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

Sarah MacQueen[1] · Ben Bradford[2]

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-randomised (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, nor 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.

✉ Sarah MacQueen
Sarah.MacQueen@ed.ac.uk

[1] Scottish Centre for Crime and Justice Research, University of Edinburgh Law School, 15 Buccleuch Place, Edinburgh EH8 9LN, UK

[2] Centre for Criminology, University of Oxford, Oxford, UK

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 Randomised, 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 hypothesised 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, nor on drivers' legitimacy judgments. Results from the ScotCET experiment therefore suggest that operationalisation 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. 2013; 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 2008).

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 operationalised at the level of everyday policing have also been rare (Mazerolle et al. 2013a, b).

To address such questions, the Queensland Community Engagement Trial (QCET) adopted randomised 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 experimental 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, for example, 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 experimental 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 (Mazerolle et al. 2011, 2012; Murphy et al. 2014).

The key limitation of QCET has been the specificity of its context and the lack of replication to confirm the results achieved: the generalisability 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 (Fig. 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 drivers' 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.

The dotted lines in Fig. 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.
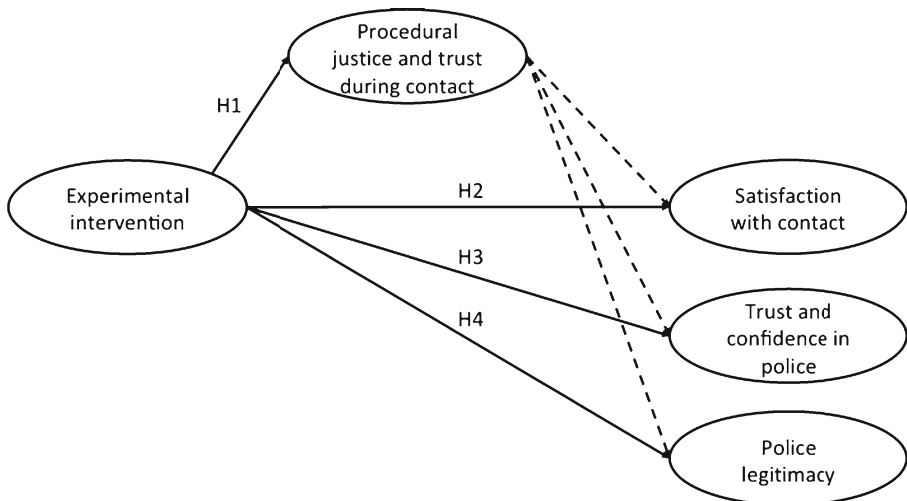

Fig. 1 ScotCET conceptual map

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 organisation 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 reorganisation 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 reorganisation had significant implications for the experiment that we describe below. In short, while nominally one organisation, the 'legacy forces' within Police Scotland retained significant affective meaning for police and, perhaps, the 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 organisation.

Like QCET, ScotCET was designed in a road policing context, with experimental 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 randomisation in the RCT (see below). As in the rest of the organisation, 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 drivers' 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 experimental 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 experimental 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. 2011, 2012). 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 experimental 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 standardise 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 drivers' 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].

---

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

Implementing the study

For implementation, random assignment of encounters to experimental 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 experimental 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 experimental 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

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 experimental condition and 335 responses from the control condition. Thus, the overall response rate was 6.6 %. At the baseline, the response rate was 7.2 % (6.2 % within the experimental group and 8.2 within the control group), and during the post period it was 6.2 % (5.2 % within the experimental group and 6.9 % within the control group).

This is substantially lower than the response rate achieved in the original QCET (13 % 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 experimental 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 % 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.

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

Table 1 Response rates across the RPUs

	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 %

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.

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 randomised design, which should have ensured all 'types' of potential respondents and non-respondents were evenly distributed between experimental 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 experimental and control groups – see below). Moreover, as noted above, the response rates for the experimental group were 6.2 % in the 'pre' period and 5.2 % 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 % to 6.9 %, with the difference statistically significant at the 10 % level ($z=-1.95$; $p=.1$). The fact that response rates fell in *both* experimental *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 % had a first degree or higher, while 12 % reported they had no qualification. Seventy-one per cent were in employment, 21 % were retired, and 73 % 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 experimental and control groups.

Table 2 Demographic breakdown of sample (% of respondents)

	'Pre' survey		'Post' survey		Total
	Experiment	Control	Experiment	Control	
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

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 7 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 2014).

Analytical technique

To compare the difference between reported perceptions among drivers in the experimental group and those in the control group, a 'difference in differences' approach was used with the unit level randomisation. Use of a matched pairs design in allocation also had to be taken into account in case organisational 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 experimental 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 experimental 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

---

[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.

assessments of police behaviour in the experimental sites compared to the control sites. The value for coefficient (4) is simply (3) minus (2); it represents the difference between change in the experimental 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 experimental 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 experimental areas relative to the control areas (effects significant at the 10 % level).

A brief examination of the contextual information gathered in the questionnaire that may help to explain this finding reveals no significant difference between experimental and control groups. People in the experimental areas were no more likely to think they had been stopped for the 'wrong' reason, nor where they more likely to have been breathalysed. It therefore seems that something in the experimental intervention *damaged* the perceived procedural justice of officers' actions, and this effect dampened, in

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	

**$p<.05$; *$p<.1$

the experimental areas, the improvement in opinions that occurred in the control areas. That is, since the only systematic difference between the experimental 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 experimental 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 the 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).

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.

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 dichotomised 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.

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 period				
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 period				
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	

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 experimental group. No evidence in support of this hypothesis was found; instead, the evidence suggests the intervention led to *diminished* feelings of procedural justice. 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 be an active participant both during and after the encounter (Murphy et al. 2008; Tyler 2006;

Tyler and Huo 2002). 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 were 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 that 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 (Mazerolle et al. 2012; Murphy et al. 2014). 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 utilisation 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 (Bradford et al. 2009; Skogan 2006; although see Myhill and Bradford 2012). When people are already generally positive about police, further improving opinions may be even more challenging, and, arguably, less meaningful as an endeavour. 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 experimental 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 begins to diminish. It appears that, perhaps unsurprisingly, delaying drivers unnecessarily results in negative reaction or 'backfire' effect. It may be that drivers in our experimental 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 'bureaucratised' 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, 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 experimental 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 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 posttest period had lasted longer, delivery of key messages and materials would have been more 'naturalised' in practice and scores across the experimental 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 % ($n=20$) 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 experimental 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 generalisable 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 utilise 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 to espouse that even the shortest encounters between citizens and police can provide 'gain' (Mazerolle et al. 2013a, b), 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 are, and are 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 and 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.

Appendix 1

Table 6 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 specifically 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

Fig. 2 ScotCET Leaflet

Appendix 3

Table 7 Procedural justice during the encounters

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

	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:

	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

Table 7 (continued)

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

	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

References

Antrobus, E., Elffers, H., White, G., & 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., & Farrington, D. P. (2010). Randomised block designs. In A. R. Piquero & D. Weisburd (Eds.), *The handbook of quantitative criminology.* New York: Springer.

Boruch, R., Weisburd, D., & Berk, R. (2010). Place randomised trials. In A. R. Piquero & D. Weisburd (Eds.), *The handbook of quantitative criminology.* New York: Springer.

Bradford, B., Jackson, J., & Stanko, E. A. (2009). 'Contact and confidence: Revisiting the impact of public encounters with the Police'. *Policing and Society, 19*(1), 20–46.

Bradford, B., Huq, A., Jackson, J., & Roberts, B. (2014a). 'What price fairness when security is at stake?Police legitimacy in South Africa. *Regulation and Governance, 8*(2), 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., & Bynum, T. (1994). Global and specific attitudes toward the police: Disentangling the relationship'. *Justice Quarterly, 11*(1), 119–134.

Hinds, L., & Murphy, K. (2007). Public satisfaction with police: using procedural justice to improve police Legitimacy'. *Australian and New Zealand Journal of Criminology, 40*, 27–42.

Hohl, K., Bradford, B., & 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., & Bradford, B. (2013). The governance of criminal justice, legitimacy and trust'. In S. Body-Gendrot, R. Lévy, M. Hough, S. Snacken, & K. Kerezsi (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. *British Journal of Criminology, 52*(6), 1051–1071.

Jackson, J., Bradford, B., Stanko, B., & Hohl, K. (2013). *Just authority? trust in the police in England and wales.* London: Routledge.

Jonathan-Zamir, T., & 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., Bennett, S., Eggins, E., Antrobus, E., White, G., & 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.

Mazerolle, L., Bennett, S., Antrobus, E., & 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), 343–366.

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

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

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

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

Murphy, K., Hinds, L., & Fleming, J. (2008). Encouraging cooperation and public support for Police'. *Policing and Society, 18*, 138–157.

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

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

Papachristos, A., Meares, T., & Fagan, J. (2012). 'Why do criminals obey the Law? the influence of legitimacy and social networks on active Gun offenders. *Journal of Criminal Law and Criminology, 102*, 397–439.

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multi-level modelling using stata (2nd edition).* College Station, TX: Stata.

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., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont: Wadsworth Cengage Learning.

Skogan, W. G. (2006). 'Asymmetry in the impact of encounters with the police. *Policing and Society, 16*(2), 99–126.

Snijders, T., & Bosker, R. (2012). *Multilevel analysis* (2nd ed.). London: Sage.

Sunshine, J., & Tyler, T. (2003). 'The role of procedural justice and legitimacy in public support for policing. *Law and Society Review, 37*(3), 513–548.

Tankebe, J. (2008). Police effectiveness and police trustworthiness in Ghana: an empirical appraisal. *Criminology and Criminal Justice: An International Journal, 8*, 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., & Huo, Y. J. (2002). *Trust in the Law: Encouraging public cooperation with the police and courts*. New York: Russell Sage Foundation.

Weisburd, D. L., & Gill, C. (2014). 'Block randomised trials at places rethinking the limitations of small N experiments. *Journal of Quantitative Criminology, 30*(1), 97–112.

Wünsch, D., & 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), 331–339.

Sarah MacQueen is a Research Fellow with the Scottish Centre for Crime and Justice Research, based at the University of Edinburgh Law School. Her most recent research explores the impact of the quality of interaction between police and citizens on public perception and judgements on procedural fairness and police legitimacy, testing the possibility of operationalizing the procedural justice framework in routine police encounters. She is also exploring the barriers and pitfalls in experimental policing research. Other research includes an examination of differential risk and experience of domestic abuse, and the ways in which particular groups of victims respond.

Ben Bradford is Departmental Lecturer in Criminology at the Centre for Criminology, University of Oxford. His research revolves around issues of procedural justice, trust, legitimacy, identity and compliance in criminal justice contexts, particularly in regard to the relationship between police and public. He has conducted studies on these and related issues in the UK, Australia, South Africa, Pakistan and other countries, and cross-national comparison is a growing aspect of his work. He also has an interest in organizational justice and the internal dynamics of police organizations, and in elements of operational policing, particularly stop and search, community patrol and other elements of 'public-facing' police work.