



Blurring the Distinction Between Empirical  
and Normative Legitimacy?

A Commentary on 'Police Legitimacy and  
Citizen Cooperation in China'

Jonathan Jackson and Ben Bradford

LSE Law, Society and Economy Working Papers 5/2019

London School of Economics and Political Science

Law Department

This paper can be downloaded without charge from LSE Law, Society and Economy Working Papers at: [www.lse.ac.uk/collections/law/wps/wps.htm](http://www.lse.ac.uk/collections/law/wps/wps.htm) and the Social Sciences Research Network electronic library at: <http://ssrn.com/abstract=3335644>.

© Jonathan Jackson and Ben Bradford. Users may download and/or print one copy to facilitate their private study or for non-commercial research. Users may not engage in further distribution of this material or use it for any profit-making activities or any other form of commercial gain.

# Blurring the Distinction Between Empirical and Normative Legitimacy? A Commentary on ‘Police Legitimacy and Citizen Cooperation in China’

Jonathan Jackson and Ben Bradford\*

**Abstract:** In a fascinating paper on the nature of police legitimacy in Southern China, Sun et al. (2018) present evidence that what have previously been treated as *possible sources of legitimacy*—public perceptions of police procedural justice, distributive justice, effectiveness and lawfulness—are in fact *constituent components of legitimacy*. We argue in this paper, that the empirical strategy used to reach this conclusion is *not fit for purpose* because both conceptual stances—*possible sources of legitimacy* or *constituent components of legitimacy*—are consistent with the same fitted statistical model. Ironically, therefore, Sun and colleagues end up assuming rather than discovering the normative bases on which people judge police legitimacy. To be sensitive to cultural context means using a methodology that does not *a priori* impose the preconditions of legitimacy. We illustrate this general point by analysing nationally representative data from 30 countries across Europe and beyond.

**Key words:** legitimacy, police, measurement, procedural justice, distributive justice.

---

\* Jonathan Jackson is Professor of Research Methodology, Head of Department of Methodology, and member of the Mannheim Centre for Criminology at the London School of Economics and Political Science. Ben Bradford is Professor at the Department of Security and Crime Science, University College London. The analysis and writing of this paper was funded by the United Kingdom’s Economic and Social Research Council: ES/R011397/1, From Coercion to Consent: Social Identity, Legitimacy, and a Process Model of Police Procedural Justice; and ES/L011611/1, Legal Norms and Crime Control: A Comparative, Cross-national Analysis. Funding for the three projects were provided by: European Commission and various national funding bodies on the European Social Survey; grants obtained by Professor Tom Tyler from Yale Law School and New York University; and South Africa Human Sciences Research Council.

The concept of legitimacy has moved centre-stage in police research. Police are empowered to use whatever level of force is deemed necessary to deal with issues of crime and disorder, so questions concerning their rightful use of power are never far from the political surface. Normative concerns about police powers are ever-present in the debates that almost continuously roil around policing. Yet, police in liberal democracies rely on the legitimacy they command and the public cooperation, deference and compliance it engenders, and this raises *empirical* concerns about the extent to which the policed hold the police legitimate.

The past decade and more has seen the pathbreaking United States-based work of Tom Tyler and colleagues into the causes and consequences of perceived police legitimacy (Sunshine & Tyler 2003; Tyler, 2006a, 2006b) spread across the world. Studies have been carried out in social, political and legal contexts as diverse as Ghana, Finland, the Russian Federation, the United Kingdom, Pakistan, Sweden, Japan, Israel, Australia, Turkey, South Africa, France, Ukraine, China, and Nigeria. In many of these countries procedural justice seems to be the most important predictor of legitimacy: namely, legitimacy seems to rest a good deal on the extent to which police officers act in fair, neutral, transparent and trustworthy ways when making decisions and interacting with the public (see Jackson 2018, for a review of the international literature).

Sun et al.'s (2018) recently published paper into police legitimacy in a coastal city in Southern China is an excellent example of the increasingly international nature of this field of enquiry (see also Tsushima & Hamai 2015; Kim et al. 2018; Akinlabi & Murphy 2018). Sensitivity to context is important to their work. China is an authoritarian regime, so it has 'low accountability and high coercion', hence people's feelings of obligation to obey external legal authority may be complex and varied. But people may also judge the legitimacy of the police on the basis of procedural justice (fair process) *and* distributive justice (fair aggregate allocation of outcomes) *and* effectiveness against crime *and* lawfulness:

the police in an authoritarian state are commonly empowered with excessive authorities that do not match normative expectations of democratic policing (e.g., procedural fairness, institutional transparency, and accountability). Authoritarian policing is thus prone to abusive treatments of the public and state manipulative efforts of performance. Lawfulness, distributive justice, and effectiveness, originally proposed by Tyler as less imperative than procedural justice, could play a different or even an enlarged role in shaping police legitimacy under an authoritarian setting. (Sun et al. 2018: 2).

Drawing on data from a city-wide survey, they conclude from their analysis that these four judgements are so strongly bound up with legitimacy that they *collectively constitute the perceived right to rule*. In other words, rather than legitimacy being an overarching judgement about the right to power and the authority to govern—that may or may not be influenced by public judgements about whether police tend to

act in procedurally just, distributively just, effective and lawful ways—legitimacy *is* procedural justice, distributive justice, effectiveness and lawfulness. Their study raises important questions regarding the importance of context in police-citizen relations, the conceptualisation and measurement of legitimacy, and the role of the researcher in (a) allowing the preconditions of legitimacy to be an empirical question discovered *bottom-up* or (b) imposing the preconditions of legitimacy onto a given political community *top-down*.

Our goals in this paper are twofold. First, while Sun et al. (2018) makes a notable contribution to the literature, we question in these pages whether the analytical strategy used to assess the legitimacy measurement model is *fit for purpose*. More specifically, we argue that the analytical strategy is not a good adjudication tool because both conceptual stances (procedural justice, distributive justice, effectiveness and lawfulness as either *possible sources of legitimacy* or *constituent components of legitimacy*) are consistent with the same fitted confirmatory factor analysis model. Second (and relatedly) we make the case that their approach lacks cultural sensitivity because it ends up being the outside expert who is deciding what criteria people use to judge institutional legitimacy, not the people themselves. We illustrate this general point by analysing data from a 30-country study.

The paper proceeds as follows. By way of conceptual ground-clearing we distinguish between the normative concept of legitimacy of political philosophers and the empirical concept of legitimacy of social scientists. After summarising the standard approach to studying empirical legitimacy in the context of the police, we review the original work of Tankebe (2013) whom Sun and colleagues sought to replicate, and then discuss Sun et al.'s (2018) approach. Following the findings from our own empirical study into the measurement of legitimacy in 30 countries, we conclude with some thoughts on future directions of research in this area.

## **NORMATIVE AND EMPIRICAL LEGITIMACY**

Political philosophers often employ legitimacy as a normatively-laden term to describe whether state institutions meet an inherently value-based set of substantive criteria regarding how they ought to be configured if their power is to be judged as rightfully held. In the context of the criminal justice system, a Western democratic conception of normative legitimacy might involve a group of outside experts deciding that institutions should be judged by according to principles of independence, transparency, accountability and other features of the rule of law.

By contrast, social scientists typically employ legitimacy as an empirically-laden concept to describe whether—as a matter of fact—those that are subject to authority confer legitimacy on that authority (Tyler 2006a, 2006b; Calderia & Gibson 1995; Gibson et al. 2003; Justice & Meares 2014; Meares 2017; Trinkner et

al. 2018). The empirical concept of legitimacy focuses on whether an institution finds ‘the approval of those who have to abide by it’ (Hinsch 2010: 40). Legitimacy is premised on a *fundamental accord* between rulers and ruled (Filiangeiri 1783-88, in Pardo 2000: 5) that is founded in shared norms and values and established via the ‘moral performance’ (Liebling 2004) of power-holders.

In the context of the police, social scientists typically operationalize empirical legitimacy along two connected lines:

1. normative justifiability of power in the eyes of citizens (the right to rule): do citizens believe that the police as an institution is just, proper and appropriate?
2. recognition of rightful authority (the authority to govern): do citizens believe that police officers are entitled to be obeyed?

The empirical legitimacy process involves acceptance (or rejection) of the implicit and explicit claims that police make to be a morally appropriate institution that has the right to expect obedience and support from the public. On the one hand, people judge the normative appropriateness of the police as an institution against societal norms of conduct of officials (e.g. do police officers make neutral and objective decisions when dealing with citizens?) and draw relevant conclusions regarding the legitimacy of the institution that officers embody and represent. On the other hand, the sources of legitimacy (i.e. the bases on which its presence or absence is judged by citizens) emerge from empirical enquiry. They are not assumed *a priori* on the basis of political, moral, legal, religious or some other philosophy.

#### THE STANDARD APPROACH TO STUDYING EMPIRICAL LEGITIMACY

Researchers typically (a) operationalise legitimacy as a psychological construct, (b) treat the normative appropriateness part of the legitimacy construct as an overarching judgment, and (c) allow the criteria that people use to judge legitimacy to be an empirical question. This strategy depends on distinguishing between potential sources of legitimacy (how officers are perceived to act) and overarching legitimacy judgements (whether the institution is seen to have the right to power and authority to govern). Because of the abstract nature of institutional legitimacy, researchers focus on the general appropriateness of the actions of officers who embody the institution, and to avoid imposing specific value content, measures of legitimacy are worded in general ways (like ‘the police usually act in ways that are consistent with your sense of right and wrong’ and ‘the police generally have the same sense of right and wrong as I do’) rather than in more specific ways (like ‘how often do the police treat people with respect and dignity?’ and ‘how effective are the police in preventing burglaries from happening?’). Researchers then use statistical analysis to assess whether, for instance, procedural justice and

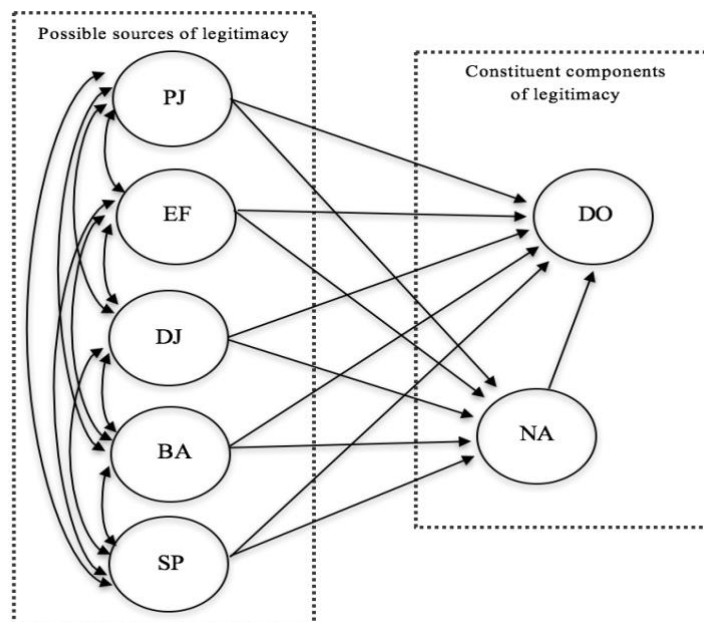
effectiveness judgements are the most important predictors of legitimacy, and if they are, then the inference is that these are two sources of police legitimacy.

In a recent example Huq et al. (2017) presented findings from a nationally representative sample survey of the UK that measured the following constructs:

1. Public attitudes towards whether officers act in procedurally just ways;
2. Public attitudes towards whether officers act in distributively just ways;
3. Public attitudes towards whether officers act in effective ways;
4. Public attitudes towards whether officers respect the limits of their rightful authority;
5. Public attitudes towards whether officers use appropriate surveillance powers;
6. Perceived institutional legitimacy: normative alignment; and,
7. Perceived institutional legitimacy: duty to obey.

Aziz Huq and colleagues first used confirmatory factor analysis to assess whether the seven constructs can reasonably be treated as empirically distinct albeit positively correlated constructs (an important first stage of analysis that we will return to). Having found evidence for empirical distinctiveness, they then used structural equation modelling to assess the extent to which each of the first five constructs predicted normative alignment and duty to obey (Figure 1).

Figure 1 Testing the possible sources of legitimacy<sup>1 2</sup>



<sup>1</sup> PJ = procedural justice; DJ = distributive justice; EFF = effectiveness; BA = bounded authority; SP = surveillance practices; NA = normative alignment; DO = duty to obey.

<sup>2</sup> CFA indicated that the constructs are equally distinct and positively correlated, and that there are good scaling properties.

Huq et al. (2017) found (a) that procedural justice and bounded authority were the key predictors of normative alignment and (b) that normative alignment and effectiveness were the key predictors of duty to obey. The conclusion was that the legitimacy (operationalised as normative appropriateness) of the police as an institution in the UK is judged most strongly on the basis of officers acting in procedurally just ways and respecting the limits of their rightful authority (two legitimating societal norms). Put another way, citizens seem to judge the legitimacy of the police as an institution in part on whether officers respect the limits of their rightful authority, treat people with respect and dignity, talk and listen to people, and act in unbiased, transparent, and accountable ways. In turn, normative alignment and police effectiveness were both predictors of willing consent to rightful authority.

The majority of studies using this approach have found that the most important predictor of legitimacy is the extent to which people think that officers act in procedurally just ways. This is the case in the US (Sunshine & Tyler 2003; Reisig et al. 2007; White et al. 2016), UK (Huq et al. 2011; Bradford 2014), Australia (Murphy & Cherney 2012; Murphy et al. 2018), Israel (Mentovich et al. 2018), China (Sun et al. 2017), and in Continental Europe (Hough et al. 2013; Dirikx & Ven den Bulck 2013). Notably, however, effectiveness and lawfulness judgements do seem to play a more important role in predicting empirical legitimacy in Pakistan and South Africa (Jackson et al. 2014; Bradford et al. 2014). The idea here is that normative judgements about fair process can to some degree be crowded out by concerns about police effectiveness and corruption, the sheer scale of the crime problem, and/or the association of the police with a historically oppressive and underperforming state.

## **POLICE LEGITIMACY AND PUBLIC COOPERATION IN SOUTHERN CHINA**

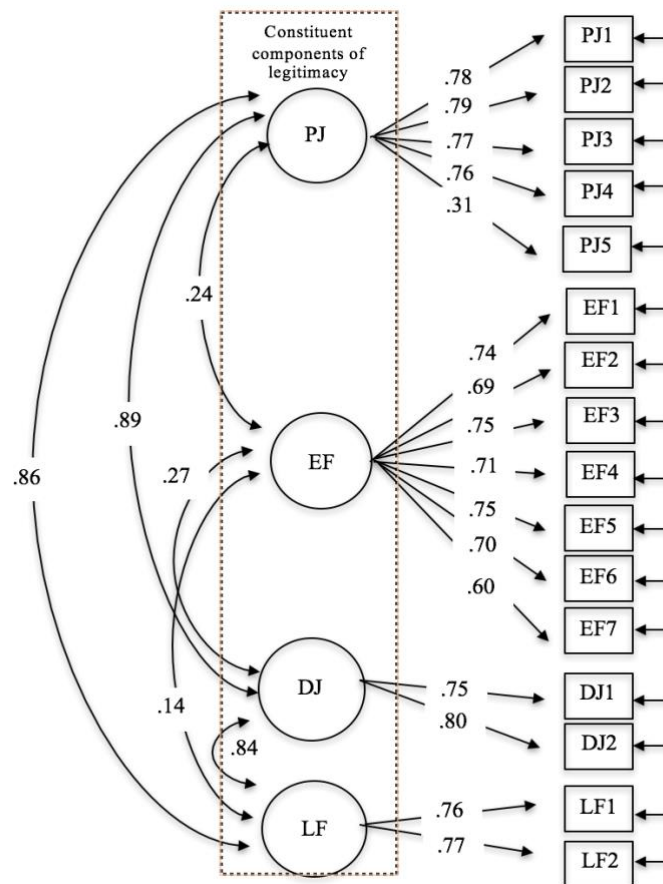
What, then, of Sun et al.'s (2018) study? Rather than using the above approach to assess whether legitimacy depends not only on procedural justice but also effectiveness, distributive justice and lawfulness, they used Tankebe's (2013) empirical strategy to adjudicate between the following two conceptual stances:

1. procedural justice, effectiveness, distributive justice and lawfulness are possible sources of legitimacy (where one then uses statistical modelling to determine the empirical importance of each one, see e.g. Huq et al. 2017);  
or,
2. procedural justice, effectiveness, distributive justice and lawfulness are constituent components of legitimacy (where they are so important that they collectively constitute the perceived right to power).

The details of the methodology used to adjudicate between these two positions are important. Sun et al. (2018) attempt to replicate Tankebe’s (2013) London-based study, so we first review the original study.

Tankebe (2013) drew upon data from a survey of Londoners that measured among other things people’s perceptions of police procedural justice, distributive justice, effectiveness and lawfulness (using multiple indicators of each). He found that a four-factor model that distinguished between the four constructs fitted the data reasonably well (summarized in Figure 2). Given the constraints in the fitted CFA model—e.g. no cross-loadings or error covariances and conditional independence of items given the four latent factors—one can treat these four judgments as distinct, albeit correlated, latent constructs.

Figure 2 Tankebe’s (2013) measurement model of legitimacy<sup>3 4 5</sup>



<sup>3</sup> Source: Tankebe J. (2013). Viewing things differently: The dimensions of public perceptions of legitimacy. *Criminology*, 51, 103–135.

<sup>4</sup> PJ = procedural justice; EF = effectiveness; DJ = distributive justice; LF = lawfulness.

<sup>5</sup> Fit statistics:  $\chi^2 = 1056$ ,  $df=98$ ,  $p<.005$ ; CFI = .970; RMSEA = .044 (95% CI .041-.046). The model fitted the data, and the measures had good scaling properties. For Tankebe (2013), this constitutes evidence that what is being measured here must be *constituent components of legitimacy* not *possible sources of legitimacy*.



Tankebe (2013) then tested whether effectiveness should be treated as a possible source of legitimacy or a constituent component of legitimacy. He fitted a three-factor CFA model without the effectiveness indicators and *effectiveness* latent variable (i.e. dropping EF1-EF7 and the EF latent variable in Figure 2). He found that the three-factor model fitted the data, ran a chi-square difference test to compare the relative fit of the three-factor and four-factor models, and noting that both the three-factor model (when indicators of procedural fairness, distributive fairness, and lawfulness were included) and the four-factor model (when indicators of effectiveness, procedural fairness, distributive fairness, and lawfulness were included) both fitted the data, he deduced from this that:

Effectiveness has to be viewed as a component of legitimacy; police organizations that seek legitimacy must demonstrate effectiveness as a normative requirement. Coicaud ... has put this well: “Every political ruler who seeks to prove he possesses the right to govern [that is, is legitimate] has to satisfy, to try to satisfy, or to pretend to satisfy the needs of the members of the community.” For the police, those needs include safety and security. (Tankebe 2013: 121).

Tankebe’s (2013) reasoning was simple. If the four-factor CFA model (Figure 1) fitted the data, then these four constructs should be labelled *constituent components of legitimacy*, not *possible sources of legitimacy*. In other words, the fitted model is incompatible with the notion that they are possible sources of legitimacy. The fact that the items scaled well; the fact that the analysis supported the idea that there were four underlying dimensions to the data; and the fact that these four factors were strongly and positive correlated (and that including effectiveness did not decrease the fit of the model) – all this means that legitimacy *is* procedural justice, distributive justice, effectiveness and lawfulness.

Overall, the findings suggest that what police researchers have persistently tended to use as predictors of legitimacy (procedural fairness, distributive fairness, lawfulness, and effectiveness) are rather the constituent parts of legitimacy ... The results of the confirmatory factor analysis presented in this study suggest that the debate [about whether legitimacy causes procedural justice or procedural justice causes legitimacy] might be redundant because procedural fairness is a constituent part of legitimacy rather than something apart from it. (Tankebe 2013: 125)

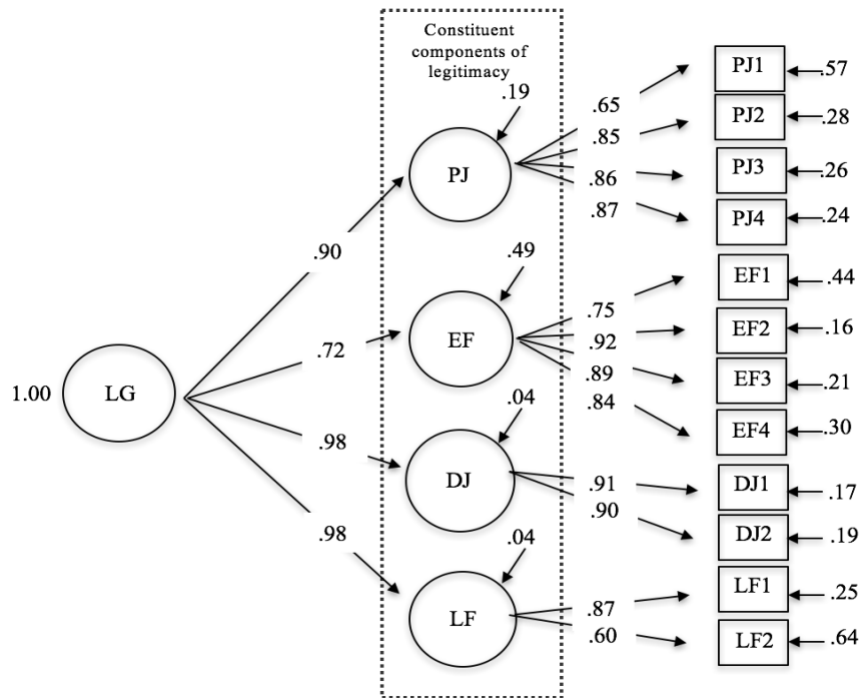
In their replication study, Sun et al. (2018) also measured public attitudes towards police procedural justice, distributive justice, effectiveness and lawfulness. Like Tankebe (2013), they also used confirmatory factor analysis to test a four-factor model. This time, however, they included a second-order factor that they labelled *legitimacy*, which means testing the idea that the second-order factor explains the

correlations between procedural justice, distributive justice, effectiveness and lawfulness (specifically that the bivariate correlations can be modelled according to one underlying latent construct). Of note is the fact that Sun and colleagues could just as reasonably tested a model with two or three second-order factors, since there is no requirement in the literature for legitimacy to be uni-dimensional (see Jackson & Gau 2015).

Sun et al. (2018) found that the model (reproduced in Figure 3) fitted the data, with (a) good scaling properties for each of the four sets of indicators, (b) four factors that were strongly and positively correlated, and (c) a second-order factor linked to each of the four first-order factors. Following the line of reasoning of Tankebe (2013) (see also Tankebe et al. 2016), they interpreted the findings as follows:

the convergent validity, discriminant validity, and internal consistency of all key measures were supported in the CFA analysis and the reliability tests...Substantively, these results mean two things: (1) procedural justice, distributive justice, effectiveness, and lawfulness are four distinct sub-constructs of legitimacy, and each sub-construct is well explained by its own corresponding observed variables, rather than by variables from a different sub-construct, and (2) the four sub-constructs correlate well with one another within their latent construct legitimacy. In short, Tankebe’s argument that procedural justice variables should be considered as indicators, rather than antecedents, of legitimacy, is supported. (Sun et al. 2018: 14).

Figure 3 Sun et al.’s (2018) measurement model of legitimacy<sup>6 7 8</sup>



<sup>6</sup> Source: Sun, I. Y., Li, L., Wu, Y., & Hu, R. (2018). Police legitimacy and citizen cooperation in China: testing an alternative model. *Asian Journal of Criminology*, 13(4), 275-291.

<sup>7</sup> LG = legitimacy; PJ = procedural justice; EF = effectiveness; DJ = distributive justice; LF = lawfulness.

<sup>8</sup> Fit statistics:  $\chi^2 = 206$ ,  $df = 50$ ,  $p < .005$ ; CFI = .982; TLI = .976; RMSEA = .058; and SRMR = .030.

Because this model fitted the data, and because the measures had good scaling properties, Sun et al. (2018) argued that the second-order factor must be labeled *possible sources of legitimacy* not *constituent components of legitimacy*.

It seems, then, that we have empirical evidence on what police legitimacy *is*, at least in the two current contexts. Legitimacy is not some overarching judgement about the normative appropriateness of the police, coupled with a felt moral duty to obey legal authorities, as specified by Tyler (2006a, 2006b), Huq et al. (2017) and others. We should not be assessing whether procedural justice, distributive justice, effectiveness and/or lawfulness explain variation in legitimacy. Legitimacy may not be *predicted* more strongly by procedural justice than by distributive justice, effectiveness and lawfulness (as typically found in extant work). Strikingly, the CFA modelling proves that legitimacy *is* procedural justice, distributive justice, effectiveness and lawfulness: if the police are to be viewed as legitimate in the current two contexts (see also Tankebe et al.'s 2016 work in the US and Ghana), police officers need to be seen to act in ways that are (a) procedurally just, (b) distributively just, (c) effective and (d) lawful.

### NORMATIVE OR EMPIRICAL CONCEPTS OF LEGITIMACY?

Their paper was so thought-provoking paper that we asked Sun and colleagues if they would share the data. Given the fundamental importance of the finding regarding measurement, we were especially interested in the CFA modelling. But while the data were not forthcoming, it turned out that we did not need the data – and this only underlines the point we would like to make.

To explain, imagine you are embarking on a new study into police legitimacy in a coastal city in South China. You begin with the *received wisdom* on the nature of legitimacy, i.e. Sun et al.'s (2018) study that showed that the residents of this coastal Chinese city judge the legitimacy of the police on the (roughly equal) bases of procedural justice, distributive justice, effectiveness and lawfulness. This key piece of work treated legitimacy as the joint distribution of these four constituent parts (i.e. it was represented as a second-order latent variable), legitimacy predicted cooperation, and the statistical effect was partly mediated by obligation to obey, which was treated as a potential outcome rather than constituent part of legitimacy (note that this second part of the study does not concern us here).

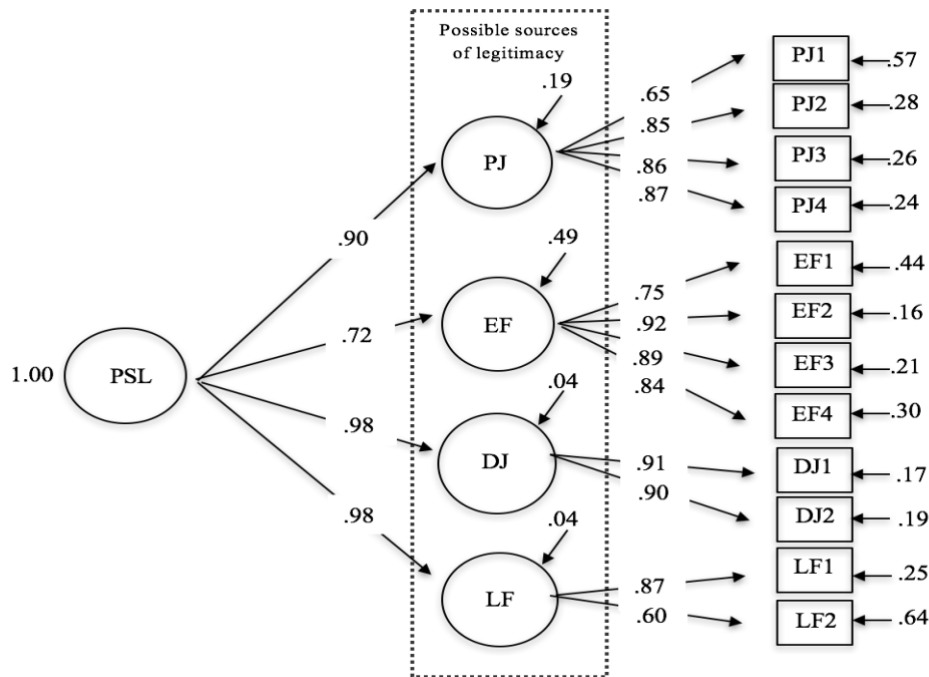
You are interested in questioning the status quo. You want to reverse the logic of Tankebe (2013) and Sun et al. (2018). Rather than legitimacy being procedural justice, distributive justice, effectiveness and lawfulness, you want to test a measurement model that (a) operationalises legitimacy as a more general

belief that the police are morally entitled to dictate appropriate behaviour, (b) treats procedural justice, distributive justice, effectiveness and lawfulness as possible sources of legitimacy, and (c) allows it to be an empirical question whether citizens of this coastal Chinese city judge police legitimacy on the basis of the procedural justice and/or distributive justice and/or effectiveness and/or lawfulness displayed by police officers, using the sort of statistical analysis employed by Huq et al. (2017) and others (e.g. Sunshine & Tyler 2003).

Recall that Huq et al. (2017) found evidence for the empirical distinctiveness of various constructs using CFA (Figure 1). They called the constructs on the left-hand side of the model *possible sources of legitimacy* and called the constructs on the right-hand side of the model *constituent components of legitimacy*. They did this because of prior theory, not because the CFA provided empirical proof of this conceptual designation. Your approach is different. Mirroring the reasoning of Tankebe, Sun and others, you start with an *a priori* conceptual stance that procedural justice, distributive justice, effectiveness and lawfulness are possible sources of legitimacy and conduct CFA to test whether you are in fact correct *in the first place*.

So you obtain Sun et al.'s (2018) data, you fit a confirmatory factor model, and you obtain the results summarized in Figure 4. You find that the measures of the four constructs (a) scale well, (b) can be represented as four latent variables, (c) are strongly and positively correlated, and (d) that these correlations between the latent variables can be modelled according to a second-order factor. Having *a priori* labelled the second-order factor *possible sources of legitimacy* you argue that the finding constitutes empirical proof that they are possible sources of legitimacy not constituent components of legitimacy (note that this mirrors the reasoning in Sun et al. 2018). You claim that the finding over-turns current thinking on police legitimacy in this coastal Chinese city. You then move on to test a model linking the potential predictors of legitimacy to obligation to obey (Figure 5), possibly showing that procedural justice is the strongest predictor (see Sun et al. 2017 for exactly that analysis).

Figure 4 An alternative measurement model of legitimacy<sup>9 10 11</sup>

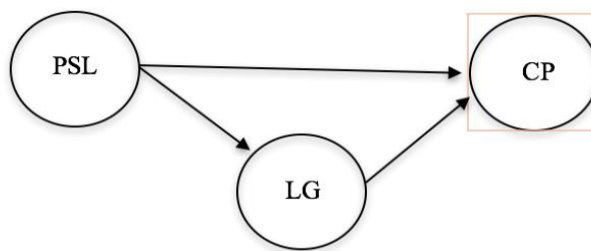


<sup>9</sup> Source: Sun, I. Y., Li, L., Wu, Y., & Hu, R. (2018). Police legitimacy and citizen cooperation in China: testing an alternative model. *Asian Journal of Criminology*, 13(4), 275-291.

<sup>10</sup> PSL = possible sources of legitimacy; PJ = procedural justice; EF = effectiveness; DJ = distributive justice; LF = lawfulness.

<sup>11</sup> Fit statistics:  $X^2 = 206$ ,  $df = 50$ ,  $p < .005$ ; CFI = .982; TLI = .976; RMSEA = .058; and SRMR = .030. Because this model fitted the data, and because the measures had good scaling properties, you argue that the second-order factor must be labeled *constituent components of legitimacy* not *possible sources of legitimacy*.

Figure 5 Theoretical model of possible sources of legitimacy, legitimacy and cooperation<sup>12</sup>



<sup>12</sup> PSL = possible sources of legitimacy; LG = legitimacy; CP = willingness to cooperate with the police.

Imagine, then, taking a moment to reflect. Tankebe (2013) and Sun et al. (2018) began by saying that prior studies have treated procedural justice, distributive justice, effectiveness and lawfulness as possible sources of legitimacy. This was the status quo for them. They then claimed that the findings from their CFA modelling showed that these four constructs were in fact constituent components of legitimacy, presumably since the fitted CFA models in Figures 2 and 3 must be

incompatible with labelling the constructs *possible sources of legitimacy*. In the words of Sun et al. (2018: 14), procedural justice, distributive justice, effectiveness and lawfulness are ‘indicators, rather than antecedents of legitimacy’.

You are doing the opposite. You start with a different status quo, i.e. that procedural justice, distributive justice, effectiveness and lawfulness are constituent components of legitimacy (Sun et al. 2018). Because your CFA modelling shows good convergence validity, discriminant validity and internal consistency (to paraphrase Sun et al., 2018 14) and because you find that a second-order factor (that you label ‘possible sources of legitimacy’) model fitted the data, what have previously been treated as constituent components of legitimacy are in fact possible sources of legitimacy. You conclude, in other words, that you were correct *in the first place* in how you labelled the constructs, and that Sun et al. (2018) were incorrect *in the first place* when they labelled the construct *constituent components of legitimacy*.

But a moment of reflection reveals that Figures 3 and 4 are identical, apart from the label given to the second-order factor. The model fits the data well regardless of what label we assign to the second-order factor. If both these different conceptual bases are consistent with the data (Figures 3 and 4), why would CFA provide empirical evidence on which of the two competing conceptual stances is ‘correct’? CFA is good at modelling correlations between variables according to some hypothesised latent structures. Whether one calls these constructs *possible sources of legitimacy* or *constituent components of legitimacy* depends on one’s conceptual stance, not on the basis of the CFA modelling. Moreover, to believe that CFA constitutes a test of the conceptual status of the constructs, you would have to reify latent variables in a particular way. Specifically, you would have to name the second-order factor before fitting the model, then interpret the fact that a second-order factor model fits the data as empirical proof that one was right to name the second-order factor in the way that you did, prior to doing the analysis.

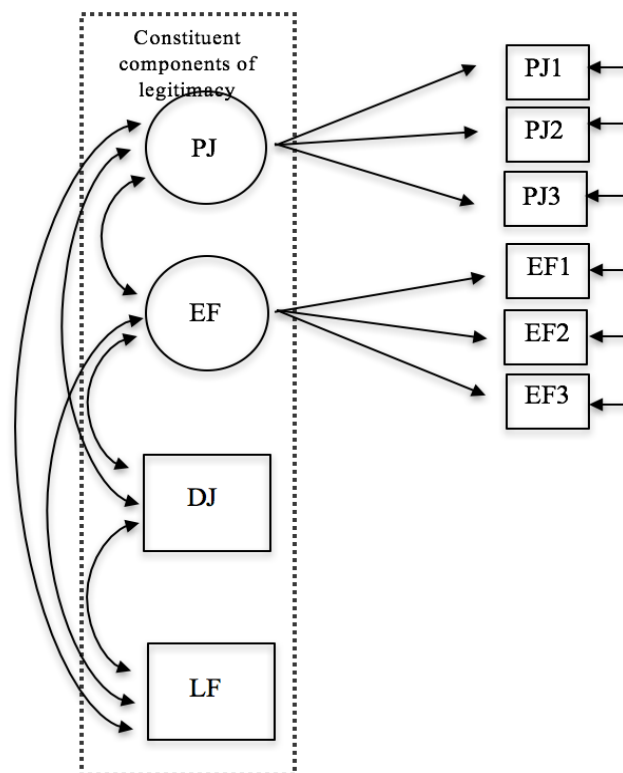
Indeed, it is difficult to know what would have needed to be present in a fitted CFA model if it was to fail the current *empirical test*. In Tankebe’s (2013) analysis it seems that if the fit of the model was reduced by adding effectiveness to the model, then this proves that the police do not have to be seen to act effectively if they are to be seen as legitimate by citizens. But why would that follow? It is a non sequitur: a less well fitting model would more likely indicate that there are cross-loadings and/or error covariances that should be added to the model. In Sun et al.’s (2018) analysis, it could be that a single second-order factor would not fit the data. But why would finding that one needs multiple second-order factors to explain why procedural justice, distributive justice, effectiveness and lawfulness are correlated somehow *prove* that procedural justice, distributive justice, effectiveness and lawfulness are potential sources of legitimacy? Another non sequitur: there is no requirement that legitimacy is unidimensional.

## TESTING THE CULTURAL (IN)SENSITIVITY OF THIS APPROACH TO MEASUREMENT: A THIRTY-COUNTRY STUDY

So far in this paper we have discussed what latent variable modelling can and cannot say about an issue that is ultimately down to conceptual analysis and operational argumentation. In the second part of this paper we turn to cultural sensitivity. Recall that Sun et al. (2018) motivated their empirical strategy by saying that the approach of Tankebe (2013) offers a greater level of cultural sensitivity than the traditional approach to measuring legitimacy. In particular, they wanted to test the idea that the legitimization of the police is more complex in China than it is in contexts like the US, UK and Australia and this meant testing the idea that legitimacy has four dimensions using Tankebe's (2013) approach.

To reflect on the general extent of cultural sensitivity, we apply the approach to 30 different countries. Using the same approach to measuring and modelling legitimacy as Tankebe (2013), we fit the same CFA model in each country separately. We distinguish between procedural justice, distributive justice, effectiveness and lawfulness (Figure 6). Following Tankebe (2013) we *a priori* call these constituent components of legitimacy.

Figure 6 Posited measurement model of potential sources of legitimacy<sup>13</sup>



<sup>13</sup>PJ = procedural justice; EF = effectiveness; DJ = distributive justice; LF = lawfulness.

## DATA

The ESS is an academically driven face-to-face interview survey that runs every two years. Charting a range of attitudes, values, behaviours and beliefs between nations and over time, it is one of the highest quality—if not the highest quality—cross-national surveys in the world, especially in terms of sampling and measurement equivalence. It employs a rigorous questionnaire translation, pre-testing and development methodology (Jowell et al. 2007). Although not all countries achieve it, the aspiration is that countries should have probability samples of the adult (16+) population, with high response rates, interviewed face-to-face using CAPI (computer assisted personal interviewing). The questionnaire comprises an invariant core of questions asked of all respondents in each round. Also included in some rounds are rotating modules that focus in detail on a particular issue. Academics are invited to bid for space on the questionnaire in each round.

In Round 5 a module on *trust in justice* containing 45 questions was included (European Social Survey 2011; Jackson et al. 2011; Hough et al. 2013). Fieldwork for Round 5 of the ESS was done in 2010/11 (European Social Survey 2010, 2018). A total of 28 countries took part, some of which were *European* in only a loose sense: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Lithuania, Netherlands, Norway, Poland, Portugal, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland, UK, and Ukraine. The probability samples are representative of all persons aged 15 and older resident within the borders of the nation, regardless of nationality, citizenship, language or legal status. The smallest sample size was 1,083 in Cyprus and the largest sample size was 3,031 in Germany.

The US data come from an internet-based survey fielded to a random selection of individuals drawn from a GFK Knowledge Networks research panel of U.S. adults (Tyler & Jackson 2014; Tyler et al. 2015). Knowledge Networks uses random digit dialling and address-based sampling methods to construct and maintain the panel. A total of 2,561 respondents were initially selected from the larger panel. The study was described, an offer of compensation extended, and a reminder email was sent to all people on the list who had not responded after three days. The survey was fielded in August and September of 2012, either in English or in Spanish. A total of 1,603 individuals completed the survey, representing a response rate of 62.5% from the existing internet panel.

The South Africa data come from the 2012 round of the South African Social Attitudes Survey (SASAS), a repeated cross-sectional survey conducted annually by the Human Sciences Research Council. The survey round consisted of a nationally representative probability sample of 3,183 South African adults aged 16 years and over living in private households. Each SASAS round of interviewing consists of a sub-sample of 500 Population Census enumeration areas (EAs), stratified by province, geographical sub-type and majority population group. The SASAS aims



to provide a long-term account of change in public values and the social fabric of modern South Africa. Given the importance of issues of crime and policing in South African society, permission was secured to field police-related questions from the trust in justice module included in the fifth round of the European Social Survey in 2010/11 (see Bradford et al. 2014).

## MEASURES

Procedural justice was measured by asking respondents how often (from 1 ‘not at all often’ to 4 ‘very often’) they think that officers in their country:

- treat people with respect;
- make fair and impartial decisions; and,
- explain their decisions and actions when asked.

Police effectiveness was measured by asking respondents: ‘Based on what you have heard or your own experience, how successful do you think the police in this country are at...’ (on a scale from 0 to 10):

- preventing crimes where violence is used or threatened;
- catching people who commit house burglaries; and,
- how slowly or quickly police would arrive at the scene if a violent crime were to occur near your house.

Distributive justice was measured using two questions. The introduction was: ‘Now some questions about whether or not the police in [country] treat victims of crime equally. Please answer based on what you have heard or your own experience.’ The first question was: ‘When victims report crimes, do you think the police treat rich people worse, poor people worse, or are rich and poor treated equally?’ The second question was: ‘when victims report crimes, do you think the police treat some people worse because of their race or ethnic group or is everyone treated equally?’ These two indicators were combined to form a single variable, with 0 equalling ‘neither poor nor minority group members are treated worse’, 1 equalling ‘either poor or minority group members are treated worse’ and 2 equalling ‘both poor and minority group members are treated worse’. Finally, lawfulness was measured by asking people (a) to agree or disagree with the statement ‘Decisions and actions of police are unduly influenced by political pressure’ on a 5-point Likert scale and (b) ‘How often do police [in your country] take bribes’ on a scale from 0 ‘never’ to 10 ‘always’. These two variables were rescaled, such that high scores equals the belief that the police act lawfully. A single index was created by taking the mean of the two variables (having divided the bribery variable by two, to put it on a comparable scale to the political pressure variable).

Note that a strength of the data is that the same measures were fielded in nationally representative sample surveys of 30 countries. A weakness is that the measures of distributive justice and lawfulness were limited to two each. The measures of distributive justice were nominal with three categories each, and the measures of lawfulness were one five category variable and one ten category variable. As such, it makes sense to produce a single derived variable for distributive justice and a single derived variable for lawfulness. In the context of confirmatory factor analysis, this does mean we can test a model exactly like Tankebe's (2013). But it is close enough because the CFA is still assessing, among other things, whether these are four empirically distinct constructs (Figure 6).

Note, also, that we do not fit a second-order factor. This is because the extra testable assumption at the heart of a second-order factor is unnecessary. It involves testing whether the correlations between the four latent constructs can be modelled according to a single underlying second-order latent construct, but as detailed above, if one finds that the second-order factor model fits, why would this increase one's confidence that the four constructs are constituent components of legitimacy rather than possible sources of legitimacy? The researcher is free to call the second-order factor whatever she likes. In addition, Sun et al. (2018) sought to replicate Tankebe (2013) and he presented legitimacy as multi-dimensional and did not fit a second-order factor.

## RESULTS

Table 1 provides the exact and approximate fit statistics for the CFA model (Figure 6) that differentiates between procedural justice (a latent construct with three indicators), effectiveness (another latent construct with three indicators), distributive justice (a manifest indicator calculated using responses to two indicators) and lawfulness (another manifest indicator calculated using responses to two indicators). We discount the exact fit statistics because the Chi-square tests is extremely sensitive to sample size (very small deviations in the match between the hypothesised and saturated models can be highlighted as statistically significant). The approximate fit statistics are adequate in all countries, with RMSEA < .08, CFI > .95 and TLI > .95 in every case.

Table 1 Fit statistics for the fitted confirmatory factor analysis model in each of the 30 countries<sup>14</sup>

Country	Chi-square	df	p-value	RMSEA 95% CI	CFI	TLI
Austria	149	16	<0.001	.061 [.052, .070]	.987	.978
Belgium	146	16	<0.001	.069 [.059, .080]	.956	.923
Bulgaria	110	16	<0.001	.050 [.041, .059]	.992	.986
Croatia	154	16	<0.001	.073 [.063, .083]	.973	.953
Cyprus	69	16	<0.001	.055 [.042, .069]	.991	.985
Czech Republic	157	16	<0.001	.061 [.063, .070]	.982	.968
Denmark	103	16	<0.001	.059 [.048, .070]	.968	.945
Estonia	91	16	<0.001	.051 [.042, .062]	.979	.963
Finland	78	16	<0.001	.046 [.036, .056]	.980	.965
France	120	16	<0.001	.061 [.051, .072]	.974	.954
Germany	143	16	<0.001	.051 [.044, .059]	.974	.955
Greece	99	16	<0.001	.044 [.036, .052]	.995	.991
Hungary	51	16	<0.001	.038 [.026, .050]	.989	.981
Ireland	142	16	<0.001	.055 [.047, .064]	.986	.975
Israel	184	16	<0.001	.068 [.060, .077]	.980	.965
Lithuania	74	16	<0.001	.047 [.036, .058]	.986	.976
Netherlands	129	16	<0.001	.062 [.052, .072]	.962	.933
Norway	131	16	<0.001	.068 [.058, .079]	.957	.924
Poland	83	16	<0.001	.049 [.039, .060]	.985	.974
Portugal	186	16	<0.001	.070 [.061, .080]	.973	.953
Russian Federation	174	16	<0.001	.062 [.054, .071]	.984	.971
Slovakia	139	16	<0.001	.065 [.055, .075]	.982	.969
Slovenia	93	16	<0.001	.059 [.047, .071]	.982	.968
South Africa	182	16	<0.001	.065 [.056, .073]	.986	.976
Spain	50	16	<0.001	.033 [.023, .044]	.994	.990
Sweden	101	16	<0.001	.060 [.049, .071]	.963	.935
Switzerland	70	16	<0.001	.047 [.036, .059]	.978	.962
Ukraine	95	16	<0.001	.051 [.041, .061]	.990	.982
United Kingdom	99	16	<0.001	.046 [.038, .055]	.988	.979
United States	96	16	<0.001	.056 [.046, .067]	.989	.981

<sup>14</sup> Data: Round 5 European Social Survey, Yale xxx Survey and SASAS 2010.

Table 2 provides details of the scaling properties of procedural justice and effectiveness in each country. Each cell gives the standardized factor loading (left) and R<sup>2</sup> (right) for each particular indicator (PJ1, PJ2, PJ3, EFF1, EFF2 and EFF3) in each of the 30 different countries. For procedural justice, the standardized factor loadings range from .59 to .94, and the R<sup>2</sup>s range from .35 to .88. For

effectiveness, the standardized factor loadings range from .44 to .90, and the R<sup>2</sup>s range from .20 to .80. Because the factors loadings and R<sup>2</sup>s are all relatively high, this indicates good scaling properties.

Table 2 Standardized factor loadings and R<sup>2</sup>s for the PJ and EFF indicators<sup>15</sup>

Country	PJ1	PJ2	PJ3	EFF1	EFF2	EFF3
Austria	.88 / .77	.93 / .87	.72 / .52	.80 / .65	.78 / .61	.78 / .60
Belgium	.78 / .61	.73 / .53	.65 / .42	.75 / .56	.70 / .49	.52 / .28
Bulgaria	.90 / .81	.93 / .87	.80 / .63	.85 / .72	.79 / .63	.71 / .51
Croatia	.84 / .71	.83 / .69	.78 / .60	.86 / .75	.83 / .69	.70 / .48
Cyprus	.88 / .77	.93 / .86	.80 / .64	.83 / .69	.79 / .63	.75 / .56
Czech Republic	.88 / .78	.88 / .77	.71 / .50	.78 / .61	.68 / .46	.70 / .50
Denmark	.77 / .59	.80 / .64	.70 / .48	.77 / .59	.53 / .28	.56 / .31
Estonia	.81 / .65	.81 / .65	.65 / .43	.70 / .49	.67 / .45	.61 / .37
Finland	.74 / .55	.80 / .64	.65 / .43	.70 / .49	.70 / .49	.54 / .29
France	.77 / .60	.84 / .70	.63 / .40	.77 / .59	.74 / .54	.63 / .39
Germany	.74 / .55	.82 / .67	.59 / .35	.67 / .45	.66 / .43	.55 / .30
Greece	.89 / .79	.94 / .88	.77 / .60	.86 / .74	.90 / .80	.81 / .66
Hungary	.82 / .68	.87 / .75	.62 / .38	.80 / .64	.75 / .56	.57 / .33
Ireland	.83 / .69	.91 / .82	.74 / .54	.80 / .63	.69 / .47	.66 / .43
Israel	.86 / .74	.86 / .73	.80 / .63	.78 / .60	.68 / .47	.61 / .37
Lithuania	.75 / .56	.83 / .68	.78 / .60	.83 / .69	.79 / .63	.64 / .41
Netherlands	.78 / .61	.79 / .62	.62 / .39	.71 / .50	.64 / .41	.56 / .32
Norway	.85 / .72	.68 / .47	.62 / .38	.73 / .53	.69 / .48	.62 / .39
Poland	.81 / .66	.84 / .71	.76 / .58	.81 / .65	.78 / .61	.65 / .42
Portugal	.77 / .60	.87 / .76	.79 / .62	.77 / .59	.78 / .61	.73 / .53
Russian Federation	.84 / .70	.89 / .80	.77 / .59	.81 / .66	.80 / .64	.70 / .48
Slovakia	.83 / .69	.91 / .83	.76 / .58	.83 / .69	.79 / .62	.65 / .43
Slovenia	.83 / .69	.92 / .85	.67 / .45	.80 / .63	.78 / .60	.62 / .38
South Africa	.87 / .76	.91 / .83	.79 / .62	.83 / .69	.78 / .61	.80 / .64
Spain	.84 / .71	.89 / .79	.70 / .49	.82 / .67	.76 / .57	.66 / .44
Sweden	.70 / .48	.73 / .53	.70 / .49	.76 / .58	.70 / .50	.56 / .31
Switzerland	.76 / .58	.76 / .58	.68 / .46	.74 / .55	.78 / .61	.44 / .20
Ukraine	.83 / .69	.92 / .84	.84 / .70	.84 / .70	.84 / .70	.75 / .56
United Kingdom	.86 / .73	.87 / .76	.70 / .49	.75 / .56	.71 / .50	.64 / .41
United States	.92 / .84	.79 / .63	.84 / .70	.85 / .71	.73 / .53	.72 / .52

<sup>15</sup>Data: Round 5 European Social Survey, Yale xxx Survey and SASAS 2010.

Table 3 provides the correlations between the four constructs. Looking across the 30 countries, the pair of constructs with the strongest correlation is procedural justice and effectiveness (ranges from .45 to .77) and the pair of constructs with the weakest correlation is distributive justice and lawfulness (ranges from .06 to .42).

Table 3 Correlations between constructs from the fitted measurement model<sup>16</sup>

Country	PJ WITH EFF	PJ WITH DJ	PJ WITH LF	EFF WITH DJ	EFF WITH LF	DJ WITH LF
Austria	.63	.54	.43	.47	.42	.42
Belgium	.51	.32	.36	.23	.28	.19
Bulgaria	.76	.49	.55	.47	.52	.42
Croatia	.71	.54	.48	.49	.40	.32
Cyprus	.70	.44	.52	.41	.46	.34
Czech Republic	.70	.46	.53	.39	.39	.31
Denmark	.58	.38	.27	.35	.25	.22
Estonia	.67	.39	.47	.34	.41	.35
Finland	.52	.32	.38	.27	.21	.25
France	.64	.45	.40	.30	.38	.32
Germany	.55	.35	.32	.36	.29	.28
Greece	.77	.40	.56	.37	.49	.33
Hungary	.65	.47	.51	.40	.38	.37
Ireland	.68	.41	.43	.42	.39	.33
Israel	.71	.40	.32	.38	.27	.19
Lithuania	.69	.42	.45	.41	.36	.29
Netherlands	.45	.38	.29	.26	.16	.19
Norway	.53	.31	.31	.29	.16	.17
Poland	.63	.46	.47	.41	.41	.37
Portugal	.51	.41	.37	.44	.47	.34
Russian Federation	.75	.42	.41	.40	.36	.25
Slovakia	.64	.39	.45	.40	.42	.29
Slovenia	.64	.43	.42	.43	.36	.36
South Africa	.54	.43	.33	.41	.27	.24
Spain	.68	.45	.45	.40	.41	.34
Sweden	.55	.33	.29	.34	.09	.18
Switzerland	.52	.36	.34	.30	.25	.29
Ukraine	.69	.30	.31	.32	.25	.06
United Kingdom	.62	.38	.35	.37	.33	.28
United States	.67	.48	.61	.34	.49	.40

<sup>16</sup>Data: Round 5 European Social Survey, Yale xxx Survey and SASAS 2010.

Applying the reasoning of Tankebe (2013) and Sun et al. (2018) we could interpret the fact that the four-factor model fitted the data in each of the 30 countries as evidence that what have previously treated as possible sources of legitimacy are in fact constituent components of legitimacy. This would imply that legitimacy rests on the same normative bases in Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Lithuania, Netherlands, Norway, Poland, Portugal, Russian Federation, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, UK, Ukraine and US. It follows that in each of these diverse social, political and legal contexts, the police need to be seen to act in ways that are procedurally just, distributively just, effective and lawful if they are to be seen as legitimate.

Indeed, it is not that procedural justice, distributive justice, effectiveness and lawfulness are all strong predictors of legitimacy in each of the 30 countries, it is that they are so fundamental to the perceived right to power that they collectively constitute the construct. This would contradict existing work that (a) assumes that procedural justice is not the same as (or not part of) legitimacy, (b) finds that procedural justice is a stronger predictor of legitimacy than distributive justice, effectiveness or lawfulness, and (c) highlights country-level differences in the extent to which each explains variation in legitimacy. We return to this point in the discussion below.

## DISCUSSION

We began this commentary by describing how the empirical concept of legitimacy (Hinsch 2008) specifies the right to power as a property of public opinion (e.g. people in Japan may view their police to be more legitimate than by people in Russian Federation) and treats it as an empirical question the criteria that citizens of a given country use to judge the legitimacy of the institution (e.g. people in Japan may judge the legitimacy of the police according to a different set of criteria compared to people in Russian Federation). This is in contrast to the normative concept of legitimacy, which involves an outsider observer determining the substantive requirements for legitimacy in different contexts (e.g. police in both Japan and Russian Federation may be judged by the outside expert according to the same criteria like, for instance, independence, accountability and other indicators of the rule of law).

We then described the methodology used by Sun et al. (2018) and Tankebe (2013) to test the validity and reliability of a new approach to measuring empirical police legitimacy. In Sun et al.'s (2018) study the scales of procedural justice, distributive justice, effectiveness and lawfulness had good measurement properties; they loaded on four strongly correlated latent variables (labelled *procedural justice*, *distributive justice*, *lawfulness* and *effectiveness*); and these four latent constructs were themselves regressed onto a single second-order factor that the

researchers *a priori* labelled *legitimacy*. The researchers subsequently argued that, on this evidence base, they are such strong criteria on which people use to judge police legitimacy that they can be treated as constituent components of legitimacy.

We then took the reader through a hypothetical reanalysis of Sun et al.'s (2018) data. Taking a different starting point but following the same logical sequence, we reached the conclusion that what has previously been viewed as four constituent components of legitimacy in this coastal Chinese city (procedural justice, distributive justice, effectiveness and lawfulness) are in fact possible sources of legitimacy. The scales of procedural justice, distributive justice, effectiveness and lawfulness had good measurement properties; they loaded on four strongly correlated latent variables (labelled *procedural justice*, *distributive justice*, *lawfulness* and *effectiveness*); and these four latent constructs were themselves regressed onto a single second-order factor that you *a priori* labelled *possible sources of legitimacy*. This thought experiment illustrated the simple point that the findings of the CFA modelling do not tell whether one is right *in the first place* to define procedural justice, distributive justice, effectiveness and lawfulness as (a) potential normative criteria that people use to judge legitimacy or (b) actual constituent components of legitimacy. This is not a good adjudication tool because the same fitted model is consistent with both conceptual stances.

We then investigated whether the approach generally lacks cultural sensitivity. We linked Round 5 of the European Social Survey (ESS) to two matching representative sample surveys of US and South Africa. This produces a 30-country data dataset spanning countries as diverse as Croatia, Denmark, France, Greece, Israel, South Africa, Russian Federation, Spain, UK, US and Ukraine. Analysing data from these 30 countries to test whether Tankebe's (2013) four-factor model fitted in each social, political and legal context, we found that the model (specified in Figure 6) did indeed fit the data in each of the 30 countries. In each and every country, measures of procedural justice, distributive justice, effectiveness and lawfulness scaled reasonably well; reflected or formed four empirically distinct constructs; and were strongly and positive correlated with each other. Applying the reasoning of Sun and colleagues, this would imply that legitimacy is, in each and every one of those countries, the same *thing*, i.e. it is comprised of public assessments of procedural justice, distributive justice, effectiveness, and lawfulness. By extension, the legitimating norms that people expect police to abide by (before the institution is to be viewed as legitimate) are the same in all countries and relate to these four areas.

We believe this would be a misfounded view. The fact that the model fitted the data well in each country says little about whether we are measuring, accurately or not, legitimacy. The claim that we are doing so is a purely conceptual matter based on an *a priori* assumption about what constitutes legitimacy in a particular context and how it can reasonably be operationalised. Moreover, because the approach ends up imposing onto 30 different social, political and legal contexts the idea that that people in each different country judge the legitimacy of the

police on roughly speaking the same bases (procedural justice, distributive justice, effectiveness and lawfulness), we have what is rather curious mix of the normative concept and the empirical concept of legitimacy. Public opinion matters with regard to levels of legitimacy in a given society, but it is the researcher who is imposing the substantive requirements for empirical legitimacy (i.e. procedural justice, distributive justice, effectiveness and lawfulness) in each and every context.

This is where the problem arises. As shown in the analysis of the ESS data, this would lead to the rather odd conclusion that legitimacy is constituted *in the same way* in UK, US, Russia, South Africa and Israel, to say nothing of the other countries shown. In other words, in each of these diverse countries the claim would be that as long as someone believes the police are procedurally and distributively just, effective, and behave in a lawful manner, then they believe the police to be legitimate. In practice, therefore, far from being sensitive to cultural variation in the composition of legitimacy, the model proposed by Tankebe, Sun and colleagues flattens out the possibility of variation because it assumes beforehand that these judgements concern procedural justice, distributive justice, effectiveness and lawfulness.

There are three obvious problems with such a proposition. First, one could take a quite reasonably normative view that police should act in procedurally just, distributively just, effective and lawful ways, and moreover that being seen to do so is so fundamental that their perceived right to power rests directly on of the four public perceptions. If policy makers in a given country do recognise the importance of being seen by those they serve, protect and police as legitimate, then framing these as constituent components of legitimacy could initially be an effective rhetorical device. But why would one need a potentially spurious reason to argue that police should act in procedurally just, distributively just, effective and lawful ways when a more convincing general normative case can be made? Moreover, studies showing, for instance, that effectiveness is not a particularly important predictor of legitimacy (defined and measured as a more general perception of normative appropriateness) could subsequently undermine the empirical argument.

Second, and relatedly, there is little or no possibility of assessing which if any is the most important component of legitimacy. What type or aspect of police behaviour is most important in generating a sense among the policed that police activity is normatively justifiable? Do people value procedural justice most? Or are they more concerned with effectiveness? These seem to us important questions, both theoretically and from a policy perspective. Yet the approach to measuring legitimacy taken above makes it difficult if not impossible to answer them. If one wants to test whether people judge the legitimacy of the police on different criteria according to local societal norms, values, institutional structures and context, one needs a strategy that has the requisite cultural sensitivity.

The third problem with the approach taken above is that it leaves no room for the fact that other judgements might come into play when people are thinking about the normative appropriateness of police activity. These might be many and



varied, and some might be morally troubling from a normative perspective. Some white US citizens, for example, might believe police are behaving appropriately when they target black US citizens, not because they think this makes policing more effective or fair – although they might also believe this – but because they have been socialised or otherwise come to believe this is just the way *police should behave*. Given the history of the US *vis-à-vis* many other liberal democracies (Alexander 2012), there is no necessary reason to assume this would be the case elsewhere, although of course it might be. This is, to our minds, an empirical question worthy of investigation. But the model specified above renders such investigation analytically and conceptually difficult. One would have to assume, for example, that racially targeted policing in a given context is only generative of legitimacy to the extent that it influences beliefs about fairness, effectiveness and lawfulness. If it does not, then it has no effect on legitimacy. In fact, nothing could have an effect on legitimacy if it did not have an effect on at least one of the four constructs.

The answer to all these problems is obvious and is already employed in much of the literature. If legitimacy is conceptualized and measured as something distinct from assessments of fairness, effectiveness and lawfulness, then it is possible to assess which if any of these is most important as a predictor of (formative judgements constituting) legitimacy. Similarly, if legitimacy is distinct and different from the four factors, then the influence of other variables is conceptually and analytically far easier to assess, since other judgements of the normative appropriateness of police activity are allowed to have effects distinct from any correlation with perceptions of fairness, effectiveness and lawfulness. For instance, respecting the limits of one's rightful authority may be important to legitimacy above and beyond perceptions of the procedural justice of the police (Huq et al. 2017; Trinkner et al. 2018).

To close, we do think there is space for alternative approaches to measuring legitimacy. Legitimacy is an abstract and unobservable psychological construct, and there are numerous ways to operationalise the perceived right to power, aside from the standard ways of institutional trust and/or normative alignment and/or obligation to obey (Tyler & Jackson 2013). But researchers need to be aware of the limits of different approaches when trying to decide which approach to measurement is preferable. In the current case, the methodology of Sun and colleagues means ironically imposing *top-down* rather than discovering *bottom-up* the normative content of legitimacy. There may be normative and/or theoretical and/or practical reasons to decide this but it is, in our view, crucial to be clear that this is what one is doing, especially when the analytical strategy offers little empirical support for the central conceptual claim.

**REFERENCES**

- Akinlabi, O. M. & Murphy, K. (2018). Dull compulsion or perceived legitimacy? Assessing why people comply with the law in Nigeria. *Police Practice and Research*, online advance access.
- Bradford, B. (2014). Policing and social identity: Procedural justice, inclusion and cooperation between police and public. *Policing and Society*, 24(1), 22-43.
- Bradford, B., Murphy, K. & Jackson, J. (2014). Officers as mirrors: Policing, procedural justice and the (re)production of social identity. *British Journal of Criminology*, 54(4), 527-550.
- Bradford, B., Huq, A., Jackson, J., & Roberts, B. (2014). What price fairness when security is at stake? Police legitimacy in South Africa. *Regulation & Governance*, 8(2), 246-268.
- Caldeira, G. A. & Gibson, J. L. (1995). The legitimacy of the Court of Justice in the European Union: Models of institutional support. *American Political Science Review*, 89(2), 356-376.
- Dirikx, A., & Van den Bulck, J. (2013). Media use and the process-based model for police cooperation: An integrative approach towards explaining adolescents' intentions to cooperate with the police. *British Journal of Criminology*, 54(2), 344-365.
- European Social Survey (2011). Trust in justice: Topline findings from the European Social Survey, ESS Topline Results Series Issue 1. By Jackson, J., Hough, M., Bradford, B., Pooler, T. M., Hohl, K. & Kuha, J.
- European Social Survey Round 5 (2010). Data file edition 3.4. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
- European Social Survey Round 5 (2018): ESS-5 2010 Documentation Report. Edition 4.2. Bergen, European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.
- Gibson, J. L., Caldeira, G. A. & Spence, L. K. (2003). Measuring attitudes toward the United States supreme court. *American Journal of Political Science*, 47(2), 354-367.
- Hinsch, W. (2008). Legitimacy and justice, in Kuhnelt, J. (ed.) *Political legitimization without morality?* London: Springer.
- Hinsch, W. (2010). Justice, legitimacy, and constitutional rights. *Critical Review of International Social and Political Philosophy*, 13, 39-54.
- Hough, M., Jackson, J. & 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.) *Routledge Handbook of European Criminology*, Oxon: Routledge, pp. 243-265.
- Huq, A. Z., Jackson, J. & Trinkner, R. (2017). Legitimizing practices: Revisiting the predicates of police legitimacy. *British Journal of Criminology*, 57(5), 1101-1122.
- Huq, A. Z., Tyler, T. R., & Schulhofer, S. J. (2011). Mechanisms for eliciting

- cooperation in counterterrorism policing: A study of British Muslims. *Journal of Empirical Legal Studies*, 8, 728–761.
- Jackson, J. (2018). Norms, normativity and the legitimacy of legal authorities: International perspectives. *Annual Review of Law and Social Science*, 14, 145-165.
- Jackson, J., Asif, M., Bradford, B. & Zakar, M. Z. (2014). Corruption and police legitimacy in Lahore, Pakistan. *British Journal of Criminology*, 54, 1067–1088.
- Jackson, J., Bradford, B., Hough, M., Kuha, J., Stares, S. R., Widdop, S., Fitzgerald, R., Yordanova, M. & Galev, T. (2011). Developing European indicators of trust in justice. *European Journal of Criminology*, 8(4), 267-285.
- Jackson, J. & Gau, J. (2015). Carving up concepts? Differentiating between trust and legitimacy in public attitudes towards legal authority, in Shockley, E., Neal, T. M. S., PytlikZillig, L. and Bornstein, B. (eds.) *Interdisciplinary perspectives on trust: Towards theoretical and methodological integration*. New York: Springer, pp. 49-69.
- Jowell, R., Roberts, C., Fitzgerald, R. & Eva, G. (eds.) (2007). *Measuring attitudes cross-nationally: Lessons from the European Social Survey*. London: Sage.
- Justice, B. & Meares, T. (2014). How the criminal justice system educates citizens. *Annals of the American Academy of Political and Social Sciences*, 651(1), 159–177.
- Kim, Y. S., Ra, K. H. & McLean, K. (2018). The generalizability of police legitimacy: Procedural justice, legitimacy, and speeding intention of South Korean drivers. *Asian Journal of Criminology*, advance online access.
- Liebling, A. (2004). *Prisons and their moral performance*. Oxford: Oxford University Press.
- Meares, T. (2017). Policing and procedural justice: shaping citizens' identities to increase democratic participation. *Northwestern University Law Review*, 111(6), 1525–1536.
- Mentovich, A., Ben-Porat, G., Levy, N., Goff, P. A. & Tyler, T. (2018). Policing alienated minorities in divided cities. *Regulation & Governance*, advance online access.
- Murphy, K. and Cherney, A. (2012). Understanding cooperation with police in a diverse society. *British Journal of Criminology*, 52(1), 181-201.
- Murphy, K., Cherney, A. & Teston, M. (2018). Promoting Muslims' willingness to report terror threats to police: Testing competing theories of procedural justice. *Justice Quarterly*, advance online access.
- Pardo, I. (2000). *Morals of Legitimacy: Between Agency and System*. New York: Berghahn Books.
- Reisig, M. D., Bratton, J. & Gertz, M. G. (2007). The construct validity and refinement of process-based policing measures. *Criminal Justice Behavior*, 34, 1005–1027.
- Sun, I. Y., Wu, Y., Hu, R. & Farmer, A. K. (2017). Procedural justice, legitimacy, and public cooperation with police: Does Western wisdom hold in China?

- Journal of Research in Crime and Delinquency, 54(4), 454–478.
- Sun, I. Y., Wu, Y. & Hu, R. (2013). Public assessments of the police in rural and urban China: A theoretical extension and empirical investigation. *British Journal of Criminology*, 53(4), 643–664.
- Sun, I. Y., Li, L., Wu, Y., & Hu, R. (2018). Police legitimacy and citizen cooperation in China: testing an alternative model. *Asian Journal of Criminology*, 13(4), 275-291.
- Sunshine J. & Tyler, T. R. (2003). The role of procedural justice and legitimacy in public support for policing. *Law & Society Review*, 37, 513–548.
- Tankebe J. (2009). Public cooperation with the police in Ghana: Does procedural fairness matter? *Criminology*, 47, 1265–1293.
- Tankebe J. (2013). Viewing things differently: The dimensions of public perceptions of legitimacy. *Criminology*, 51, 103–135.
- Trinkner, R., Jackson, J. & Tyler, T. R. (2018). Bounded authority: Expanding “appropriate” police behavior beyond procedural justice. *Law & Human Behavior*, 42(3), 280-293.
- Tsushima, M. & Hamai, K. (2015). Public cooperation with the police in Japan testing the legitimacy model. *Journal of Contemporary Criminal Justice*, 31(2), 212–228.
- Tyler, T. R. & Jackson, J. (2013). Future challenges in the study of legitimacy and criminal justice, in Tankebe, J. and Liebling, A. (eds.) *Legitimacy and Criminal Justice: An International Exploration*. Oxford: Oxford University Press, pp. 83-104.
- Tyler T. R. & Jackson J. (2014). Popular legitimacy and the exercise of legal authority: Motivating compliance, cooperation and engagement. *Psychology, Public Policy and Law*, 20, 78–95.
- Tyler, T. R., Jackson, J. & Mentovich, A. (2015). On the consequences of being a target of suspicion: Potential pitfalls of proactive police contact. *Journal of Empirical Legal Studies*, 12(4), 602-636.
- Tyler, T. R. (2006a). Legitimacy and legitimation. *Annual Review of Psychology*, 57, 375-400.
- Tyler, T. R. (2006b). *Why people obey the law*, Princeton: Princeton University Press.
- White, M. D., Mulvey, P. & Dario, L. M. (2016). Arrestees’ perceptions of the police: Exploring procedural justice, legitimacy, and willingness to cooperate with police across offender types. *Criminal Justice and Behavior*, 43, 343-364.