in police fairness (4)	0.80	0.71	0.76	1			
Trust in police effectiveness (5)	0.69	0.63	0.68	0.80	1		
Legitimacy - duty to obey (6)	0.59	0.60	0.61	0.64	0.61	1	
Legitimacy - moral alignment (7)	0.76	0.75	0.78	0.84	0.73	0.85	1