SOCIAL TRUST IN CRIMINAL JUSTICE: A METRIC

Joshua Kleinfeld* & Hadar Dancig-Rosenberg**

What is the metric by which to measure a well-functioning criminal justice system? If a modern state is going to measure performance by counting something—and a modern state will always count something—what, in the criminal justice context, should it count? Remarkably, there is at present no widely accepted metric of success or failure in criminal justice. Those there are—like arrest rates, conviction rates, and crime rates—are deeply flawed. And the search for a better metric is complicated by the cacophony of different goals that theorists, policymakers, and the public bring to the criminal justice system, including crime control, racial justice, retributive justice, and social solidarity.

This Article proposes a metric based on the concept of social trust. The measure of a well- or poorly functioning criminal system is its marginal effects on (1) the level of trust a polity's members have toward the institutions, officials, laws, and actions that comprise the criminal justice system; (2) the level of trust a polity's members have, in virtue of the criminal system's operations, toward government generally (beyond the criminal justice system); and (3) the level of trust a polity's members have toward one another following incidents of crime and responses to crime. Social trust, we argue, both speaks to an issue at the philosophical core of crime and punishment and serves as a locus of agreement among the many goals people bring to the criminal justice system. The concept can thus be a site of overlapping consensus, performing the vital function of enabling liberal societies to make policy despite disagreement about first principles.

---

* Professor of Law and (by courtesy) Philosophy, Northwestern University.

** Professor of Law & Associate Dean for Research (2019–21), Bar-Ilan University; Helen Diller Institute Visiting Professor, UC Berkeley School of Law (2021–23); Co-chair of the Israeli Criminal Law Association.

INTRODUCTION

What is the metric by which to measure a well-functioning criminal justice system? If a modern state is going to measure performance by counting something—and a modern state will always count *something*—what, in the criminal justice context, should it count?

Governments at present commonly focus on crime rates, recidivism rates, conviction rates, arrest rates, clearance rates, and cost. As we argue below, those metrics, while important and illuminating for certain purposes, are misleading in terms of what they are commonly taken to show and pernicious in terms of the incentives they create. Conviction rates, for example, are typically used to assess prosecutors' performance: they measure the proportion of people found guilty once prosecutors have decided to file charges. The effect is to incentivize prosecutors to focus on easy-to-prove cases rather than socially serious cases and, since the most certain conviction is a plea deal, to encourage excessive and sometimes unsavory forms of plea bargaining.[1] Another example: crime rates are commonly used to evaluate

---

1   *See* David Alan Sklansky, *The Changing Political Landscape for Elected Prosecutors*, 14 OHIO ST. J. CRIM. L. 647, 649, 669 (2017).

police forces,[2] yet crime rates are affected by such a complex array of societal factors beyond policing (broken families, poverty, poor mental health services, etc.) that it is perfectly possible for a community to have high crime notwithstanding good policing or low crime notwithstanding bad policing. Crime rate metrics can also incentivize harshness, since the intuitive way to prevent crime, at least in the short term, is to give offenders long sentences.

We are not the first to notice this problem, and in response to it a scholarly literature has formed, largely in empirical journals, around finding something better. Yet the question of what to measure in criminal justice is fundamentally a philosophical problem. It turns on what constitutes the *excellence* of a criminal system, even what constitutes *justice* in a criminal system—matters native to the theory of punishment. And therein lies the rub. The empiricists typically do not engage with the philosophical problem, and so construct metrics that nip around the edges of considerations that should matter. And the philosophers are commonly uninterested in problems of measurement, and so ignore the bureaucratic conditions of modern life—the fixation with quantitative metrics and, consequently, the power of metrics to structure incentives, determine the allocation of resources, and redefine what governments care about or even, in a sense, "see."[3] Given these conditions, which prevail with particular force in the institutional and bureaucratic landscape of modern criminal systems, a good metric might have more impact on justice and even, potentially, more interest as a matter of first principles, than any statement of values abstractly understood. The real goal is to translate values into bureaucratic operating procedures; the "killer app" of modern life is the bureaucratic operationalization of values. But the philosophers are on the whole not communicating with the empiricists about what to measure.

The goal of this Article is to propose a certain concept of social trust and offer an argument as to why social trust, so understood, should be the north star in measuring the success or failure of criminal justice institutions. Social trust, as we discuss in Section II.A below, is anything but a novel concept. It refers broadly to an individual's beliefs about the general trustworthiness of other people, social institutions, and government: on a large scale, it refers to ambient levels of trust within a society toward other citizens, toward institutions, and

---

2   *See* Richard Rosenfeld & Joel Wallman, *Did De-Policing Cause the Increase in Homicide Rates?*, 18 CRIMINOLOGY & PUB. POL'Y 51, 67 (2019).

3   *Cf.* JAMES C. SCOTT, SEEING LIKE A STATE: HOW CERTAIN SCHEMES TO IMPROVE THE HUMAN CONDITION HAVE FAILED (Veritas paperback ed. 2020) (1998) (examining the way governments use metrics to make the social world legible from a bureaucratic point of view, distorting the improvisational character of ungoverned social life and, because governmental metrics affect the flow of power, altering what they were meant only to measure).

toward government generally. It has been extensively studied in economics, political science, sociology, and psychology, and it has proven to be one of the best predictors of overall societal well-being known to social science, with remarkable and well-documented effects on economic success, effective government, levels of participation in civil society, and compliance with the law. Social trust is of particular interest now, in our politically polarized era, as it seems to be the very thing partisan polarization destroys. Others in the fledgling literature on criminal justice metrics have started to explore it or concepts related to it.[4] And yet those efforts have been hampered by an inadequate understanding of *why* social trust matters in criminal justice, and thus how best to conceptualize it and approach its measurement. Our goal in this Article is therefore to justify social trust as the lodestar of criminal justice metrics, and, in justifying it, to clarify the concept in ways that could contribute to future empirical work.

As we use the term in the context of criminal justice, social trust refers to (1) the level of trust a polity's members have toward the institutions, officials, laws, and actions that comprise the criminal justice system; (2) the level of trust a polity's members have, in virtue of the criminal system's operations, toward government generally (beyond

---

4   A small number of scholars have thought about trust in connection to criminal systems, but usually in specific contexts (e.g., the efficacy of prisons or the theory of retributivism), and never as a metric by which to measure criminal justice outcomes. *See, e.g.*, DAVID BOONIN, THE PROBLEM OF PUNISHMENT 143–49 (2008) (citing Daniel Korman, *The Failure of Trust-Based Retributivism*, 22 LAW & PHIL. 561, 562 (2003)) (discussing trust-based conceptions of retributivism); LAWRENCE W. SHERMAN, TRUST AND CONFIDENCE IN CRIMINAL JUSTICE 14–15 (2002) (citing Tom R. Tyler, *Trust and Democratic Governance*, *in* TRUST AND GOVERNANCE 269, 269–94 (Valerie Braithwaite & Margaret Levi eds., 1998)) (noting the relevance of the concept of "social trust" to criminal justice); ALEKSANDAR FATIĆ, PUNISHMENT AND RESTORATIVE CRIME-HANDLING: A SOCIAL THEORY OF TRUST (1995); Jim Staihar, *Trust in Enforcement: A Unified Theory of Sanctioning Certified Public Accountants*, 26 J.L. BUS. & ETHICS 19, 22–23 (2020) (advancing a trust-restoration-based normative theory of the sanctioning of Certified Public Accountants); Jim Staihar, *Proportionality and Punishment*, 100 IOWA L. REV. 1209, 1216–23 (2015) (describing an unfair advantage theory of punitive desert that assumes one's unexcused criminal act reduces one's trustworthiness); Jim Staihar, *Punishment as a Costly Signal of Reform*, 110 J. PHIL. 282, 284 (2013) (arguing that the state should provide offenders with an opportunity to undertake punishment as a means to signaling their reform, in order to achieve the "expected benefits from restoring [their] trustworthiness"); Alison Liebling, *'Legitimacy Under Pressure' in High Security Prisons*, *in* LEGITIMACY AND CRIMINAL JUSTICE: AN INTERNATIONAL EXPLORATION 206, 206–27 (Justice Tankebe & Alison Liebling eds., 2013); Jim Staihar, *A New Systematic Explanation of the Types and Mitigating Effects of Exculpatory Defenses*, 12 NEW CRIM. L. REV. 205, 205 (2009) (arguing that one's blameworthiness and desert corresponds to the severity of the burden one must undertake to restore the conditions of trust undermined by a criminal act); R.A. Duff, *Who is Responsible, for What, to Whom?*, 2 OHIO. ST. J. CRIM. L. 441, 458 (2005) (noting the potential causal effect between crimes victimizing individuals and "social volatility" or "loss of trust"); Susan Dimock, *Retributivism and Trust*, 16 LAW & PHIL. 37, 39 (1997).

the criminal justice system); and (3) the level of trust a polity's members have toward one another following incidents of crime and responses to crime. The goal of measurement in criminal justice, we submit, is to construct a tool by which to measure people's responses to the operation of the criminal system along these three dimensions. For example—and it is just an example—someone who is a crime victim, witness, community stakeholder, or even offender might get a survey in the wake of a contact with the criminal system asking whether the contact increased or decreased her sense of trust in the criminal system, in government generally, and in other people. Our central claim is that a well-functioning criminal system is one whose operation increases people's sense of trust along those three lines; a poorly functioning criminal justice system is one whose operation diminishes trust along one or more of those three lines. Marginal effects on social trust are the key measure of whether a criminal system is functioning well or poorly. Thus, in criminal justice, the output that is most important to measure and incentivize is something affective and interpersonal, not material.

These claims are not as absolute as they might sound, for two reasons discussed at length below. First, we are not arguing that other metrics should be done away with; the way social trust should "fit" into other metrics is more nuanced than that. Crime control, for example, is indispensable to measure; it just turns out that crime control depends on social trust to such an extent that, if one wants to minimize crime, there is good reason to measure trust. Second, we are not arguing for pure maximization of social trust. Pure maximization runs into problems of both the distribution of trust (for example, it can be dysfunctional to increase average trust if doing so creates radical distrust within a minority group) and non-trust-based constraints of justice (for example, scapegoating and show trials would violate basic principles of justice even if they were secret enough to increase trust in the short term). We will later argue that the distribution of trust matters, especially in conditions of an alienated minority, and that principles of basic justice constrain the maximization of trust. Those caveats in place, however, we *do* claim that social trust is sufficiently central to understanding what makes a criminal justice system function well that a metric based on social trust should take pride of place among all metrics.

A criminal justice metric based on social trust would have nine virtues, which we preview here and defend below, and which set it apart from any other available metric.

First, a metric based on social trust would radically alter criminal justice officials' incentives. Rather than encouraging excessively harsh policing and punishment practices in order to increase arrest or

conviction rates, or excessively mild policing and punishment practices in order to cover up rising crime or curry favor with interest groups, criminal justice officials would have reason to handle crime and punishment in ways that victims, community members, and even offenders themselves find sensible given local community standards. Police officers, prosecutors, prison wardens, and others would be incentivized to care whether societal stakeholders would view their decisions as just, effective, and fair. The result, we predict, would be a criminal system that substantially reflects local views of what it means to be moral and effective in responding to crime. In short, *a metric based on social trust would align official incentives with community norms.*

Whether that is a good or bad thing depends to some extent on what one thinks of the local community's norms. But it does not *entirely* turn on that question. For one thing, the question isn't just the merits and demerits of community norms but what the alternative would be if community norms do *not* prevail (e.g., bureaucratic imperatives or special interest group ideologies). For another, reflecting community norms has positive downstream effects even in conditions of flawed norms. It overcomes in a single gesture the current array of bureaucratic imperatives, like certainty, punitiveness, and cost savings, that currently structure incentives. It makes official behavior subject to moral thinking and democratic influence rather than purely instrumental thinking and insider influence. Above all, it makes officials treat criminal justice as an instrument of social solidarity. If you, like us, think social solidarity is central to punishment's goals, to do this is to accomplish something deeply important: aligning criminal justice officials' incentives with the goals of punishment. Perhaps the best thing any social institution can do is to align agents' incentives with the goals of the institution. But even if you do not share our solidaristic conception of punishment's goals, the capacity of a metric based on social trust to overcome the dysfunction of current incentives might yet be valuable.

Second, a metric based on social trust would indirectly but substantially contribute to a low crime rate. A striking array of different types of empirical research about the causes of crime, including work in psychology, sociology, history, and comparative law, have come to a common conclusion: public perceptions of legitimacy, based on whether the public thinks the law is just, fair, sensible, and grounded in rightful authority, are key to legal compliance.[5] Furthermore, it

---

5   TOM R. TYLER, WHY PEOPLE OBEY THE LAW 161 (2006); Anjuli Van Damme & Lieven Pauwels, *Explaining and Preventing Distrust in the Criminal Justice System—a Combined Quantitative and Qualitative Approach*, *in* SOCIAL ANALYSIS OF SECURITY 258, 267 (Paul Ponsaers ed., 2012); Tracey L. Meares & Dan M. Kahan, *When Rights Are Wrong: The Paradox of Unwanted Rights*, *in* URGENT TIMES: POLICING AND RIGHTS IN INNER-CITY COMMUNITIES 3,

does not appear to be the case that the two are merely correlated—that trust goes up in low-crime conditions and down in high-crime conditions. The evidence is that people's levels of trust affect their propensity to commit crimes; in a phrase, *increasing trust decreases crime.* Indeed, another benefit of steering the criminal system toward community norms—another benefit that does not depend on whether the norms are flawed—is that doing so tends to increase *legitimacy*, which in turn decreases crime.

Third, from the standpoint of criminal justice communitarians—including advocates of reconstructivism, restorative justice, and nonadversarial responses to crime—social trust speaks to the very foundations of crime and punishment. According to reconstructivists, for example, "criminal law and procedure have a distinctive role to play in the social world: where a wrong has been committed that is of such a nature as to attack the values on which social life is based, it is the office of criminal law to reconstruct that violated normative order."[6] It is here that one finds theorists who, like the two of us, regard social solidarity as one of the *goals* of punishment: "The measure of success in criminal justice is social solidarity around a shared moral culture; the measure of failure is alienation and normative disintegration."[7] Likewise, restorative justice theorists conceptualize crime in terms of injury to relationships and hold that the response to crime should aim to repair those relationships.[8] And community courts aim to spur rehabilitation through a sense of belonging.[9] The key is to see that, although theorists in this communitarian family tend to express their ideas in philosophical language ("shared moral culture" and the like), that philosophical language has a *correlate* in social science, and trust is that correlate. It is, to be sure, an approximate correlate; communitarian ideas aren't simply reducible to trust. But the fit is close enough to treat trust as the appropriate metric for all such theories.

Fourth, social trust should appeal to retributivists concerned about individual responsibility, impunity, and proportionality. There is no obvious way to measure deontological justice, but decades of

---

21–22 (Joshua Cohen & Joel Rogers eds., 1999); *see also* Faith E. Gifford & Michael D. Reisig, *A Multidimensional Model of Legal Cynicism*, 43 LAW & HUM. BEHAV. 383, 383 (2019) (collecting studies that show how social control agents exercise authority influences legal cynicism, which in turn, influences criminal activity).

    6   Joshua Kleinfeld, *Three Principles of Democratic Criminal Justice*, 111 NW. U. L. REV. 1455, 1457–58 (2017).

    7   *Id.* at 1462.

    8   HOWARD ZEHR, CHANGING LENSES: A NEW FOCUS FOR CRIME AND JUSTICE 181 (1990).

    9   *See* Tali Gal & Hadar Dancig-Rosenberg, *"I Am Starting to Believe in the Word 'Justice'": Lessons from an Ethnographic Study on Community Courts*, 68 AM. J. COMP. L. 376, 409–10 (2020).

empirical research show that Americans across all demographic divisions (along with people from many other societies as well) share fundamentally retributive assumptions about justice. To a staggering extent, Americans want the law to punish in accord with relative blameworthiness, which they consider a basic requirement of justice, agree about the relative blameworthiness of different crimes, and, crucially, favor much less draconian levels of punishment than American law currently prescribes, provided the issues are presented outside of the kind of bumper-sticker political context that leads people simply to declare their partisan identities.[10] In a society with those background attitudes, a criminal system that spurs social trust will tend to favor the things retributivists want. Now, this might not give retributivists *everything* they would want. There is a relativism and a consequentialism built into the DNA of social trust that is at odds with the absolutist and deontological spirit of retributive justice. (We discuss the relativism at length below, particularly the problem of a metric based on social trust in a society with bad social norms.) But, given both the empirics and the dearth of alternatives (again, there is no obvious way to measure "justice"), social trust should track retributive values closely enough to satisfy pragmatically minded retributivists. Social trust is essentially a good *proxy* for considerations of justice as the community understands the term.

Fifth, a metric based on social trust would build concern for racial injustice into the very fabric of the criminal system. At present, concerns over racism and brutality have led to exceptionally low levels of trust toward the criminal system among black Americans and other minorities,[11] but the metrics by which governments measure criminal justice outcomes render those effects invisible. A criminal system fixated on conviction rates, arrest rates, recidivism rates, and crime rates does not "see" the consequences of its actions for minorities' sense of solidarity with the American project. An arrest based on racism or perceived as based on racism still counts as an arrest; a racist prosecution that leads to conviction still counts as a win. The destructive social

---

10   *See generally* Tracey Meares, *Policing and Procedural Justice: Shaping Citizens' Identities to Increase Democratic Participation*, 111 Nw. U. L. Rev. 1525, 1531 (2017) (summarizing decades of research on empirical tests of intuitions about procedural justice); Paul H. Robinson, *Democratizing Criminal Law: Feasibility, Utility, and the Challenge of Social Change*, 111 Nw. U. L. Rev. 1565, 1565 (2017) (summarizing decades of research on empirical tests of intuitions about retributive justice).

11   *See, e.g.*, William J. Bowers, Marla Sandys & Thomas W. Brewer, *Crossing Racial Boundaries: A Closer Look at the Roots of Racial Bias in Capital Sentencing when the Defendant Is Black and the Victim Is White*, 53 DePaul L. Rev. 1497, 1502 n.21 (2004) (collecting studies showing that African-Americans have lower trust in the criminal justice system); Sarah M. Buel, *Putting Forfeiture to Work*, 43 U.C. Davis L. Rev. 1295, 1334 (2010) (observing a lack of trust among the Hispanic community toward the criminal justice system).

alienation that results from racism and perceptions of racism simply does not register within conventional metrics. Thus, those concerned about reforming criminal justice institutions for reasons of racial justice often find themselves, in a sense, on the outside of the system, trying to explain that something might be wrong even if, say, crime rates are going down. A metric based on social trust would change this dynamic: to the extent a criminal system diminishes trust among members of a racial group, that fact becomes *visible* and, assuming the commitment to trust reflected in the metric is genuine, *important.* A police chief or district attorney would get a report showing quantitatively that, say, trust in the criminal justice system among black Americans is low. That police chief or district attorney would have reason—indeed, incentive-based reasons—to try to figure out why trust is low and how to address the problem.

Sixth, social trust should speak to criminal justice liberals in the John Stuart Mill tradition of protecting liberty against arbitrary power. Libertarianism's starting point, for many, is *distrust* of the state—and distrust of the community, conceptualized as a mass or crowd, as well. That starting-point has led many in the libertarian tradition to imagine that the end-point—a society in which individual freedom is actually treasured and enjoyed—is characterized by a similar kind of mistrust. That is the error. A life of genuine freedom cannot be lived in a society in which one rationally believes that agents of the state or other citizens are likely to invade or neglect one's rights—conditions, that is, of low trust. It is not as though people are free in a violent society and a failed or corrupt state. Real freedom requires just the opposite: a rational expectation that neither state officials nor other citizens will do us wrong—conditions, that is, of high trust. There is, in other words, a deep and underappreciated link between libertarian freedom and social trust. Libertarianism *starts* from a keen apprehension of how state and community power can oppress, but it *ends* by devising ways to minimize and overcome the risk of that oppression (albeit backstopped by a keen apprehension of how state and community power can go wrong). Its endgame is to produce a society in which high levels of trust are rationally justified. The criminal context demonstrates this point with particular force, for there may be no other site in public life in which the twin risks of oppression by the state and invasion by other citizens are more acute. If one can assume that responses to crime that increase trust will tend to be those that (a) make citizens feel safe from private criminals, and (b) make citizens feel that their state is not arbitrary or oppressive, and vice versa, then a metric based on social trust will incentivize conditions of individual liberty. Add one more factor— that any policy proposal must be judged relative to its alternatives, and the present array of criminal justice metrics is arguably bad at

controlling crime and clearly bad at controlling arbitrary and oppressive state power—and the argument is done. Social trust should be a libertarian's preferred criminal justice metric.

Seventh, as the last several paragraphs demonstrate, social trust is a locus of agreement among the cacophony of different goals people bring to the criminal justice system, including individual freedom, racial justice, retributive justice, community solidarity, and crime control. This trading-post quality is central to our larger argument: a good metric in criminal justice, we submit, must be one that functions for policy *despite underlying theoretical disagreement.* In that sense, metric-construction in criminal justice is different from metric-construction in other contexts, where it is possible first to settle goals and then determine how to measure them. A firm's shareholders might agree that their goal is to maximize profits and construct a metric to track costs and revenue; a hospital's doctors might agree that their goal is to optimize health outcomes and construct a metric accordingly. But as we show below, the situation in criminal justice is less like that of tracking profit at a firm than, say, trying to measure the success of U.S. foreign policy: the underlying disputes about what constitutes success are so extreme as to stymie a decision procedure that requires settling goals first. In such a situation, one can either tie oneself to the mast of a particular theoretical perspective and construct a metric that speaks only to loyalists of that view, or construct a metric that tracks something important to virtually everyone—building a zone of "overlapping consensus," to use Rawls's term,[12] or "incompletely theorized agreement," to use Sunstein's.[13]

The latter is our approach here: as we will show, virtually everyone in criminal justice has reason to care about social trust. Indeed, social trust not only bridges gaps between different theoretical perspectives, like retributivism and communitarianism: it also transcends the unproductive oscillation between tough-on-crime and soft-on-crime, Left/Right thinking that currently structures the politics of criminal law. Social trust is neither tough-on-crime nor soft-on-crime. It speaks equally to crime control and community solidarity, to retributive justice and racial justice. The concept can be a site of overlapping consensus, performing the vital function on which both Rawls and Sunstein were focused: enabling liberal societies to make policy despite disagreement about first principles.

Eighth, a good metric in criminal justice must be *practicable*: without slighting the moral complexities of crime and punishment, a

---

12   John Rawls, *The Idea of an Overlapping Consensus*, 7 OXFORD J. LEGAL STUD. 1, 1 (1987).

13   Cass R. Sunstein, *Incompletely Theorized Agreements in Constitutional Law*, 74 SOC. RSCH. 1, 1 (2007).

metric that works for governance must be reasonably straightforward and inexpensive to assess, must be quantifiable, must make sense to criminal justice officials, must apply to different parts and stages of the criminal process, and must give officials information that translates to action. Philosophically sophisticated insights and empirically sophisticated models that cannot be translated into usable metrics are, for present purposes, beside the point. The goal is something that flawed, resource-limited, actual governments can use to figure out whether criminal justice officials are doing their job well and provide actionable information about how to do it better, whether the party in view is an individual officer in a traffic stop or a police department, prosecutor's office, court system, or prison system as a whole. Social trust has that flexibility and capaciousness, making virtually every aspect of the criminal system subject to a common standard of evaluation based on a relatively simple set of survey questions. This simplicity is, again, a feature not a bug. Some think that what leads to uptake of new ideas in government policy is that the intervention be small. That is a misconception. Better that a policy proposal be simple than that it be small. The sweeping simplicity of the social trust metric—a few questions that can be asked at any part or stage of the criminal justice process and that will, if we are right in our various contentions in this Article, thereby provide information useful to action—is key to getting policy change.

Finally, social trust has been so extensively studied in domains outside of criminal justice that scholars know a great deal already about *how* to study it. Survey instruments, experimental techniques, and other methods are already available and have already been validated, subjected to scholarly analysis, and refined over decades of use.[14] In the criminal justice context, studies have measured trust-related factors

---

14   *See, e.g.*, J. David Lewis & Andrew Weigert, *Trust as a Social Reality*, 63 SOC. FORCES 967 (1985) (reviewing sociological methods of studying trust); Jonathan Jackson, Ben Bradford, Mike Hough, Jouni Kuha, Sally Stares, Sally Widdop, Rory Fitzgerald, Maria Yordanova & Todor Galev, *Developing European Indicators of Trust in Justice*, 8 EUR. J. CRIMINOLOGY 267 (2011) (describing the methodological development process of a forty-five-item module in Round 5 of the European Social Survey, which fields the core survey indicators). For examples of validated social trust questions, see Ilir Gedeshi, Paul M. Zulehner & David Rotman, *EVS Trend File 1981–2017*, GESIS (July 7, 2021), https://search.gesis.org/research_data /ZA7503 [https://perma.cc/QAB2-S85W]; Rory Fitzgerald, *ESS10-2020*, EUR. SOC. SURV. (Jun. 6, 2022), https://ess-search.nsd.no/en/study/172ac431-2a06-41df-9dab-c1fd8f3877e7 [https://perma.cc/9WCM-LVQZ]; *EQLS 2016—Questionnaire*, EUROFOUND (2016), https://www.eurofound.europa.eu/surveys/european-quality-of-life-surveys/european-quality-of-life-survey-2016/questionnaire/ [https://perma.cc/72MP-NWPN]; *World Values Survey Wave 7 (2017–2022)*, WORLD VALUES SURV. (2022), https:// www.worldvaluessurvey.org/WVSDocumentationWV7.jsp [https:// perma.cc/R6X5-4NFC]; *GSS Documentation*, NORC, https://gss.norc.org/Get-Documentation [https:// perma.cc/3ZQZ-DADA]; GALLUP, WORLDWIDE RESEARCH: METHODOLOGY AND CODE-BOOK, (2017).

in the context of both policing[15] and criminal courts.[16] The existing literature should shift, we think, to the modified conception of trust we propose herein. But there is no need to start from scratch.

---

15   For a sample of studies on the relationship between police encounters and general attitudes toward the police (not necessarily trust), see Wesley G. Skogan, *Citizen Satisfaction with Police Encounters*, 8 POLICE Q. 298, 298–99, 316–17 (2005) (examining the character and consequences of encounters between police and residents of the city of Chicago and finding that citizens' satisfaction with police was determined by their interaction with police during encounters); Wesley G. Skogan, *Asymmetry in the Impact of Encounters with Police*, 16 POLICING & SOC'Y 99, 112–13 (2006) (finding, based on survey data on police-initiated and citizen-initiated contacts with police in Chicago, that the relationship between how people recall being treated and their general confidence in the police was asymmetrical so that negative personal experiences with police were far more influential on attitudes than positive ones); Lyn Hinds, *Public Satisfaction with Police: The Influence of General Attitudes and Police-Citizen Encounters*, 11 INT'L J. POLICE SCI. & MGMT. 54, 62–64 (2009) (finding that in Australia satisfaction from citizen-initiated contact with police made people more satisfied overall, but prior views of police performance, police legitimacy, and police use of procedural justice were stronger predictors); Dennis P. Rosenbaum, Amie M. Schuck, Sandra K. Costello, Darnell F. Hawkins & Marianne K. Ring, *Attitudes Toward the Police: The Effects of Direct and Vicarious Experience*, 8 POLICE Q. 343, 354–55 (2005) (finding, based on the measurement of attitudes before and after encounters with the police among residents of Chicago, that vicarious experience of the police (e.g., learning that someone else has had a good or bad encounter with the police) has a stronger influence than individual direct contact with the police over changes in people's attitudes toward police); Kathryn Foster, Melissa S. Jones & Hayley Pierce, *Race and Ethnicity Differences in Police Contact and Perceptions of and Attitudes Toward the Police Among Youth*, 49 CRIM. JUST. & BEHAV. 660, 675 (2022) (exploring how effects of direct and vicarious police stops on youth attitudes toward the police vary by race/ethnicity and finding that direct and/or vicarious police contact can generate negative attitudes toward police among black, Hispanic, and in some cases, white youth, though these effects vary across type of police stop and type of attitude. When a direct stop involved more officer intrusiveness, black youth reported less respect and more negative perceptions of procedural justice.); Kristina Murphy, Lorraine Mazerolle & Sarah Bennett, *Promoting Trust in Police: Findings from a Randomised Experimental Field Trial of Procedural Justice Policing*, 24 POLICING & SOC'Y 405, 405–06 (2014) (testing whether procedural justice could be used by police agencies during short, routine traffic stops to increase public trust and confidence in police and finding, using survey data from 2,762 Australian drivers who had been exposed to either a procedural justice script (experimental condition) or a standard police procedure (control condition), that trust and confidence in police was higher in the experimental condition, even after respondents' demographic background and general perceptions of police were taken into account).

16   For studies on how procedural justice perceptions and experience with criminal courts affect the attitudes toward these courts and the views of their legitimacy see, for example, Jane B. Sprott & Carolyn Greene, *Trust and Confidence in the Courts: Does the Quality of Treatment Young Offenders Receive Affect Their Views of the Courts?*, 56 CRIME & DELINQ. 269, 275–76 (2010) (interviewing a sample of youths at their first appearance in court and then again at the sentencing using the same questionnaire to explore changes in their views over time, and finding that the way youths feel they have been treated—specifically, by their own lawyer and by the judge—affected broad views of legitimacy, even when controlling for their overall satisfaction of the outcome of their case); Galma Akdeniz & Seda Kalem, *How Going to Court Affects the Attitudes Towards Courts*, J. SOCIO. RSCH., Oct. 2020, at 1, 9, 23 (arguing, based on a secondary analysis of data collected in 2006–2007 using a nationally

Part I, below, focuses on the problem of metric construction facing criminal justice today. It critiques existing metrics (Section I.A) and surveys the variety of different goals to which a metric must answer if it is to be politically practicable and a site of overlapping consensus (Section I.B). Part II focuses on our proposed, trust-based solution to the problem. It takes up the existing literature showing the central place of trust in social science today (Section II.A); the reasons to think social trust is of particular importance in the context of criminal law (Section II.B); the relationship between social trust and communitarian theories (subsection II.C.1); the effect of social trust on crime control (subsection II.C.2); the relationship of social trust to considerations of retributive justice (subsection II.C.3); the relationship of social trust to considerations of racial justice (subsection II.C.4); the relationship of social trust to considerations of political liberalism (subsection II.C.5); the effect of social trust on the incentive structures facing criminal justice officials (Section II.D); and a series of objections and responses to the social trust metric (Section II.E).

Finally, the Conclusion begins the process of thinking through what an instrument designed to measure social trust would look like. This Article's focus is on why social trust is the right thing to measure; figuring out how best to measure it—the question of implementation—will be the subject of follow-up work. But it is difficult to think through the first without at least considering the second, and the Conclusion provides a preview of our plans in that regard. The basic thought is to use survey techniques to ask people who have had recent encounters with the criminal justice system whether the experience left them with more or less trust in the criminal system and its officials, in government generally, and in their fellow citizen. If you got a ticket, or reported a crime, or in any other way encountered the criminal system, you'd get a follow-up survey assessing how your level of trust was affected by the encounter. One could also survey randomly selected members of the public, asking them about recent contacts with the criminal justice system or informing them about decisions taken within the criminal justice system and asking them about their reactions. But

---

representative sample of adults in Turkey, that attitudes toward courts are shaped by two components: the stable base which reflects a general and abstract trust in institutions, and a flexible component which is shaped by the nature of the experience and interaction that one has with the courts); Hilke A.M. Grootelaar & Kees van den Bos, *How Litigants in Dutch Courtrooms Come to Trust Judges: The Role of Perceived Procedural Justice, Outcome Favorability, and Other Sociolegal Moderators*, 52 LAW & SOC'Y REV. 234, 243–44 (2018) (finding, based on a sample of 483 respondents who filled out pre-hearing and post-hearing questionnaires in administrative and criminal court hearings in the Netherlands, that procedural justice is positively associated with trust in judges when outcomes are relatively favorable, and that this association is even stronger when outcomes are relatively unfavorable).

the point of presenting these tentative implementation ideas is only to fix ideas. Implementation is a hard problem—harder than it might at first appear—and it belongs in a separate article.

I.   THE PROBLEM

A.   *A Critique of Existing Metrics*

Six metrics dominate criminal law: crime rates, conviction rates, arrest rates, clearance rates, recidivism rates, and monetary cost. Those metrics are, without exception, important: *of course* we want to know how much crime there is, how many people are being arrested and for what, etc. But they are significantly misleading if treated as answers to the question of how well the criminal system (or some part of the system, like a police department or prosecutor's office) is working. They also distort incentives.

1. Crime Rates

The most common category of metric in criminal justice is crime rates—either the rate of overall crime or the rate of some specific type of crime or set of crimes. The FBI's "Uniform Crime Reporting" tool, for example, measures yearly fluctuations in homicide, rape, robbery, kidnapping, burglary, and auto theft.[17] Other metrics focus on homicide rates[18] or violent crime rates.[19] But although these metrics are vitally important for understanding matters of overall societal health and public policy, it is only a trick of the mind—a sort of intellectual optical illusion—that makes them seem like the right way to measure criminal justice systems.

The reason, mainly, is that criminal justice institutions do not control crime rates any more than (or not much more than) hospitals control disease rates. Criminal systems *respond* to crime, just as medical systems *respond* to illness. The response can have some effect on the amount of crime in society—how considerable is a matter of controversy—but what causes crime in the first place is a matter of the entire ecology of a society, from its economy to its culture to its welfare systems to its mental health systems and beyond. Doctors can do a lot

---

17   *Crime Data Explorer*, FED. BUREAU OF INVESTIGATION, https://crime-data-explorer.app.cloud.gov/pages/home/ [https://perma.cc/A5R9-DJ5M] (select "Crime Data Explorer" and scroll down to view rates of various violent and property crimes).

18   *See, e.g.*, *Homicide Mortality by State*, CTRS. FOR DISEASE CONTROL & PREVENTION, https://www.cdc.gov/nchs/pressroom/sosmap/homicide_mortality/homicide.htm [https:// perma.cc/MN4B-7ZCR].

19   *See, e.g.*, *Violent Crime in the U.S.*, STATISTA (Sept. 30, 2021) https://www.statista.com /topics/1750/violent-crime-in-the-us/#dossierKeyfigures [https://perma.cc/R2N6-SJ2D].

about cancer; they do not control whether people smoke. The intellectual optical illusion is to assume from the conceptual pairing of "crime and punishment" (or "crime and police" or "crime and criminal justice" or whatever else) that what happens in the one sphere depends on what happens in the other. The truth is that the two are only marginally coordinated.

How marginally? Do we have any idea how much criminal justice institutions affect crime? That is no less than the question of whether and how well deterrence, incapacitation, and other criminal justice interventions work. The basic answer is that the matter is empirically unsettled. Perhaps the best information available comes from studies of homicide—the category for which we have our best data, and which tends to correlate with violence generally. Those studies suggest that, while anarchic societies can have homicide rates above 100 per 100,000, "effective policing can drive homicide rates down to 10 or 20 per 100,000."[20] Impressive. But low-crime societies like those in democratic Europe today commonly achieve rates of 1 per 100,000, and even the United States, with its exceptionally high homicide rates, is only at 7 per 100,000.[21] The twenty-fold variation from 1 to 20 per 100,000—from the world's safest to nearly its most dangerous cities—is therefore *not* dependent on policing.

What does it depend on? What factors influence the homicide rate apart from effective policing? Perhaps the leading study argues, as we do later, that psychological and cultural factors related to social alienation are key, such as "[a] feeling of trust in government and the officials who run it," "fellow feeling arising from racial, religious, or political solidarity," and "[t]he belief . . . that one can command the respect of others without resorting to violence."[22] But the views in the scholarly literature are legion. Researchers focus variously on poverty,[23] unstable family structures,[24] peer influences,[25] personal stress

---

20   RANDOLPH ROTH, AMERICAN HOMICIDE 9 (2009).

21   *Id.* at 7; *see also* Daniel H. Foote, *The Benevolent Paternalism of Japanese Criminal Justice*, 80 CALIF. L. REV. 317, 384–85 (1992) (observing that the Netherlands and Japan had similar criminal justice systems, but the Netherlands had rising crime rates while Japan had low crime rates).

22   ROTH, *supra* note 20, at 18.

23   Matthew Redelings, Loren Lieb & Frank Sorvillo, *Years off Your Life? The Effects of Homicide on Life Expectancy by Neighborhood and Race/Ethnicity in Los Angeles County*, 87 J. URB. HEALTH 670, 671 (2010).

24   Kelsey A. Jonas, *Fixing the FMLA's Flaws: A Fight for Care, Adult Children, and Tax Incentives*, 118 W. VA. L. REV. 1313, 1321 (2016).

25   *See* Barry C. Feld, *Adolescent Criminal Responsibility, Proportionality, and Sentencing Policy: Roper, Graham, Miller/Jackson, and the Youth Discount*, 31 LAW & INEQ. 263, 273 (2013).

and trauma,[26] the quality of a country's welfare and mental health systems,[27] and more besides. *None* of those factors, notice, have much to do with criminal justice institutions like police, prosecutors, and courts. Those institutional actors can have some effect: they can deter lawlessness, incapacitate or rehabilitate those prone to repeat offenses, and even, as we argue later, comport themselves in office in ways that tend marginally to increase or diminish social trust. But they cannot change the economy, fix broken families, or undo personal trauma.

The point is this: with so many factors in play and so much unknown, we cannot determine from analyzing crime rates whether a criminal system is functioning well or poorly. Crime rates say a lot about the health of a society. But if we are searching, as here, for a metric by which to measure the success of criminal justice institutions and policies, crime rates do not work because we do not know what fraction of rising or falling crime to credit to criminal justice institutions. It is entirely possible to have failing criminal justice institutions in a low-crime society; the institutional failures might not be noticed when other social forces keep crime down, but they would still be failures. It is entirely possible to have successful criminal justice institutions in a high-crime society; the institutions might be seen as failing because of the optical illusion that causes us to associate crime rates and criminal systems, but the situation might be no different than good doctors alongside high obesity rates in a society that eats too much. Perhaps sophisticated empirical techniques will one day be able to disentangle the marginal effects on crime rates of criminal justice institutions relative to other factors. But those sophisticated empirics are not the norm today—the measures of crime rates that criminal justice officials actually use do not disentangle the factors—and might be impossible where data is limited or natural experiments unavailable.

The knowledge/causation problem just discussed is not the only problem with using crime rates to evaluate criminal systems. A second problem is the assumption that all we want from a criminal system is to minimize crime. Our criminal justice goals are more complicated than that. If cutting the hands off of thieves could drive theft to zero, we would not do it for reasons of justice—yet a crime rate metric would "see" the policy as a success. If imprisoning all recidivists for life could halve the crime rate, the policy would be a success in terms of crime control and a failure in terms of proportionality or mercy—yet our metric would only capture one of those dimensions. Just as we would

---

26   William B. Harvey, *Homicide Among Young Black Adults: Life in the Subculture of Exasperation, in* HOMICIDE AMONG BLACK AMERICANS 153, 164 (Darnell F. Hawkins ed., 1986).

27   *See* Steven M. Dawson, Comment, *The Promise of Opportunity—and Very Little More: An Analysis of the New Welfare Law's Denial of Federal Public Benefits to Most Legal Immigrants*, 41 ST. LOUIS U. L.J. 1053, 1064 (1997).

not measure the quality of a story solely by how gripping the plot is (while leaving out, say, good writing and character development), we should not measure a multiple-goal criminal system solely by its effect on crime.

A third problem is that measuring crime rates, even if it tracks (one) important goal, gives no guidance as to how to achieve that goal. A good metric should tell us about levers, not just outcomes. A medical system should track vaccine distribution, not just COVID infections, because there is good reason to think vaccine distribution will reduce COVID infections (and, if it doesn't, that is important information too). If we only know that the orchestra was bad, we have no idea why it was bad or how to fix it; if we know that it was out of tune, we have some guidance as to the direction of repair. We provide evidence below for the claim that social trust tends to reduce crime—that one of the best things a police officer, prosecutor, or other criminal justice official can do to reduce crime is to act in ways that increase trust. If that claim is true, social trust is a *lever* that might tell us more about how well the criminal system is functioning than crime rates *even in terms of controlling crime.* It would also give criminal justice officials more useful, actionable data than crime rates. And if that claim is false, the only way to find out that it's false is to measure both social trust and crime rates and investigate their relationship.

Finally, a metric focused on crime control alone sets into motion a cascade of distorted incentives for police, prosecutors, and judges. It incentivizes harsh punishment, for one thing: if one's metric of professional success is simply to reduce crime, the intuitive answer to every question of punishment is "more." It might be possible to push back on that impulse—to point out ways in which harsh punishment might increase crime in the long-term, for example—but that argument only works if the empirics bear it out[28] and if officials are willing to focus on the long term. Relatedly, crime control metrics incentivize incapacitative *forms* of punishment, like imprisonment, rather than liberty-preserving sanctions, like fines. Fines don't remove the possibility of a second offense and an uptick in the crime rate; imprisonment does (at least if one excludes in-prison crimes from the metric).

Another incentive problem is that crime control metrics give police and prosecutors reason to push the envelope of procedural constraints. If invasive or discriminatory stop-and-frisk policies and coercive forms of plea bargaining help get criminals off the streets, those benefits "count" in the metric of professional success, while the

---

28   There is some empirical support for the idea that harsh punishment can increase crime, but those empirics are not so definite that one would want to make all arguments against harshness depend on them. *See, e.g.*, Tim Friehe & Thomas J. Miceli, *On Punishment Severity and Crime Rates*, 19 AM. L. & ECON. REV. 464, 479 (2017).

constitutional costs do not. A pure crime control metric also incentivizes deliberately misclassifying offenses. Some of the most darkly funny scenes in *The Wire* involve "juking the stats" this way: "Making robberies into larcenies, making rapes disappear," Prez explains.[29] "You juke the stats and majors become colonels."[30] (The opposite is also possible: one could juke the stats to make crime seem worse than it is if helpful for, say, getting funding.)

Crime control metrics even distort the incentives facing legislatures, giving them reason to define crimes in ways that make crime control as easy as possible, rather than defining the thing they actually mean to stop. If drug possession with intent to distribute is a legislature's target concern, but intent is hard to prove, there is an easy solution: criminalize simple knowing possession. Or keep the intent-to-distribute element but make the sentences so long that prosecutors can secure plea bargains without needing to prove intent in court. Either way, convicting drug dealers becomes easier, and crime rates fall.

Crime rate metrics might even incentivize legislatures to *de*criminalize: if theft is no longer criminal, property crime goes down. Sound fanciful? It may have happened in recent years in California, where a statistical appearance of relatively low crime may have come from a combination of pure decriminalization (which makes offenses disappear from statistical measures altogether) and reclassifying what were formerly felonies as misdemeanors (which leads to reduced enforcement and reporting, and makes the felony rate go down even if the less visible misdemeanor rate goes up).[31] The fundamental problem is that, if the metric of success is just low crime rates, legislatures' incentives shift from writing statutes that define the social evil the legislature wants stopped to figuring out how, instrumentally, to make crime go down. The result is to distort the law itself.

2. Conviction, Arrest, and Clearance Rates

Perhaps the dominant metric of professional success among prosecutors is conviction rates—a measure of how often a prosecutor secures a conviction once he or she has chosen to indict. Rates are high—the average is 94.5%[32]—which makes prosecutors look effective.

---

29 *The Wire: Know Your Place* (HBO television broadcast Nov. 12, 2006).

30 *Id.*

31 *See* Joshua Kleinfeld & Thomas Hoyt, *A Criminal Law Based on Harm Alone: The Story of California Criminal Justice Reform*, 94 S. CAL. L. REV. 35, 37, 57–59 (2020).

32 U.S. DEP'T OF JUST., UNITED STATES ATTORNEYS' ANNUAL STATISTICAL REPORT: FISCAL YEAR 2020, at 5–7 (2020); *see also* John Gramlich, *Only 2% of Federal Criminal Defendants Go to Trial, and Most Who Do Are Found Guilty*, PEW RSCH. CTR. (June 11, 2019), https://www.pewresearch.org/fact-tank/2019/06/11/only-2-of-federal-criminal-defendants-go-to-

But the background fact is that American prosecutors have almost total charging discretion: they can decline to bring charges at any time, regardless of the underlying evidence and with no obligation to give reasons, and they can also indict for whatever charges the law and facts will reasonably support, no matter how the crime is intuitively characterized. If there is a video recording of one person shooting a gun at another, prosecutors have the authority to prosecute the crime as attempted murder, assault with a deadly weapon, simple assault, mere illegal discharge of a firearm, or simply not to bring charges at all. That background fact makes the conviction rates metric unhelpful for assessing much of anything—including prosecutorial effectiveness—and has bizarre effects on incentives.

The first problem with conviction rates is giving prosecutors an incentive to value airtight evidence (which is not the same as proof beyond a reasonable doubt) more than considerations of public safety or justice. This can lead to socially destructive patterns of both under-charging and overcharging. Imagine a defendant who has a good case for leniency based on justice or mercy but against whom the evidence of a crime is rock-solid (a drug mule, for example, who might be as much victim as offender). A prosecutor responsive to her incentives has good reason to pursue charges in that case. Now imagine a suspect who is a serious predator and against whom the evidence is solid beyond a reasonable doubt but not airtight. A prosecutor responsive to his incentives has reason *not* to indict in that case, because a conviction rate metric incentivizes him above all not to lose. This might explain some of the peculiar patterns of nonprosecution in American criminal law—why gangsters known to whole neighborhoods (or to the whole country, like Al Capone) often escape indictment or why, during the mortgage crisis a decade ago, virtually the entire financial sector avoided indictment. Another variation on the theme: rather than not charging at all, prosecutors responsive to their incentives have good reason to charge what is easiest to prove rather than what the defendant actually did. Imagine again a video showing one person shooting at another one, clearly trying to kill. Attempted murder requires proving intent, which might be difficult notwithstanding the video. Charge illegal discharge of a firearm and the problem dissipates. But what if the offender really was attempting murder? There is value in aligning crimes of conviction with what offenders actually did—rule of law norms of candor and public intelligibility, for example. The conviction rate metric incentives just getting the win.

---

trial-and-most-who-do-are-found-guilty/ [https://perma.cc/378N-WMV6] (explaining that, in 2018, fewer than 1% of criminal defendants went to trial and won).

These incentive structures become still more pernicious in light of plea bargaining, for an admission of guilt is the one certain pathway to a conviction. In a world of both plea bargaining and conviction rate metrics, a prosecutor responsive to her incentives has reason to secure plea deals, not just for familiar reasons of efficiency, but precisely to avoid the public goods that trials serve—like juries' supervision of the facts, judges' supervision of the law, and defendants' right to make their best case. The conviction rates/plea bargaining combination also incentivizes unsavory tactics to get a plea deal, such as bringing excessively punitive charges to compel a confession or (a variation on the theme) bringing excessively punitive charges initially and offering excessively lenient ones subsequently: threaten someone who committed felony assault, but did not aim to kill, with an attempted murder charge and he might well confess to felony assault—but only because he was threatened with an untrue charge. Threaten him with attempted murder and offer him misdemeanor assault and he will almost certainly confess—but what he did, really, was felony assault.

The unifying theme here is that a metric of conviction rates gives prosecutors an incentive to overemphasize the *tactics* of conviction and underemphasize the *goals* of conviction—desert, proportionality, mercy, public safety, rehabilitation, and the other moral and practical goals criminal justice is supposed to serve. A criminal system should aim to align the incentives of prosecutors and other criminal justice officials with the goals of punishment. Conviction rates do the opposite. Now, it's easy to imagine objections. "Focusing on what the evidence will prove is a good thing—an expression of fidelity to the reasonable doubt standard," one might argue. "Prosecutors are generally people of good will and good sense, who will not engage in these strategic charging practices," another might argue. Our response to both objections is this: bad incentives creep into the thinking of even good people. An incentive that rewards certainty creates a desire to think that the reasonable doubt standard is not met when the evidence is anything short of perfect.

If these incentives do affect how prosecutors behave, assuming high-crime background conditions, what we should expect to see in prosecutorial practices is low- and midlevel prosecutorial churn: lots of indictments, lots of convictions, lots of sentences, but mostly not for major charges. The leading empirical study of mass incarceration to date—John Pfaff's important book, *Locked In*—finds that this is precisely what we do see and, furthermore, that it is a major driver of mass incarceration: "When I first saw my own results," Pfaff writes, "I stared at my computer for a few minutes in disbelief. I had expected to find that changes at every level—arrests, prosecutions, admissions, even time served—had pushed up prison populations. Yet across a wide

number and variety of states, the pattern was the same: the only thing that really grew over time was the rate at which prosecutors filed felony charges against arrestees."[33] The rate grew—but lengthy sentences did not. During the explosion of incarceration in the 1970s, '80s, and '90s, what seems to have happened is that prosecutors filed a lot of felony charges, pled a lot of cases to low- and mid-level felonies, and thereby incarcerated lots of people for relatively short lengths of time.[34] That does not prove that prosecutors were motivated by the conviction rate metric. But it is certainly consistent with such a motivation.

One sees the same root problem with arrest rates and clearance rates—two of the major metrics used to evaluate police forces (both individual officers and departments). When police forces are rewarded for making lots of arrests, their incentive is to make an arrest whenever the evidence of a crime is solid, no matter how trivial the crime. If it's easier to arrest five people for minor drug charges than one major dealer, it makes sense to invest resources in the five, which has the perverse effect of simultaneously increasing the number of people caught up in the criminal net and ignoring the people who most need to be stopped.

Clearance rates—a measure of the proportion of reported crimes that lead to arrest—are better: it makes good sense to test whether the police are *solving* crimes. But even that metric incentivizes police to focus resources on the crimes that are easiest to solve rather than the ones that are most serious. It also, once again, gives police incentives to misclassify crimes, such that difficult-to-solve crimes disappear or at least appear less serious than they were. If a felony burglary is hard to solve, call it misdemeanor larceny. If a rape is hard to solve, pretend it was consensual: "Making robberies into larcenies, making rapes disappear," as Prez explained.[35]

3. Recidivism Rates

Measuring recidivism has become the "go-to" performance measurement tool in many criminal justice contexts—sometimes to measure the overall performance of the system but more commonly to measure specific programs and methods.[36] It is absolutely central to sentencing policy: the logic behind the Federal Sentencing

---

33   JOHN F. PFAFF, LOCKED IN: THE TRUE CAUSES OF MASS INCARCERATION—AND HOW TO ACHIEVE REAL REFORM 72 (2017).

34   *Id.* at 74–75.

35   *The Wire, supra* note 29.

36   Cecelia Klingele, *Measuring Change: From Rates of Recidivism to Markers of Desistance*, 109 J. CRIM. L. & CRIMINOLOGY 769, 777–78 (2019); Chris Cunneen & Garth Luke, *Recidivism and the Effectiveness of Criminal Justice Interventions: Juvenile Offenders and Post Release Support*, 19 CURRENT ISSUES CRIM. JUST. 197, 197–98 (2007).

Guidelines' striking focus on criminal history, for example, is based on the U.S. Sentencing Commission's studies concluding that criminal history predicts recidivism.[37]   It is also central to decisions about whether to put a defendant into the criminal process or into a diversion program, whether to give a defendant probation or parole, and whether diversion, probation, or parole programs are working or failing.[38]   Funds flow on the basis of recidivism: the Federal Second Chance Act, for example, provides financial grants for reintegration plans of former prisoners—provided recidivism rates are not too high.[39]   Criminal justice insiders, like state supreme court justices, care about recidivism.[40]   Criminal justice critics and reformers have mostly surrendered to it, focusing their fire on sentencing policies not shown to reduce recidivism but often accommodating policies that do reduce it.[41]   Even community courts and other experiments in nonadversarial criminal justice commonly measure success by whether offenders recidivate.[42]   As algorithmic assessment tools take hold, recidivism is becoming still more important.   The Correctional Offender Management Profiling for Alternative Sanctions tool, for example, which was used in over a million cases between 1998 and 2018, was designed around one and only one goal: predicting recidivism.[43]   A newer algorithmic movement—the "evidence-based practices" movement—likewise focuses on recidivism.[44]

Recidivism *is* important.   It is utterly crucial to public safety: there is reason to think a small group of repeat offenders cause a large

---

37   U.S. SENT'G GUIDELINES MANUAL ch. 5, pt. A, introductory cmt. (U.S. SENT'G COMM'N 2018); *id.* at ch. 4, pt. A, introductory cmt.

38   Klingele, *supra* note 36, at 778–79.

39   *Id.* at 779–80; Second Chance Act of 2007: Community Safety Through Recidivism Prevention, 34 U.S.C. § 10631 (2018).

40   Roger K. Warren, *Evidence-Based Practices and State Sentencing Policy: Ten Policy Initiatives to Reduce Recidivism*, 82 IND. L.J. 1307, 1307 (2007); *see also* TRACY W. PETERS & ROGER K. WARREN, NAT'L CTR. FOR STATE CTS., GETTING SMARTER ABOUT SENTENCING: NCSC'S SENTENCING REFORM SURVEY 5 (2006) ("Judges hearing felony cases frequently complain about the ineffectiveness of current sentencing policies and the resulting high rates of recidivism.").

41   Warren, *supra* note 40, at 1309–15.

42   WEST HUDDLESTON & DOUGLAS B. MARLOWE, NAT'L DRUG CT. INST., PAINTING THE CURRENT PICTURE: A NATIONAL REPORT ON DRUG COURTS AND OTHER PROBLEM-SOLVING COURT PROGRAMS IN THE UNITED STATES 9 (2011) (collecting studies of recidivism rates of drug courts participants); CYNTHIA G. LEE, FRED L. CHEESMAN, II, DAVID B. ROTTMAN, RACHEL SWANER, SUVI LAMBSON, MIKE REMPEL & RIC CURTIS, NAT'L CTR. FOR STATE CTS., A COMMUNITY COURT GROWS IN BROOKLYN: A COMPREHENSIVE EVALUATION OF THE RED HOOK COMMUNITY JUSTICE CENTER 156 (2013) (reporting that the community court studied reduced the recidivism rate among adult criminal defendants within two years by 10%).

43   Julia Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, SCI. ADVANCES, Jan. 2018, at 1, 1.

44   Warren, *supra* note 40, at 1308.

portion of the country's criminal harm.[45] Repeat offenses also change the framing of a crime, urging us to think about actor as well as act—about the person standing behind the crime. People commit a first offense for all sorts of reasons (momentary temptation, a flash of passion, a failure of judgment, a loss of control), but repeat offenses typically reveal either a flaw of moral character or a deep-rooted problem (addiction, trauma, mental health issues), and we cannot respond appropriately either to the moral failing or the deep-rooted problem without appreciating its persistence.

Yet important as it is, a recidivism-based metric has problems too. One is that it does not provide useful information when a lot of people are committing a first offense—social situations in which *many* people engage in crime. What is distinctive about certain kinds of societal problem is that crime has become widespread, rather than certain individuals getting away with a lot of crime. A criminal system needs metrics that capture both.

But the bigger problem is, again, incentivizing harshness: the obvious way to prevent repeat offenses is to lock recidivists up and, as the saying goes, throw away the key. Policies aimed at preventing recidivism, like three-strikes laws, commonly eventuate in life or multidecade sentences. Likewise, criminal justice officials responsive to their incentives and measured by whether those under their supervision reoffend have little reason to give someone a second chance. Now, there may be counterarguments to the intuition that preventing recidivism means punishing harshly—points about the criminogenic effects of long-term incarceration, for example. But those arguments depend on a long-term perspective and a number of empirical assumptions that are difficult to fully prove. Thus, the basic effect of a metric based on recidivism is harsh sentencing.

4. Cost

Cost-benefit analysis in government policy rose to prominence in the 1980s and almost immediately started to influence criminal law. In the tough-on-crime 1980s and 1990s, politicians commonly emphasized the cost of various criminal policies to taxpayers and academics commonly engaged in research that measured criminal justice policy in terms of cost-effectiveness (e.g., studies comparing the efficacy of different drug control programs in terms of return on investment).[46]

---

45   In a study of youth homicide victims and offenders in Boston, for example, the average killer had 9.7 prior arraignments. David M. Kennedy, *Pulling Levers: Chronic Offenders, High-Crime Settings, and a Theory of Prevention*, 31 VAL. U. L. REV. 449, 452 (1997).

46   Mark A. Cohen, *Measuring the Costs and Benefits of Crime and Justice*, 4 CRIM. JUST. 263, 265–66, 299–300 (2000).

In criminal justice today, discussion of financial costs and cost-benefit analysis are staples of government policy, academic research, and public discourse, used to measure both whether particular programs are succeeding or failing and whether the criminal system as a whole is providing a good return given its cost.[47]

In many ways, this focus on financial cost and cost-benefit analysis has been good for criminal justice policy. It pushes back against waste. It brings into sharper focus the tradeoffs between different goods, including nonobvious tradeoffs between investments in criminal justice and investments in other social goods that may bear on criminal justice (like mental health services). It is one of the few metrics that can push back against harshness (for example, long prison sentences are expensive). And, of course, if two criminal justice policies were equally beneficial, all else equal, choosing the cheaper option is just rational.

The problem is, cost-benefit analysis is incoherent until one specifies what counts as a "benefit"—and that is precisely the issue in contention here. Is the benefit in question a reduction in the crime rate? Better conviction rates for prosecutors? Reduced recidivism? Marginal increases in social trust? Cost-benefit analysis requires specifying a goal. The problem with cost-benefit analysis in criminal justice is that, all too often, a goal is assumed without reflection. Programs are measured by whether, for example, they promote the most convictions at least cost, or the most arrests at least cost, or the fewest repeat offenses at least cost. But that sort of calculus leads to all the problems of inaccuracy and bad incentives discussed above. We would not object to investigating how to maximize social trust at least cost. But that is just to say that cost-benefit analysis only comes into play once one knows one's goals.

There is also an incentive problem. Often the means by which to minimize cost in criminal justice involves mistreating offenders or otherwise violating important values. For example, as discussed above, a metric based on minimizing crime or recidivism incentivizes long prison sentences. A metric based on minimizing crime or recidivism *at least cost* incentivizes long prison sentences in the worst prison conditions that one can constitutionally get away with. Likewise, when police are measured by arrest rates, their incentive is to arrest a lot of low-level offenders rather than focus on major offenders. When police are measured by their arrests per dollar of spending, that incentive goes into hyperdrive: sustained investigation of major offenders is expensive. In other words, cost-benefit analysis pushes back against undue

---

[47]   *See, e.g.*, NAT'L INST. OF JUST., U.S. DEP'T OF JUST., DRUG COURTS: THE SECOND DECADE 27–31 (2006) (reporting that the drug court studied cost taxpayers significantly less than traditional courts); John Roman, *Can Cost-Benefit Analysis Answer Criminal Justice Policy Questions, and If So, How?*, 20 J. CONTEMP. CRIM. JUST. 257, 258 (2004).

harshness and irrational waste in some cases (e.g., emphasizing the costs of long prison sentences), but it can encourage exceedingly harsh and even irrational policies in others. Everything depends on what one treats as a benefit.

A related problem is that cost-benefit analysis is just incompatible with deontological considerations of justice. We do not starve prisoners, even if the cheapest way to hold them would be not to feed them. We do not throw suspects in prison without any sort of process, even if process costs money. Our conceptions of basic justice serve as constraints—often so intuitively that we do not notice them at all. But cost-benefit analysis cannot see those constraints and incentivizes criminal justice officials to get as close to the deontological minimum as possible.

Cost and cost-benefit analyses are best seen, then, as a supplement to other metrics. A society must first decide what it values in a given context—a goal of some kind that constitutes that society's conception of "benefit." It decides what counts as a cost, too, though usually the financial conception of costs is obvious enough to go unremarked. It decides what principles are nonnegotiable. And only then does cost-benefit analysis provide useful guidance, ensuring that we pursue our goals and uphold our principles efficiently. So conceived, cost-benefit analysis is a valuable supplement to any sort of metric, social trust included. But it can't substitute for those others.

B.   A Cacophony of Goals

Criminal justice is beset by so many and such fundamental disagreements that standard approaches to constructing a metric will not work in the criminal justice context. The normal procedure in constructing a metric is to figure out a goal and then determine how to measure it. There would be theoretical interest to doing that in criminal justice—starting from the perspective of a retributivist, for example, and then figuring out how to measure impunity, proportionality, and the other things retributivists care about. But whatever metric were to emerge from that procedure would be too disputed to become policy and would not speak to people of other theoretical points of view or to the many people who are agnostic on matters of first principles. It would also neglect the real insights different theoretical perspectives bring to the table: a theory like retributivism isn't wrong to value desert; it is wrong to *disvalue* crime control—and so it goes with many other theories as well, for they are rarely wrong to care about the things they care about. Disagreement about goals affects most every field, perhaps (medicine, financial management, accounting, astronomy), but the level of disagreement in criminal law is so extreme in degree as to be different from many other fields in kind. Recognizing

that fact sets up what is in our view the real challenge: identifying something widely shared enough, and measurable enough, to form the basis for an overlapping consensus as to metric.

Consider, for example, a dispute of criminalization—for example, the dispute over the criminalization of drugs. Advocates of Mill's harm principle argue that drug use cannot justly be criminalized because it is definitionally nonharmful: "[T]he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others," Mill wrote, and his followers ever since have developed a libertarian conception of "harm" that excludes self-regarding action (like drug use) and consensual other-regarding action (like drug sale).[48] Straightforward utilitarians, by contrast, do not agree that drug use is definitionally nonharmful: they try to measure and compare the harms or costs of allowing drug use (such as addiction and the provision of services to the addicted) with the harms or costs of prohibiting it (such as illegal markets and gangs).[49]

Retributivists, meanwhile, typically argue that the focus should not be on harmdoing but *wrongdoing*; Michael Moore, for example, argues that criminal legislation "must exclusively aim at preventing or punishing moral wrongs" by "prohibiting all and only those behaviours that are in fact morally wrong."[50] The question of drug criminalization is thus a moral question about whether using drugs is wrong. A related view—"rights theory"—focuses on whether conduct infringes someone's rights: if drug use infringes no one's rights, it cannot be criminalized.[51] Meanwhile, reconstructivists and other communitarians focus not on material harm, nor moral wrong, but on the cultural effects of drugs: "whether addiction offends or threatens a moral culture based on freedom, whether mind alteration offends or threatens a moral culture based on reason, and whether effortless forms of overpowering hedonistic pleasure create a bad environment in which to

---

48   JOHN STUART MILL, ON LIBERTY 80 (David Bromwich & George Kateb eds., Yale Univ. Press 2003) (1859); *see also* H.L.A. HART, LAW, LIBERTY AND MORALITY 4–6 (1963) (defending and developing Mill's principle in a storied debate with Patrick Devlin over the criminalization of homosexual sex); Richard A. Posner, *On Liberty: A Revaluation, in* ON LIBERTY, *supra*, at 197, 197 (defining "harm" as "temporal" like "a punch in the nose," "tangible, secular, material—physical or financial, or, if emotional, focused and direct—rather than moral or spiritual," and, crucially, "other-regarding" and "without consent").

49   Paul Smith, *Drugs, Morality and the Law*, 19 J. APPLIED PHIL. 233, 239–43 (2002).

50   Michael S. Moore, *Liberty's Constraints on What Should Be Made Criminal, in* CRIMINALIZATION: THE POLITICAL MORALITY OF THE CRIMINAL LAW 182, 192, 199–200 (R.A. Duff, Lindsay Farmer, S.E. Marshall, Massimo Renzo & Victor Tadros eds., 2014).

51   *See* Alan Brudner, *Private Law and Kantian Right*, 61 U. TORONTO L.J. 279, 284–90 (2011).

raise children,"[52] together with, insofar as drugs affect how people behave in public settings, the loss of "shared norms of decent comportment in common spaces."[53]

Notice a few features of this example. First, the conflict involves not only multiple goals and competing priorities, nor even incommensurable goods, but disputes about whether the competing goals are legitimate at all. Millian libertarians, for example, think it is wrong—not just low priority but *wrong*—to limit individual freedom because of moral or cultural disapproval.[54] Second, these deep conflicts of theory and principle cannot be set aside as merely academic concerns: societies cannot but decide whether and which drugs to criminalize, and the philosophical positions above get picked up and repackaged in public discourse and in the law itself. Third, one could recapitulate the conflict over criminalizing drugs in dozens of other contexts. Should prostitution be criminalized? What about the public disorder associated with homelessness? When if ever should speech be criminalized? What are the virtues and vices of criminalizing socially normal misbehavior, like underage drinking? These disputes over criminalization illustrate fundamental disagreements in our society over criminal law's proper function. Indeed, the first lesson of studying criminalization is that one cannot even say that an agreed-upon goal in criminal justice is to *control crime.* Nothing could be more obvious than the consensus that criminal justice should control crime . . . right? But before one can agree to control crime, one must first *define* crime. Many of our greatest disagreements are about what conduct to put in that box.

Once one has defined crime, the next issue is to enforce it—the work of police, prosecutors, and courts. Views about the goals of the police are among the most controversial of our time, marked by disagreements about both ordinary matters of good policing and attacks on the basic legitimacy of the institution.[55] As to prosecutors, familiar controversies about election versus appointment, undercharging, overcharging, and plea bargaining implicitly reflect fundamental disputes over goals. Consider a prosecutor's decision not to indict despite good evidence of a crime because she thinks a particular defendant deserves mercy, or to offer one defendant a break in order to get testimony against another. If it is prosecutors' duty to faithfully enforce the formal law, neither their personal sense of justice nor the advantages of

---

52   Joshua Kleinfeld, *Reconstructivism: The Place of Criminal Law in Ethical Life*, 129 HARV. L. REV. 1485, 1532 (2016).

53   Kleinfeld & Hoyt, *supra* note 31, at 71.

54   *See* supra note 48 and accompanying text.

55   *See* Jessica M. Eaglin, Essay, *To "Defund" the Police*, 73 STAN. L. REV. ONLINE 120, 123–24 (2021).

bargaining can justify those decisions. If their role is to do justice and protect the public, a measure of moral and practical discretion is necessary to the office. And as to criminal courts, consider familiar conflicts over the degree to which courts should focus on factual truth versus protecting defendants' rights (excluding evidence or allowing it, broadly or narrowly interpreting due process rights, etc.), as well as less familiar conflicts over the degree to which criminal offenders should be directed into traditional, adversarial courts versus nonadversarial alternatives like community courts and restorative justice.[56] Implicitly the question is one of goals: whether the court system should prioritize public safety or rehabilitation, uncovering factual truth, or protecting defendants' rights.

Once the criminal system has determined an instance of guilt, it must decide how to punish or otherwise respond to the crime. We thus come to well-known disputes over whether punishment should aim at retribution, deterrence, incapacitation, rehabilitation, or less well-known approaches like normative reconstruction,[57] restorative justice,[58] reparations,[59] and beyond. It seems obvious that these theories of how to respond to crime imply different ideas as to goals. Does the criminal system exist chiefly or solely to prevent future instances of crime (reflecting goals of deterrence, incapacitation, and, in some ways, rehabilitation)? To give wrongdoers their just deserts (retribution)? To sustain the community's normative order (reconstruction)? To restore relationships damaged by the crime (restorative justice)? To help victims (reparations)? All of the above? Such conflicts about the goals of punishment have been the stuff of academic and policy debate for several centuries.

Thus what we see at every stage of the criminal justice system, from criminalization to punishment, is significant, policy-relevant dispute about basic goals. This level of disagreement makes it impossible to construct a metric that makes sense on first principles for all concerned. But it might still be possible to measure something that most everyone has reason to care about. Imagine a dispute among parents with different educational first principles about their children's school: perhaps one group of parents thinks education exists chiefly to

---

56   *See* Hadar Dancig-Rosenberg & Tali Gal, *Guest Editors' Introduction: Multi-Door Criminal Justice*, 22 NEW CRIM. L. REV. 347, 351 (2019) (coining the term "multi-door criminal justice" to describe the multiplicity of processes, mechanisms, values, and goals that coexist in modern criminal justice systems worldwide).

57   Kleinfeld, *supra* note 52.

58   *See, e.g.*, JOHN BRAITHWAITE, RESTORATIVE JUSTICE AND RESPONSIVE REGULATION (2002).

59   *See, e.g.*, Lucia Zedner, *Reparation and Retribution: Are They Reconcilable?*, 57 MOD. L. REV. 228, 234–38 (1994).

enrich souls, another focuses on getting good jobs. Both might find that teacher quality predicts the outcomes they care about. Likewise, it might be possible to find a metric in criminal justice that tracks something many or all of the most influential perspectives care about.

II.   SOCIAL TRUST

Part II explains what the concept of social trust is and why it is so important throughout social science, why it has a special link to criminal law, and why it speaks at once to communitarian goals of repairing frayed relationships; utilitarian goals of controlling crime; retributive concerns about impunity and disproportionate punishment; racial justice goals of defeating bigoted laws and practices and addressing problems of racial alienation; and liberal goals of protecting individual liberty. In other words, we will show that social trust meets the challenge laid out at the end of Part I: the challenge of taking seriously the insights of diverse schools of thought in crime and punishment and finding a point of overlapping consensus among them. We will also examine the effect of a social trust metric on criminal justice officials' incentives and respond to objections.

A.   A Central Concept in All of Social Science

Social trust is not a new, peripheral, or idiosyncratic concept. It is the subject of a large scholarly literature, both theoretical and empirical, spanning decades, which has shown it to be among the more significant predictors of overall societal health known to social science. In proposing that the concept of social trust should play a central role in criminal justice, we are building on an existing edifice.

Frank Fukuyama defines social trust as "the expectation that arises within a community of regular, honest, and cooperative behavior based on commonly shared norms, on the part of other members of that community."[60] Margaret Levi defines it as "a holding word for a variety of phenomena that enable individuals to take risks in dealing with others, solve collective action problems, or act in ways that seem contrary to standard definitions of self-interest."[61] As such, its motivational character is distinct from the more familiar, incentive- and coercion-based motivational structures typically studied in economics and law. Rather, trust is an affective and cultural asset, which motivates and enables people to rely on each other even when law and incentives are

---

60   FRANCIS FUKUYAMA, TRUST: THE SOCIAL VIRTUES AND THE CREATION OF PROSPERITY 26 (1995).

61   Margaret Levi, *A State of Trust, in* TRUST AND GOVERNANCE, *supra* note 4, at 77, 78.

not enough. It is the piece of the societal puzzle that flows into the gaps left open by formal legal rules and self-interest.

What is startling about social trust is not so much the concept as the degree of its effect—even in contexts that one might think to be rule- and incentive-determined. Start with economics: all else equal, high levels of social trust predict national wealth. "[O]ne of the most important lessons we can learn from an examination of economic life," Fukuyama concludes, "is that a nation's well-being, as well as its ability to compete, is conditioned by a single, pervasive cultural characteristic: the level of trust inherent in the society."[62] His comparative study of different national economies argues that high levels of trust characterize and unify the most successful economies of the modern world, including Japan, Germany, and, contrary to stereotype, the United States[63] (though distrust in the United States is on the rise).[64] Small-scale studies of individual firms have likewise shown that intrafirm trust predicts productivity.[65]

The mechanism both at large and small scales seems to be one of enabling people to rely on each other in transactions even where ordinary incentives are unavailable and formal legal protections insufficient or costly to invoke. In particular, trust enables people to "solve information problems" (buying and selling without fear of tricks in the exchange) and "provide credible assurances that the trustee will follow through on her obligation" (without relying on suits for breach of contract).[66] These factors drive down transaction costs and enable commerce between larger, more anonymous groups. Indeed, one of the interesting findings in the literature is that, as general social trust declines, economic organization changes *form*: from impersonal corporations to family firms and from arms-length transactions among strangers to trade within kinship networks.[67] Essentially, when general social trust is lacking, people fall back on family and tribe.

Trust is also central to civil society: as Robert Putnam famously observed, one of the most extreme and well-studied differences between high- and low-trust societies is the flourishing sphere of nonstate civic organizations (professional associations, charitable clubs, parent-teacher groups, etc.) that dot high-trust societies like freckles on a

---

62   FUKUYAMA, *supra* note 60, at 7.

63   *Id.* at 7–8.

64   *See, e.g.*, LEE RAINIE, SCOTT KEETER & ANDREW PERRIN, PEW RSCH. CTR., TRUST AND DISTRUST IN AMERICA 20 (2019).

65   *See, e.g.*, Oliver E. Williamson, *Calculativeness, Trust, and Economic Organization*, 36 J.L. & ECON. 453, 463–66 (1993).

66   Levi, *supra* note 61, at 79.

67   *See* FUKUYAMA, *supra* note 60, at 25.

cheek, and that are largely absent from low-trust societies.[68] It seems that people are inclined to cooperate with others in various causes based on personal enthusiasm or altruism so long as they can trust those others to behave reciprocally. Thus, social trust leads to the formation of "social capital"—a related concept about the resources of healthy societies to solve collective action problems outside the state.[69] Fukuyama conceptualizes social trust as an input into social capital: "Social capital is a capability that arises from the prevalence of trust in a society or in certain parts of it."[70] Thus trust undergirds social capital and civil society both.

Most significantly for this paper, trust is a vital factor in legal compliance: as intersecting streams of research have shown, what chiefly makes people obey the law in the many situations in which sanctions are unlikely is the belief that the state's demands are reasonable, legitimate, and fair, and that other citizens can be expected to obey them as well. One important stream of research supporting this point is Tom Tyler's procedural justice theory (pioneered, appropriately enough, in a book entitled *Why People Obey the Law*).[71] Based on a set of empirical studies, Tyler essentially shows that fair procedures contribute to a sense of governmental legitimacy, which in turn has a massive effect on legal compliance—far more massive than those who think of human motivation in terms of force and incentives would predict.[72] Indeed, fair procedures have more effect on compliance than substantive outcomes.[73]

---

68   *See* ROBERT D. PUTNAM, BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY 88 (2000); ROBERT D. PUTNAM MAKING DEMOCRACY WORK: CIVIC TRADITIONS IN MODERN ITALY 98–99 (1993).

69   *See* Glenn C. Loury, *A Dynamic Theory of Racial Income Differences, in* WOMEN, MINORITIES, AND EMPLOYMENT DISCRIMINATION 153, 176 (Phyllis A. Wallace & Annette M. LaMond eds., 1977); James S. Coleman, *Social Capital in the Creation of Human Capital*, 94 AM. J. SOCIO. S100–S101 (1988). *See generally supra* note 67 and accompanying text; Tom R. Tyler & Roderick M. Kramer, *Whither Trust?, in* Trust IN ORGANIZATIONS: FRONTIERS OF THEORY AND RESEARCH (Roderick M. Kramer & Tom R. Tyler eds., 1996); TRUST: MAKING AND BREAKING COOPERATIVE RELATIONS (Diego Gambetta ed.,1988); DOUGLASS C. NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE (1990); John Orbell & Robyn M. Dawes, *A "Cognitive Miser" Theory of Cooperators' Advantage*, 85 AM. POL. SCI. REV. 515 (1991); John M. Orbell & Robyn M. Dawes, *Social Welfare, Cooperators' Advantage, and the Option of Not Playing the Game*, 58 AM. SOCIO. REV. 787 (1993).

70   FUKUYAMA, *supra* note 60, at 26.

71   TYLER, *supra* note 5.

72   *Id.* at 175; *see also* Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283, 284, 300 (2003); TOM R. TYLER & YUEN J. HUO, TRUST IN THE LAW: ENCOURAGING PUBLIC COOPERATION WITH THE POLICE AND COURTS 156 (2002); Levi, *supra* note 61, at 11.

73   TYLER, *supra* note 5, at 175.

Other streams of research, less procedurally focused, show similar patterns. Studies of disaffected minorities in Israel, for example, find that negative feelings toward government tend to spur low-salience forms of law breaking, such as traffic violations.[74] Studies of compliance with COVID rules, where coercive enforcement was often difficult and societies depended on voluntary compliance, show that people have tended to resist pandemic restrictions to the extent they saw them and the people imposing them as unreasonable, dishonest, or hypocritical—to the extent, that is, that the government and the expert class had lost trust.[75] Indeed, there is reason to think that the fundamental story of the COVID-19 pandemic—the whole worldwide saga of expert advice and government directives, of lockdowns and vaccines, and of defiance and more-or-less compliance—is in substantial part a tale about social trust. As Fukuyama argued (twenty-five years after writing his book about trust), what has determined a society's success or failure in dealing with the pandemic "is not a matter of regime type. Some democracies have performed well, but others have not, and the same is true for autocracies. The factors responsible for successful pandemic responses have been state capacity, social trust, and leadership."[76]

Other lines of research about legal compliance focus not on citizens' views of government, but on their views *of one another.* For example, research on "contingent consent" shows that a person's willingness to comply with governmental demands depends *both* on seeing the government as trustworthy and being "satisfied that other citizens are also engaging in ethical reciprocity."[77] As Levi explains, "[c]ontingent consenters are strategic but ethical actors; they want to cooperate if others are also cooperating."[78] In fact, people prove willing to shoulder quite heavy burdens in quite varied situations (think of, for example, complying with a military draft) so long as they are convinced that others are doing the same: "Promoting trust, in the form of reason to believe that fellow citizens are contributing their fair share, is thus a

---

74   *See* Roni Factor, David Mahalel, Anat Rafaeli & David R. Williams, *A Social Resistance Perspective for Delinquent Behaviour Among Non-Dominant Minority Groups*, 53 BRITISH J. CRIMINOLOGY 784, 796 (2013).

75   *See* Anne Leonore de Bruijn, Yuval Feldman, Malouke Esra Kuiper, Megan Brownlee, Chris Reinders Folmer, Emmeke Kooistra, Elke Olthuis, Adam Fine & Benjamin van Rooij, Why Did Israelis Comply with COVID-19 Mitigation Measures During the Initial First Wave Lockdown? 18 (June 25, 2021) (unpublished manuscript), https://www.papers.ssrn.com/sol3/papers.cfm?abstract_id=3681964; Ori Aronson, Ido Baum & Dalit Gafni, *Batei-Mishpat Kehilati'im Be'Israel: Mechkar Mellaveh* [*COVID-19 Regulations and The Rule of Law: A Proposal for First-Wave Clemency*], 24 MISHPAT UMEMSHAL 121 (2022) (Isr.).

76   Francis Fukuyama, *The Pandemic and Political Order: It Takes a State*, FOREIGN AFFS., July/Aug. 2020, at 26, 26.

77   Levi, *supra* note 61, at 88–89.

78   *Id.* at 89.

potential alternative to costly incentive schemes for solving societal collective action problems."[79] Indeed, this impulse—the impulse to look to one's peers and comply only if one can trust that others are also complying—is so strong that efforts to enforce legal obligations by means of sanctions have been shown to backfire when the sanctions "giv[e] citizens reason to doubt that other citizens are contributing voluntarily to societal collective goods."[80] For example, an IRS advertising campaign that emphasized the seriousness of the problem of tax cheating and the sanctions tax officials would impose on offenders turned out to *reduce* compliance because viewers who were not cheating on their taxes concluded from the ads that others were doing so and that they were playing the fool by complying.[81] The fix was to change the ads to make cheaters look like social misfits.[82]

A great deal of this trust-and-compliance research focuses on compliance specifically with criminal law, and we discuss it further in the subsection below on social trust and crime control (subsection II.C.2). For now, the important thing is to see that trust comes in two forms, which one might term "vertical" and "horizontal." The first, "vertical" dimension is one's attitude toward the state: Can one rely on it to behave intelligently, decently, and effectively? Are its procedures fair? Can one regard it as legitimate? The second, "horizontal" dimension is one's attitude toward fellow citizens: can one rely on them to follow the rules, interact fairly, and show goodwill? Our proposal that a criminal justice metric based on trust measure trust in the criminal system itself, trust in government generally, and trust in one another is meant to capture both of these dimensions (splitting the vertical prong into two).

One final aspect of the connection between trust and the state deserves special note: that citizens come to know the state, and to trust or distrust it, not in the abstract, but through encounters with *officials.* Officials win trust, Levi argues, when they make credible their commitment to enforcing the law, demonstrate competence in doing so, honor agreements, uphold standards, and create "bureaucratic arrangements that reward competence and relative honesty by

---

79   Herbert Gintis, Samuel Bowles, Robert Boyd & Ernst Fehr, *Moral Sentiments and Material Interests: Origins, Evidence, and Consequences*, *in* MORAL SENTIMENTS AND MATERIAL INTERESTS: THE FOUNDATIONS OF COOPERATION IN ECONOMIC LIFE 3, 32–33 (Herbert Gintis, Samuel Bowles, Robert Boyd & Ernst Fehr eds., 2005) (citing Dan M. Kahan, *The Logic of Reciprocity: Trust, Collective Action, and Law, in* MORAL SENTIMENTS AND MATERIAL INTERESTS: THE FOUNDATIONS OF COOPERATION IN ECONOMIC LIFE, *supra*, at 339).

80   *Id.*

81   *See* Kahan, *supra* note 78, at 347–51.

82   *Id.*

bureaucratic agents."[83] They lose it through "promise breaking, incompetence, and the antagonism of government actors toward those they are supposed to serve."[84] There is evidence that when the state builds trust in one context, that trust transfers to other contexts: "For example, citizens who trust the government or a major agent as a protector of legal rights may also trust the government as a fair conscripter for the military."[85] And in this dynamic of gaining, losing, and transferring trust, one particular group of officials has an outsized effect: police. Encounters with police are both relatively frequent and relatively vivid. As Tracey Meares argues, police "powerfully and pervasively provide both people who are processed by the system and those who are not with a formal education in what it means to be a citizen."[86] So powerful is this research that President Obama's Task Force on 21st Century Policing, which included police chiefs, sheriffs, civil rights activists, and union representatives among its members, made recommendations to build trust and legitimacy of policing "the foundational pillar on which all of our other recommendations rested."[87]

B.   *Trust and the Nature of Crime and Punishment*

The discussion in the last Section—social trust's relationship to economics, civil society, and legal compliance—is at once intuitive and surprising, like finding out that early childhood piano practice predicts later medical school performance. Once it is said, one can see why it would be so, but there doesn't seem to be an a priori reason why it should be so; the connection is contingent. Not so with the link between social trust and criminal law. The connection is stronger—not a matter of conceptual necessity, perhaps, but of what one might call *anthropological necessity.* Trust is relevant to criminal law because crime, properly understood, is a kind of wrongdoing that puts the basic terms of social life in doubt, and in response to which it is therefore *rational* to reduce trust. Crime is a challenge to trust, a *reason* to distrust. An effective response to crime—*punishment,* taking that word in its broadest signification—is therefore also a matter of trust: it should be a reason to trust again.[88]

One of the persistent misunderstandings of crime and punishment is to imagine that the harm caused by crime can be understood

---

83   Levi, *supra* note 61, at 85–89.

84   *Id.*

85   *Id.* at 84.

86   Meares, *supra* note 10, at 1527.

87   *Id.* at 1532.

88   For other criminal law scholars reflecting on this link, albeit usually in more specific theoretical contexts (such as the connection between retributivism and trust), see *supra* note 5.

in purely material (e.g., physical or financial) terms and that the necessary responses can be likewise purely material (e.g., measures solely of compensation or control). The misunderstanding represents a failure to grasp the nature of criminal wrongdoing and the problem it presents. Fundamentally, the problem of criminal wrongdoing is this: whatever society we build, however good and just, it is the human condition that some people some of the time will violate the rules and principles on which it is based in such a way that their actions performatively deny those rules and principles.[89] The norms in question are relative to the kind of society one has: blasphemy in a theocratic society denies the authority of the society's religion; interpersonal violence in a liberal society denies the principle that every individual has an inviolable sphere of rights. But what such forms of wrongdoing have in common is that they show hostility or indifference to the basic obligations of social life; they represent an *attack* on the society's normative underpinnings. Accidentally damaging someone's property might be tortious, but it is vandalism that attacks rights of property. Falling into someone might injure them, but it is assault that denies rights of physical integrity. The difference is not material harm: there might be no difference between assault, vandalism, and accident on that score. The difference is that the form of wrongdoing properly labeled "criminal" performatively denies the social contract that governs a society of that kind.

It is rational when presented with wrongdoing of this kind—wrongdoing that breaks the social contract—to ask whether the social contract still holds. The crime is a reason to doubt that it holds. It presents victims, third parties, and even offenders themselves with questions about whether the state can protect them (for if it could, why was the norm broken?) and whether other people can be trusted to respect the basic rules of social life (for at least one of those people showed himself ready to violate those rules). And in a larger sense, the crime puts the norm itself in a liminal space: it puts into question whether that norm is actually binding in social life, whether it is, in a sense, real or fake. These questions are particularly acute when the norm-breaker is itself a representative of the state, for then the three nodes of alienation—the conviction that the state cannot protect you, that one's fellow citizens will prey on you, and that the norm is fake—collapse in upon each other. A mere accident leading to civil liability is a physical or material harm requiring repair. But crime by nature is a challenge to faith.

The questions that crime presents—questions about the state, about other people, and about the social contract itself—are not

---

89   For a philosophical exposition, see Kleinfeld, *supra* note 6, at 1459–60.

answered until society *responds* to the crime. If crime by nature is a reason to withdraw trust, the response to crime must provide a reason to restore trust. The task of the state in punishing—again, using that word in its broadest conception—is therefore not limited to the tangible and material tasks of repairing and preventing further violations. The challenge is to craft a response to wrongdoing capable of restoring trust in the state, in one's fellow human beings, and in the society of which one is a part. It must be possible postpunishment for a rational person to believe that the social contract holds.

Trust is thus conceptually integral to the whole edifice of criminal law; it is the second dimension, beyond material harm and control, that makes crime and punishment different than other areas of law and life that also involve managing risk and cost. In the pages to follow, we will try to show that this trust-crime connection holds for a series of different, more particularized theoretical perspectives, and (in the discussion of crime control) that it holds empirically as well as conceptually. But all of that is, in a sense, secondary to this point about the nature of crime and punishment. Trust is the right thing to measure in criminal justice because losing and restoring trust are what crime and punishment are fundamentally about.

C. An Overlapping Consensus

1. Communitarian Goals

We turn now to a set of communitarian perspectives on crime and punishment for which increasing trust is explicit and at the core of the theory. For some theorists, like retributivists and crime control utilitarians, trust is not intrinsically of interest; it just happens to correlate with the things the theorist really cares about. But trust is on native soil in communitarianism. Indeed, what led the two of us to this paper was the realization that the philosophical communitarianism we have each in our own way advocated in criminal theory has a *correlate* in social science, and *social trust is that correlate.*

The first such theory, "reconstructivism," holds that the central object of criminal law is to reconstruct a community's normative order in the wake of acts that violate and threaten that normative order—restitching the torn social fabric, to use the clichéd but helpful metaphor.[90] Criminal law on a reconstructive view has a distinctive social function: where a wrong has been committed that is of such a nature as to attack the values on which social life is based, criminal law's role

---

90   *See* Kleinfeld, *supra* note 52; *see also* Kleinfeld, *supra* note 6, at 1457–58; Joshua Kleinfeld, *Two Cultures of Punishment*, 68 STAN. L. REV. 933 (2016).

is to reassert and reconstruct those values.[91] Reconstructivists thus view crime as a force of alienation—it is aptly called *antisocial*—and punishment, at its best, as a perfect opposite, something *prosocial.* Punishment acts as a sort of normative immune system against ideological invasions that, if left unanswered, would tend to weaken people's sense of being bound to one another, to a shared set of values, and to a shared system of law.[92] Reconstructivism is thus a form of communitarian consequentialism oriented to social solidarity.

Consider, for example, an offender who commits a serious assault without justification or excuse, beating the victim badly, heedless of the consequences. As a reconstructivist sees it, that assault carries two layers of social meaning. First, it expressively denies the validity of moral and legal norms. Second, it denies the victim's status as someone who matters—someone whose rights and welfare have value. Punishment's distinctive function is to counteract these meanings, to deny the denial, reaffirming the validity of the norm, the authority of the law, and the dignity of the victim: "Thefts break down norms of property, and punishment rebuilds them. Burglaries deny the security of the home, and punishment affirms it. Domestic violence degrades its victims, and punishment denies their degradation."[93] Crime and punishment are thus an exchange of meanings. Their call and response define criminal justice as a distinctive mode of social ordering.

Now, as the product of an intellectual tradition founded by Hegel and Durkheim, reconstructivism's main ideas are typically expressed in a philosophical way.[94] But they don't have to be. The concepts at work in reconstructivism ("alienation," "solidarity," and the like) are not identical to the concepts swirling around the social trust literature ("social capital," "procedural justice," etc.). But they are close. The jump from solidarity to social trust and from alienation to distrust is just the jump from philosophy to social science.

The restorative justice movement also regards trust-promotion as a central goal. Sometimes the focus is explicit: restorativists not uncommonly discuss the damage crime does to the victim's and larger community's sense of trust and thematize trust-building in response.[95] More often, however, restorative justice's trust-orientation is implicit in the very process of bringing together offender, victim, family members, and community stakeholders in a conversation that simultaneously insists on the crime's wrongfulness and the need for the offender

---

91   *See* Kleinfeld, *supra* note 52, at 1499.
92   *See* Kleinfeld, *supra* note 6, at 1462.
93   Kleinfeld, *supra* note 6, at 1462; *see also* Kleinfeld, *supra* note 52, at 1509.
94   *See* Kleinfeld, *supra* note 52, at 1487.
95   Albert W. Dzur & Susan M. Olson, *The Value of Community Participation in Restorative Justice*, 35 J. SOC. PHIL. 91, 91–92, 96 (2004).

to take responsibility for it, while also aiming to restore the offender to the fold.[96] "Reintegrative shaming" approaches, for example, bring communities together to shame offenders for their deeds while at the same time expressing commitment and concern toward them as persons.[97] This joint denunciation of the wrongdoing and acceptance of the wrongdoer combines high levels of social control with high levels of social support—unlike typical punitive approaches, which provide high levels of control with low levels of support, or typical rehabilitative approaches that provide high levels of support with low levels of control. The goal is expressly to repair relationships among offenders and victims and build social capital throughout the community.[98] Any such process is a trust-building exercise. In substantial measure, what restorative justice is restoring is trust.

A third member of the communitarian family—nonadversarial, nonpunitive procedural systems like community courts—likewise treat trust as a fundamental concept. The goal of such systems is to address the root problems that lead to crime, particularly repeat crime, and thereby to offer a rehabilitative alternative to jail and prison.[99] Prosecutors, defense attorneys, probation officers, and the judge work as a team, with the judge acting as team leader, to construct a rehabilitation plan for participants, requiring things like abstaining from substance use, maintaining employment, getting mental health services, reestablishing family contacts, and checking in with participants at regular hearings to see how they're progressing.[100] To a substantial degree, these rehabilitative programs strive to embed participants in webs of relationships that enhance their sense of connection to others. Successful participants often find their way back to estranged family members. One of the interviewees in a study on community courts in Israel, a prosecutor, mentioned a story of a female defendant who committed violent offenses against her ten-year-old daughter: "The daughter wrote a letter and asked that it be read in court . . . . 'Thank you for

---

96   *See* Joanna Shapland, Anne Atkinson, Helen Atkinson, Emily Colledge, James Dignan, Marie Howes, Jennifer Johnstone, Gwen Robinson & Angela Sorsby, *Situating Restorative Justice Within Criminal Justice*, 10 THEORETICAL CRIMINOLOGY 505, 506–07 (2006); David Miers, *Situating and Researching Restorative Justice in Great Britain*, 6 PUNISHMENT & SOC'Y 23, 24 (2004).

97   *See* JOHN BRAITHWAITE, CRIME, SHAME AND REINTEGRATION 55 (1989).

98   *See id.* at 56–57; *see also* Ted Wachtel & Paul McCold, *Restorative Justice in Everyday Life, in* RESTORATIVE JUSTICE AND CIVIL SOCIETY 114, 116–17 (Heather Strang & John Braithwaite eds., 2001).

99   *See* GREG BERMAN & JOHN FEINBLATT, GOOD COURTS: THE CASE FOR PROBLEM-SOLVING JUSTICE 1 (2005) (coining the term "revolving door justice").

100   *See* JOHN FEINBLATT & GREG BERMAN, RESPONDING TO THE COMMUNITY: PRINCIPLES FOR PLANNING AND CREATING A COMMUNITY COURT 3–4 (2001); Hadar Dancig-Rosenberg & Tali Gal, *Criminal Law Multitasking*, 18 LEWIS & CLARK L. REV. 893, 902 (2014).

returning my mother to me. My mother was irritated all the time and always angry, and after she came to you, and she came here, I got my mother back and I thank you.'"[101] Relationships with authority figures within these programs are also important, in part because those officials represent the state to the defendant. As one of the community court participants said about the judge and team in his case: "This is the first time that I feel someone is interested in what I'm going through and what bothers me."[102] A probation officer added: "Suddenly the judge asks how are you, what do you want to say? The defendant is in shock. He sees that the law-enforcement and justice system cares about him. That it's not the same system they've known before, that wanted to throw them in jail and failed to really see them."[103]

Trust runs through such processes like a red thread. It is only through trust that the judge, prosecutor, and other court officials can form the kinds of bonds with participants that make rehabilitation possible. It is the sense of being trusted that motivates participants to stick with the rehabilitative program when the process works. And it is trust that participants build when they fix relationships with family members and other people in their lives.

One of us has his roots in reconstructivism, the other her roots in nonadversarial, nonpunitive approaches to criminal law. There are differences between them. But social trust is at the center of both. It is indeed at the center of any communitarian approach to crime and punishment.

2. Crime Control Goals

One of the ironies of criminal justice is that, for the reasons discussed in Section I.A, measuring crime rates tells one little about whether the criminal system is functioning well—even to reduce crime. But social trust has a strong and apparently causal relationship to crime control. As we broached initially in Section II.A's discussion of trust and general legal compliance and extend here to the more specific question of trust and compliance with *criminal* law, diverse streams of research have come to a common conclusion: breakdowns in social trust—that is, alienation from the state and its laws, from society generally, and from other people within one's polity—are one of the most

---

101   Hadar Dancig-Rosenberg & Tali Gal, Many Shades of Success: Bottom-Up Indicators of Individual Success in Community Courts 18 (unpublished manuscript) (on file with authors).

102   Gal & Dancig-Rosenberg, *supra* note 9, at 400.

103   Dancig-Rosenberg & Gal, *supra* note 101 (manuscript at 21); *see also* TALI GAL & HADAR DANCIG-ROSENBERG, A FORMATIVE STUDY ON COMMUNITY COURTS IN ISRAEL: A FINAL REPORT 68–69 (2017).

significant and underappreciated causes of crime. Furthermore, as compared to many of the other factors affecting crime rates (broken families, poverty, poor mental health services, etc.), trust and alienation are things criminal justice officials can influence: how police, prosecutors, judges, and other criminal justice officials behave toward the people they interact with—how competently they do their jobs and whether they treat people fairly and respectfully in the course of doing their jobs—meaningfully affects how people see the criminal system and, through that lens, how they see their government. In short, criminal justice officials have some significant control over trust, which in turn has some significant influence over crime. Any pragmatically minded policymaker who cares about crime control and public safety should therefore regard social trust as a key metric of interest, as should any philosophical utilitarian or economic utilitarian likewise oriented to crime control goals. Our goal in this subsection is to assemble diverse streams of evidence—each, it bears emphasis, the product of massive collective effort by other scholars, and, inevitably, some dispute within their fields—to that effect.

To start with, trust seems to be *the* key variable in homicide rates. Homicide rates are among the most reliable indicators of overall levels of violent crime.[104] They are difficult to manipulate, reporting is high, legal definitions are stable enough that one need not fear a change in rates based on a change in what counts as "homicide" (notably, social scientists exclude vehicular manslaughter and deaths in war from the category), and, crucially, we have good data.[105] There is no other category of crime for which we can compare, say, a medieval English village with the American West after the Civil War, but, in virtue of large-scale efforts by historians and social scientists, we can engage in that sort of comparison for homicide. In addition, homicide rates change enough over time and place to give researchers meaningful variation. Rates around the world today vary from as little as 1 per 100,000 per year in Western Europe and Southeast Asia to over 20 per 100,000 per year in sub-Saharan Africa and Latin America, and commonly leap to 100s per 100,000 per year during civil wars and other upheavals.[106]

To appreciate these numbers, one must bear two things in mind. First, that they are yearly *rates*. If the homicide rate is 1 per 100,000, stays constant over time, and the average person lives to be 78, a

---

104   *See, e.g.*, Orlando J. Pérez, *Gang Violence and Insecurity in Contemporary Central America*, 32 BULL. LATIN AM. RSCH. 217 (2013).

105   See *id.*; Catrin Andersson & Lila Kazemian, *Reliability and Validity of Cross-National Homicide Data: A Comparison of UN and WHO Data*, 42 INT'L J. COMPAR. & APPLIED CRIM. JUST. 287, 288–91 (2018). *See generally* Manuel Eisner, Long-Term Historical Trends in Violent Crime, 30 CRIME & JUST. 83 (2003).

106   ROTH, *supra* note 20, at 7 fig.1.4, 18.

newborn has about a 1 in a 1,000 chance of being the victim of an as-
sault leading to death in the course of his or her life. At 5 per 100,000,
those chances are about 1 in 250. At 25 per 100,000, they are about 1
in 50. Second, one must bear in mind that homicide rates typically
track levels of violent crime generally, and other violent crimes are
much more common: a 1 in 250 chance of being killed by violence
suggests an unknown but much higher chance of being the victim of
some sort of violence. It must be acknowledged that the association
between homicide and violence generally is imperfect; in particular,
gun availability drives homicide rates up and medical improvements
drive homicide rates down.[107] Nonetheless, homicide rates are proba-
bly the best window we have into how violent a society is.

    What societal factors, then, drive homicide rates? In a magisterial
work of comparative history—not undisputed, to be sure, but in our
opinion the best guide currently available on the subject—Randolph
Roth argues that many common explanations cannot account for the
data. A strong state makes some difference: "By ending outright law-
lessness, effective policing can drive homicide rates down to 10 or 20
per 100,000."[108] But, as touched on above, low-crime societies like
those in democratic Europe today commonly achieve rates of 1 per
100,000.[109] To a striking extent, economic conditions, encompassing
both poverty and economic inequality, also cannot explain the degree
of variation: too many poor and unequal societies have low rates of
homicide (as in South and Southeast Asia), too many wealthy societies
have high rates of homicide (as in the United States), and there is too
much fluctuation uncorrelated with both poverty and economic equal-
ity (for example, homicide rates dropped during the Great Depres-
sion).[110]

    Yet one important clue lurks in the data. Most of the fluctuation
over time and place is due to one *type* of homicide: killing by men of
people with whom they have had no long-term hostility. Homicides
involving people connected to one another in long-term relationships
that have become toxic—a category that includes virtually all homi-
cides committed by women and some by men—do not vary dramati-
cally across time and place and are, in any case, "rarely numerous
enough to give any society a high homicide rate."[111] What drives major
fluctuations in levels of homicide is men's propensity to kill unrelated

---

107   David Hemenway & Matthew Miller, *Firearm Availability and Homicide Rates Across 26
High-Income Countries*, 49 J. TRAUMA INJ., INFECTION & CRITICAL CARE 985, 985 (2000);
ROTH, *supra* note 20, at 12.
108   ROTH, *supra* note 20, at 9.
109   *Id.* at 7 fig.1.4.
110   *Id.* at 7 fig.1.4, 9.
111   *Id.* at 14–16.

adults in situations of predation, sudden hostility, or recent hostility.[112] In other words, the question of why homicide rates vary over time and place becomes, once one looks at the data, a question of what causes men to be predisposed to violence. As Roth puts it, men in some societies are "willing to prey on others or to view them as enemies or rivals," "emotionally prepared to be violent at the slightest provocation," and ready to "view every encounter with another man as having the potential to be a life-and-death struggle for supremacy or self-preservation," while men in other societies "refrain from violence even if they are brutalized or humiliated."[113] Why?

Criminologist Gary LaFree points out that of every variable studied by social scientists, homicide rates among unrelated adults in the United States "have correlated perfectly with only two: the proportion of adults who say they trust their government to do the right thing and the proportion who believe that most public officials are honest."[114] Expanding on that finding to include four centuries of data, encompassing both Western Europe and the United States, Roth concludes that four factors predict homicide rates:

1. The belief that government is stable and that its legal and judicial institutions are unbiased and will redress wrongs and protect lives and property.

2. A feeling of trust in government and the officials who run it, and a belief in their legitimacy.

3. Patriotism, empathy, and fellow feeling arising from racial, religious, or political solidarity.

4. The belief that the social hierarchy is legitimate, that one's position in society is or can be satisfactory and that one can command the respect of others without resorting to violence.[115]

Some of the violence that ensues when these four factors are low is driven by people taking the law into their own hands because they don't trust the state to take action on their behalf (as in feuds). Some of it is driven by the feeling that violence is a means of winning respect. Some of it is driven by class-, race-, or ideology-based hostility, and some by individual aggression. But what all of it has in common is a deep relationship to social trust. Indeed, when one takes into account that social trust can be either vertical (trust in government) or horizontal (trust in one's fellow citizens), it is fair to conclude that both of

---

112   *Id.* at 17.

113   *Id.*

114   *Id.* (citing GARY LAFREE, LOSING LEGITIMACY: STREET CRIME AND THE DECLINE OF SOCIAL INSTITUTIONS IN AMERICA 75–81, 91–113 (1998)).

115   *Id.* at 18.

LaFree's factors and all four of Roth's factors are grounded in one thing: social trust.

Turning from homicide studies to procedural justice studies, again there is extensive evidence of a trust/crime control link. In Section II.A, we discussed empirical findings connecting fair and respectful procedures, perceived legitimacy, and general legal compliance. In later studies, Tyler and others showed that the same link holds specifically with regard to people's willingness to cooperate with the police.[116] When police are courteous and fair, the studies find, the effect is to spur both "immediate decision acceptance" and a sense that police authority is legitimate, which makes people "more willing to accept the directives and decisions of the police and courts" and diminishes "the likelihood of defiance, hostility, and resistance."[117]

A third stream of research connecting trust and crime rates focuses on "legal cynicism," that is, a generalized "skepticism about law and the actors who make and enforce it."[118] Whereas procedural justice is "a specific attitude; that is, it revolves around individual, face-to-face interactions between officers and citizens," legal cynicism is "a global attitude; that is, it is a generalized belief about the typical operations and motives of police and other authority figures."[119] It can be directed at state officials' motives or competence and may cause people to feel either that their government does not care about them or is incapable of acting effectively; the response, either way, is to adopt self-reliance strategies.[120] For example, legal cynicism predicts whether people in high-crime neighborhoods call the police when they need help or assume that the police do not care and will not come.[121] As a global attitude toward state institutions and officials, legal cynicism

---

116   *See* TYLER & HUO, *supra* note 72, at 26; Tyler, *supra* note 72, at 284; Jason Sunshine & Tom R. Tyler, *The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing*, 37 LAW & SOC'Y REV. 513, 534 (2003); *see also* Valerie Braithwaite, Kristina Murphy & Monika Reinhart, *Taxation Threat, Motivational Postures, and Responsive Regulation*, 29 LAW & POL'Y 137, 149 (2007) (showing that lower perceptions of procedural justice among a sample of Australian taxpayers who were surveyed on their attitudes toward the Australian Taxation Office were associated with greater resistance and reduced cooperation).

117   Tyler, *supra* note 72, at 284 fig.1, 286.

118   Jacinta M. Gau, *Procedural Justice, Police Legitimacy, and Legal Cynicism: A Test for Mediation Effects*, 16 POLICE PRAC. & RSCH. 402, 404 (2015); *see also* Robert J. Sampson & Dawn Jeglum Bartusch, *Legal Cynicism and (Subcultural?) Tolerance of Deviance: The Neighborhood Context of Racial Differences*, 32 LAW & SOC'Y REV. 777, 786 (1998) (defining legal cynicism as "the sense in which laws or rules are not considered binding in the existential, present lives of respondents").

119   Gau, *supra* note 118, at 404.

120   *Id.*

121   *See generally* Patrick J. Carr, Laura Napolitano & Jessica Keating, *We Never Call the Cops and Here Is Why: A Qualitative Examination of Legal Cynicism in Three Philadelphia Neighborhoods*, 45 CRIMINOLOGY 445 (2007).

might be shaped by a range of factors, including general social aliena-
tion, a lack of bonds to institutions, and negative interactions with var-
ious kinds of officials.[122] Satisfaction with law enforcement authorities
and views of the legitimacy of the criminal justice system, however, play
a particularly big role.[123] Essentially, legal cynicism represents a gen-
eral breakdown in trust between citizen and state, which in turn affects
people's willingness to obey the law and cooperate with law enforce-
ment.[124]

　　Social resistance theory is a fourth body of research connecting
trust and crime. Focused on lawbreaking among nondominant minor-
ity groups, social resistance researchers argue that feelings of social al-
ienation and experiences of injustice damage minorities' feelings of
attachment to the state, the law, and the larger society, which leads to
everyday behaviors of "social resistance," including criminal activity.[125]
Social resistance particularly, though not exclusively, predicts *malum
prohibitum* rulebreaking and crimes with no immediate victim—such
as, for example, vandalism or aggressive driving.[126] As one study con-
cluded, "lack of attachment to the country is correlated with social re-
sistance, which in turn is associated with a willingness to express this
resistance in different ways, which eventually is correlated with traffic
violations."[127] Notably, there is evidence that positive one-on-one ex-
periences with criminal justice officials can increase an alienated mi-
nority person's sense of belonging to the larger society and in turn his
or her willingness to abide by the law and cooperate with police.[128] In
other words, social resistance seems to be moderately sensitive to re-
cent experiences with state officials, particularly criminal justice offi-
cials.

　　Studies of rehabilitation within community courts is a fifth line of
research supporting the trust-crime connection.[129] The psychological

---

　　122   Amy E. Nivette, Manuel Eisner, Tina Malti & Denis Ribeaud, *The Social and Develop-
mental Antecedents of Legal Cynicism*, 52 J. RSCH. CRIME & DELINQ. 270, 272 (2015).

　　123   *See, e.g.*, Carr et al., *supra* note 121.

　　124   *See, e.g.*, Nivette, *supra* note 122, at 271. Additionally, based on a survey of attitudes
toward police and police service among 273 citizens in four neighborhoods in Pittsburgh,
Pennsylvania, Scaglion and Condon found that respondents' perceptions of the way in
which specific officers have related to them personally in previous encounters is a more
significant determinant of general attitudes toward police than were other socioeconomic
variables, including race and income. Richard Scaglion & Richard G. Condon, *Determinants
of Attitudes Toward City Police*, 17 CRIMINOLOGY 485, 486–87 (1980).

　　125   Factor et al., *supra* note 74.

　　126   *Id.* at 786.

　　127   *Id.* at 796.

　　128   Ben Bradford, *Policing and Social Identity: Procedural Justice, Inclusion and Cooperation
Between Police and Public*, 24 POLICING & SOC'Y 22, 37–38 (2014).

　　129   *See* Dancig-Rosenberg & Gal, *supra* note 56, at 358.

pattern is related to the one seen in social resistance theory: individuals living on the margins of society often feel uncared for or put upon by society, and those feelings lead to patterns of self-destructive and other-destructive behavior, which are often criminal. And, again, substantial evidence shows that criminal justice officials can, to some extent, turn those feelings around,[130] and that doing so has some rehabilitative effect.[131] As one community court participant remarked: "I don't like to do stupid things because everyone here supports me so much. . . . I don't want to disappoint."[132] The matter is not simple: people with long histories of mental illness, addiction, or bad choices do not always respond to trust-enhancing environments. But it works for some participants and, if one thinks of rehabilitation not as an all-or-nothing matter, but in realistically moderate terms—as an additional year without a relapse, say, or a better relationship with family members despite ongoing troubles—the evidence of rehabilitative effect becomes substantial.[133]

Research on restorative justice processes provide a sixth body of support for the idea that increasing trust decreases crime. Restorative justice processes are an important test case. They have had major effects on policy, particularly outside of the United States, and are

---

130   *See also* Mark Halsey & Melissa de Vel-Palumbo, *Courts as Empathic Spaces: Reflections on the Melbourne Neighbourhood Justice Centre*, 27 GRIFFITH L. REV. 182 (2018). *See generally* M. SOMJEN FRAZER, THE IMPACT OF THE COMMUNITY COURT MODEL ON DEFENDANT PERCEPTIONS OF FAIRNESS: A CASE STUDY AT THE RED HOOK COMMUNITY JUSTICE CENTER (2006).

131   *See, e.g.*, LEE ET AL., *supra* note 42 (showing, in an evaluation study of the Red Hook Community Justice Center in Brooklyn, that the chances of being rearrested two years after being processed were 10% lower for community court graduates than for defendants indicted for similar offenses who had been processed in other Brooklyn courts); STUART ROSS, EVALUATING NEIGHBOURHOOD JUSTICE: MEASURING AND ATTRIBUTING OUTCOMES FOR A COMMUNITY JUSTICE PROGRAM, 5–6 (2015) (finding a similar effect of 10% reduction in repeat conviction rates in a comparison conducted in Melbourne, Australia, between a local community court and a mainstream court); WESTAT, EAST OF THE RIVER COMMUNITY COURT (ERCC) EVALUATION 59 (2012) (reporting 42% reduced risk of being reconvicted among program graduates in an evaluation study of the East of the River community court program, in Washington, D.C.). However, some studies have found no significant effect. *See, e.g.*, Eric Grommon, Natalie Kroovand Hipple & Bradley Ray, *An Outcome Evaluation of the Indianapolis Community Court*, 28 CRIM. JUST. POL'Y REV. 220 (2017); LUCY BOOTH, ADAM ALTOFT, RACHEL DUBOURG, MIGUEL GONÇALVES & CATRIONA MIRRLEES-BLACK, NORTH LIVERPOOL COMMUNITY JUSTICE CENTRE: ANALYSIS OF RE-OFFENDING RATES AND EFFICIENCY OF COURT PROCESSES (2012); DARRICK JOLIFFE & DAVID FARRINGTON, INITIAL EVALUATION OF RECONVICTION RATES IN COMMUNITY JUSTICE INITIATIVES (2009).

132   Dancig-Rosenberg & Gal, *supra* note 101 (manuscript at 21).

133   *Id.*

currently being used in thousands of programs worldwide,[134] and their very premise, as discussed above, is trust-based. Whether they work in the sense of reducing crime has been the subject of extensive research, and the basic finding is that they work to a degree. Offenders who take part in well-operated restorative justice processes often report positive feelings of fairness, transparency, neutrality, and the opportunity to make a difference, which encourages them to cooperate with the reparation plan requirements and to engage in prosocial behaviors.[135] Studies consistently indicate the effectiveness of restorative processes in reducing recidivism rates, although the effect is typically modest.[136] Furthermore, "[s]ystem-wide restorative justice may potentially achieve general deterrence as well, by raising the level of certainty that offenses will be reported and handled."[137]

These six streams of research collectively represent a massive body of empirical study by different scholars using different methodologies and starting from different questions. Yet they come to a common conclusion: that increasing trust decreases crime, and that the behavior of criminal justice officials can affect levels of trust. Anyone who

---

134   Mark S. Umbreit, Robert B. Coates & Betty Vos, *The Practice of Victim Offender Mediation: A Look at the Evidence, in* INTERNATIONAL HANDBOOK OF PENOLOGY AND CRIMINAL JUSTICE 691, 692 (Shlomo Giora Shoham, Ori Beck & Martin Kett eds., 2008).

135   *See* HEATHER STRANG, REPAIR OR REVENGE: VICTIMS AND RESTORATIVE JUSTICE 170–77 (2002) (presenting findings from the Reintegrative Shaming Experiments comparing cases randomly assigned to either court or conference in Canberra, Australia); *see also* Tom R. Tyler, *Restorative Justice and Procedural Justice: Dealing with Rule Breaking*, 62 J. SOC. ISSUES 307, 308–09 (2006) (stating that respectful and inclusive processes are more effective than threatening, aggressive ones).

136   *See, e.g.*, James Bonta, Rebecca Jesseman, Tanya Rugge & Robert Cormier, *Restorative Justice and Recidivism: Promises Made, Promises Kept?, in* HANDBOOK OF RESTORATIVE JUSTICE: A GLOBAL PERSPECTIVE 108, 113–15 (Dennis Sullivan & Larry Tifft eds., 2006) (presenting a meta-analytic review of findings regarding recidivism rates following punitive and restorative justice processes, and concluding that restorative interventions have an approximately 7% reduction impact on recidivism); Jeff Latimer, Craig Dowden & Danielle Muise, *The Effectiveness of Restorative Justice Practices: A Meta-Analysis*, 85 PRISON J. 127, 137–39 (2005) (presenting a meta-analytic review of findings regarding the effectiveness of restorative justice programs in comparison with formal punitive processes. Effectiveness was measured by victim and offender satisfaction, restitution compliance, and recidivism rates. The meta-analysis concluded that restorative justice interventions resulted in small, but significant reductions in recidivism and were more effective with low-risk offenders.); Lawrence W. Sherman & Heather Strang, RESTORATIVE JUSTICE: THE EVIDENCE 68–71 (2007) (presenting a systematic review of evidence drawn from reasonably unbiased tests comparing court and conferencing processes in Canberra, Australia, and concluding that, with only one exception, rigorous tests of restorative justice showed significant reductions in recidivism rates).

137   Hadar Dancig-Rosenberg & Tali Gal, *Restorative Criminal Justice*, 34 CARDOZO L. REV. 2313, 2332 (2013).

cares about crime and public safety has good reason to want a metric that measures social trust and incentivizes increasing social trust.

3. Retributive Goals

Following Michael Moore, we use "retributivism" to refer to the family of theories that understand it to be a requirement of justice that people are held responsible as individuals for their choices to engage in wrongdoing, and are blamed or punished for that wrongdoing, to the extent deserved, and no more, without regard to the costs or benefits that might follow from blame or punishment.[138] Retributivism is thus a deontological theory committed to principles of anti-impunity, personal responsibility, desert, and proportionality, and it takes a morally absolute rather than relativist view of what constitutes right, wrong, responsibility, and proportional desert. Both the moral absolutism and the anticonsequentialism are at odds with the relativism and consequentialism of a metric based on social trust. Of all the prominent perspectives on criminal justice, this one might be furthest from social trust in spirit. Can this gap be bridged? Does a metric based on social trust have something to offer even a retributivist?

Our answer is that the gap can be bridged in a contingent yet stable way, based on a large body of empirical research about how typical members of the public think about criminal justice.[139] Led by the work of Paul Robinson, the basic thrust of these studies is that Americans, along with most other people in most other countries, are intuitive retributivists: the principles of anti-impunity, personal responsibility, desert, and proportionality defended by philosophical retributivists reflect—as an empirical matter—extremely widely shared attitudes and social norms. It follows that a system of criminal law that inspires social trust will reflect these principles.

For example, in one study, Robinson and Kurzban show that participants, when presented with a series of scenarios involving core criminal offenses and asked to rank them in terms of seriousness and proper punishment, reach a highly stable and widely agreed upon ranking, with little variation across virtually all demographic and

---

138   MICHAEL S. MOORE, PLACING BLAME: A THEORY OF THE CRIMINAL LAW 153 (1997) ("[R]etributivism is the view that we ought to punish offenders because and only because they deserve to be punished."). *See generally* ANDREAS VON HIRSCH, DESERVED CRIMINAL SENTENCES (2017).

139   *See generally* PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW (1995) (summarizing the research as it stood in 1995); Paul H. Robinson, *Democratizing Criminal Law: Feasibility, Utility, and the Challenge of Social Change*, 111 NW. U. L. REV. 1565 (2017) (summarizing the research as it stood in 2017).

national variables.[140] Similar studies by others concur in that result.[141] Based on their own studies and a review of many others, Robinson and Kurzban conclude:

> [A]vailable evidence suggests that human intuitions of justice about core wrongdoing . . . are deep, predictable, and widely shared. While there are disagreements about the relative blameworthiness of wrongdoing outside the core, the core wrongs themselves—physical aggression, takings without consent, and deception in exchanges—are the subject of nuanced and specific intuitions that cut across demographics.[142]

So great is the degree of agreement on these moral matters that it equals people's agreement about basic matters of perception, like whether a line is straight or a color is red.[143] Robinson and Kurzban thus argue that such rankings must be determined by an innate moral mechanism.[144]

Robinson and others have turned this body of accumulated empirical findings into a theory, which they term "empirical desert."[145] As opposed to deontological desert, empirical desert supports having rules of criminal liability and punishment reflect lay intuitions of retributive justice as a means of increasing compliance. As Robinson explains, "by tracking society's intuitions of desert, the criminal justice system increases its moral credibility and therefore its normative influence over individuals in the community. This normative influence can, in turn, result in material crime control gains."[146] In other words, people's retributive intuitions are so strong that effective crime control depends on having a criminal system that reflects them.

Three caveats are in order. First, people do not universally agree about wrongdoing outside the "core."[147] Murder and theft are one thing; marijuana smoking and copyright violations are another. Second, people do not universally agree about proper punishment in an *absolute* sense: they do not agree, for example, that the sentence for

---

140   Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829, 1867–80 (2007).

141   *See, e.g.*, Dogan D. Akman & Andre Normandeau, *The Measurement of Crime and Delinquency in Canada: A Replication Study*, 7 BRIT. J. CRIMINOLOGY 129, 144, 147 (1967); Sandra S. Evans & Joseph E. Scott, *The Seriousness of Crime Cross-Culturally*, 22 CRIMINOLOGY 39, 48–50 (1984); Sergio Herzog, *Public Perceptions of Crime Seriousness: A Comparison of Social Divisions in Israel*, 39 ISR. L. REV. 57, 59, 66 (2006).

142   Robinson & Kurzban, *supra* note 140, at 1892.

143   *See id.* at 1855.

144   *See id.* at 1892.

145   *Id.* at 1830 & n.1.

146   Paul H. Robinson, Joshua Samuel Barton & Matthew Lister, *Empirical Desert, Individual Prevention, and Limiting Retributivism: A Reply*, 17 NEW CRIM. L. REV. 312, 313 (2014).

147   Robinson & Kurzban, *supra* note 140, at 1892.

murder should be life imprisonment or anything of that sort.[148] What they agree about is the *relative* seriousness of various types of wrongdoing.[149] Third, there is disagreement in the field about the appropriate interpretation of the relevant studies. Some argue that, despite a consensus on the ordinal ranking of traditional crimes, there is no consensus about the appropriate punishments and that people are willing to depart from principles of desert provided the crime is not too serious and they believe preventive goals can be achieved in some other way.[150] Others argue that, "although moral judgments depend on numerous cognitive and physiological mechanisms that are presumably the product of evolutionary pressures, they are not innate insofar as they depend crucially on social meaning that varies across cultural groups."[151]

For present purposes, we are agnostic on these disagreements. What is striking about them, in our view, is the degree of scholarly nuance they reflect: even if the moral consensus around retributivism is not as total as Robinson and others writing in the same vein think, it is substantial enough that the scholarly disputes are about relatively fine distinctions and subtleties. In broad strokes, people really are intuitive retributivists. It follows that people will not agree with, and will not trust, a criminal system that regularly violates retributive principles. And it follows in turn that a metric based on social trust would give retributivists the anti-impunity, personal responsibility, desert, and proportionality they want.

Would the side effect be to make criminal law unduly harsh? That is, are people's retributive intuitions so punitive that social trust can only be sustained by great severity? Many suspect so; the view is encouraged by diverse media images of a bloodthirsty public. Yet further empirical research by Robinson and others appears to refute the stereotype. In one study, for example, participants were presented with a set of scenarios based on real criminal cases and asked to indicate the punishment they thought appropriate for the crimes described.[152] The average punishments given by survey respondents were dramatically *less* severe than the ones actually prescribed by law and given in the real cases.[153] For example, the actual court sentence given for possession of cocaine in one case was life without parole; the mean sentence

---

148   *Id.* at 1881.

149   *Id.*

150   *See, e.g.*, Christopher Slobogin & Lauren Brinkley-Rubinstein, *Putting Desert in Its Place*, 65 STAN. L. REV. 77, 94–95 (2013).

151   Donald Braman, Dan M. Kahan & David A. Hoffman, *Some Realism About Punishment Naturalism*, 77 U. CHI. L. REV. 1531, 1532 (2010).

152   *See* Paul H. Robinson, Geoffrey P. Goodwin & Michael D. Reisig, *The Disutility of Injustice*, 85 N.Y.U. L. REV. 1940, 1961 (2010).

153   *See id.* at 1972 tbl.4.

survey participants favored was 4.2 years.[154] The actual court sentence in a case of felony murder was life at hard labor without parole; the mean sentence laymen favored was 17.7 years.[155]

Based on these findings, the researchers argue that American criminal justice's present severity cannot be the product of the community's sense of justice—for it doesn't reflect the community's sense of justice. Rather, they argue, the harshness comes from misguided crime control motivations and distortions in the politics around crime and punishment.[156] Were punishment based on the retributive views most Americans actually hold, it would be far less harsh. If that is true, it follows that a metric based on social trust would not encourage the level of severity in criminal punishment America now has. It would encourage a moderate form of retributive proportionality in line with Americans' or other communities' moral intuitions.

4. Racial Justice Goals

Issues of racism and racial alienation are at the center of American criminal justice reform today, and any metric that speaks to all major voices in this space must address racial justice. But precisely because those issues are so front and center in contemporary America, it is important to appreciate that the problem of group-based oppression or alienation, particularly but not exclusively along racial or ethnic lines, is a problem of widespread and lasting significance in crime and punishment. Criminal law is and has been entangled with group identity—with racial or ethnic minorities, immigrants, religious dissenters, the poor, and beyond—throughout much of the world today, and throughout history. In France today, Arabs are disproportionately arrested and imprisoned;[157] in Germany, that is true of Turks;[158] a century ago in the United States, that was true of Irish and Italians.[159] Given the universality of the issue, it follows that a metric by which to measure

---

154   *Id.*

155   *Id.*

156   *Id.* at 1947–48.

157   *See* David A. Guba, *Cannabis Prohibition in France over the Past 50 Years Has Disproportionately Punished Its Muslim Minority*, RELIGION NEWS SERV. (July 13, 2022), https://religionnews.com/2022/07/13/cannabis-prohibition-in-france-has-disproportionately-punished-its-muslim-minority/ [https://perma.cc/2N39-V98B].

158   *See* THE OXFORD HANDBOOK OF ETHNICITY, CRIME, AND IMMIGRATION 5 (Sandra M. Bucerius & Michael Tonry eds., 2014).

159   *See* Christopher Muller, *Northward Migration and the Rise of Racial Disparity in American Incarceration, 1880–1950*, 118 AM. J. SOCIO. 281, 293–94 (2012); Carolyn Moehling & Anne Morrison Piehl, *Immigration, Crime, and Incarceration in Early Twentieth-Century America*, 46 DEMOGRAPHY 739, 759 (2009); *see also* WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE 15–40 (2011).

how well a criminal system is functioning—a metric that works not only in the contemporary United States but across time and place—should have some built-in way of highlighting problems of group-based oppression or alienation. One of the great strengths of a metric based on social trust, we submit, is the way it illuminates and prioritizes these problems.

Our argument starts from two assumptions. First, we assume a criminal system that takes social trust seriously. It uses a trust-based metric and its commitment to the metric is genuine: the officials staffing the system care about increasing social trust (among other reasons, because they are incentivized to care about it). Second, we assume that if a criminal system mistreats a racial minority, it will diminish social trust within that minority: members of the racial minority, when asked if they trust criminal justice officials, their government more generally, or their fellow citizens, will report lower levels of trust than if they were not mistreated.

What will happen in a criminal system that takes trust seriously when members of a racial minority report intense distrust of the system? First, criminal justice officials and the wider public will become *aware* that minorities feel alienated. That awareness should not be taken for granted: racial justice concerns might be hard to miss in contemporary American discussions of criminal justice reform, but they were utterly overlooked in earlier eras of American history, and they are missing in many other countries still. There is also an important distinction to note here between awareness that a topic is a matter of concern in the broader culture (for example, that it is discussed in classrooms and newspapers) and awareness in the sense that a fact is made part of the formal processes that structure institutional decisionmaking. As James Scott argues in *Seeing Like a State*, bureaucracies see what they can measure.[160] In addition, insofar as the results of the social trust metric are made public (as crime rates are now), low trust in the criminal system among minorities will be reported and analyzed by journalists and academics, making a democratic public aware of the problem. Indeed, the metric implies that low levels of trust among minorities are a problem—it "problematizes" low trust, to use the unlovely but helpful academic term—and that problematizing of low trust among minorities is a facet of awareness in itself. A metric based on social justice makes problems of race *visible*.

Second, when members of a racial minority report intense distrust in a criminal system that takes trust seriously, criminal justice officials will have reason to care about why minorities feel alienated and to direct policy—including, crucially, the allocation of resources—to

---

160   *See supra* note 3 and accompanying text.

repairing the conditions that lead to those feelings. Both their professional commitments and their incentives will require trust building. Presumably, officials will try to figure out why minorities feel as they do and make repairs they consider feasible. At the very least, in making cost-benefit calculations, officials will register diminished trust among minorities *as* a cost.

Consider, for example, stop-and-frisk policies in the contemporary United States. One of the greatest problems with those policies is that they communicate a message to black Americans: "You are an 'other,' a permanent suspect, and we, the dominant class, can use the police to investigate you at any moment." That is an intensely alienating message. But what existing metric has the capacity to see that alienation as a problem? Not arrest rates or clearance rates: they might go up with stop-and-frisk policies. Not crime rates: they might go down with stop-and-frisk policies, particularly in the short term. The cost in terms of racism and racial alienation is invisible to those metrics. But on a social trust metric, a police commissioner in a stop-and-frisk city will see data to the effect that stop-and-frisk policies diminish black Americans' trust in their criminal justice institutions and in government more generally. If, as we have assumed, trust-building is part of his incentive structure, that commissioner's career may depend on fixing the trust problem. At the least, when he weighs the costs and benefits of stop-and-frisk policy, he will count diminished trust among minorities as a cost. The commissioner will therefore have good reason to abandon stop-and-frisk policies or at least confront the hard choices of balancing questions of trust against questions of crime control. At bottom, what the social trust metric does is *translate* issues of race and bias from the realm of moral protest, where they can seem vague or easily disputed, into the realm of data that institutional bureaucracies find intelligible and incentives to which institutional actors are motivated to respond.

In fact, an affective, trust-based measure of racial problems in criminal justice has two advantages that other, nonaffective metrics (like disproportional incarceration rates) lack. First, measures of trust are agnostic with respect to questions of "actual" versus "perceived" racism. Measures of objective factors like disproportionate incarceration lead to arguments about whether the disproportion is justified by underlying rates of criminal offending, but if high social trust is the goal, those arguments become irrelevant. If rates of minority incarceration are causing black Americans not to trust their government, it is no answer to say that the incarceration is a result of disproportionate offending.

Second, one of the most pernicious effects of racism and perceived racism in criminal justice is the way in which the feelings of

anger and alienation that *start* within the criminal system don't *stay* within it. Those feelings are contagious, undermining the targeted minority's feelings about all of government and society. That in turn leads to the series of other consequences of low trust discussed in Sections II.A and subsection II.C.2: lower levels of economic success, participation in civil society, compliance with government urgings and directives outside of criminal law (such as getting vaccines), and, tragically and ironically, compliance with criminal law itself. Low trust *causes* elevated crime, which in turn leads to high rates of arrest and incarceration, which in turn leads to low trust, which in turn elevates crime—and so the cycle goes. A metric and incentives oriented to trust-building would break this most vicious of vicious cycles.

We can now imagine a series of scenarios involving fissures between minority and majority populations' levels of social trust—scenarios that present difficult and interesting questions about the *distribution* of trust in a society. Consider first a situation in which members of a racial minority report lower levels of trust in the criminal system than members of the racial majority, but the minority's level of trust can be improved without any concomitant decrease in the majority's level of trust. This might happen in a society with a large number of false convictions of minorities or slow response rates to 911 calls: minorities would not trust the system, but the fix—improvements in standards of evidence and procedure or police response times, for example—might well increase or at least not decrease trust within the majority as well. The solution in such a case is clear: the criminal system should do what it can to increase minority trust.

But now imagine a harder case, in which any increase in the level of trust among racial minorities would exact a concomitant diminishment of trust within the racial majority. This would be a society in which the perceptions or desires of the racial groups were so at odds that what the one wants or believes right is precisely what the other does not want or believes wrong. One might see this in a situation like the O.J. Simpson case, where the verdict that would reaffirm white faith in the criminal system was the very verdict that would undermine black faith in the criminal system. One might see the same structure in, for example, the Jim Crow South, where oppressing black people was precisely what the white population expected and wanted their criminal system to do. In such cases, trust-building in one group is trust-diminishing in the other. What should a criminal justice system oriented to trust do in that situation?

One option would be to undertake whatever policy would maximize social trust on the whole. On such a decision procedure, the racial majority would typically prevail just in virtue of being more numerous than the minority, but there could be exceptions: a criminal justice

policy with small effects on majority trust and large effects on minority trust could go the other way, depending on how big the minority group is and how strongly its members feel about the issue. The problem with this approach, in our view, is threefold. First, it ignores the possibility that the high levels of alienation among minorities might be based on actual injustice. One of the benefits of the social trust metric is the spotlight it puts on actual injustice (it is very difficult otherwise to figure out how to measure injustice) and the incentives it creates to make repairs; those benefits would dissipate if diminishment in social trust among minorities could just be outweighed. Second, pure trust maximization creates incentives to abuse minorities where doing so would please majorities. If we are thinking seriously and cynically about incentives, the incentive to curry favor with a racial majority by mistreating a racial minority should loom large. Third, pure maximization fails to recognize the socially destructive effects of radical distrust within any identity-based group. It is dangerous to create a situation in which a people or caste within a society feels that the society in which they live is their enemy. Criminal systems are unusual in the degree to which they can spur such feelings and have special reason to guard against the tendency.

Another option would be a trust-based version of Rawls's difference principle: changes in trust should be "to the greatest benefit of the least advantaged" members of society.[161] On such a principle, the criminal justice policies that prevail would be those that maximally benefit minority trust, without regard to effects on majority trust. The problem here, it seems to us, is not taking seriously enough the destructive effects of diminishing majority social trust. Imagine a policy of acquitting every member of a racial minority accused of a crime and convicting every member of a racial majority accused of a crime, regardless of the facts. Even if that policy would increase minority trust, it seems plainly unwise, as well as unjust.

We would endorse a third, "prioritarian" approach, in which policymakers typically focus on aggregate social trust but with a side-constraint or priority around the special problem of radical distrust among racial or other identity groups. In other words, the goal is to maximize aggregate social trust with the side-constraint of avoiding group-based feelings that the criminal system, government, or society generally is their enemy. Imagine, for example, a criminal system in which social trust is high within the majority population but an abused minority group utterly hates the system. Imagine a criminal justice policy that would enormously enhance trust within that small group, only modestly diminish trust within the larger group, but, because the majority

---

161   JOHN RAWLS, A THEORY OF JUSTICE 83 (1971).

is so much larger than the minority, reduce trust overall. The prioritarian approach would favor the policy nonetheless.

It should be emphasized, however, that these distributional questions only arise in distinctively pathological societal situations—situations in which the policies that would satisfy group A would offend group B, perhaps for the very reason that they would offend group B. There are many opportunities outside of those zero-sum situations to enact policies that would increase trust among minorities without decreasing it among majorities. A criminal justice official armed with a trust-oriented metric might find out, to use an earlier example, that a minority neighborhood does not trust the police to come when called and aim to quicken 911 response times. Outside of a pathologically biased society, the majority would not object; indeed, the reform might build trust for all.

5. Liberal Goals

By "liberalism," we refer, not to contemporary progressive political views, but to the philosophical tradition of treasuring individual autonomy and minimizing the oppressive potential of the state that has been a staple of criminal theory since John Stuart Mill.[162] Liberalism is a comprehensive political tradition, but in the criminal context, it derives its intellectual structure from the recognition that criminal law deals with *two* threats to liberty, both at a sort of apex: the risk of private violence and property invasion by other citizens, to which the state may legitimately respond with force, and the risk of violence and rights-invasion by the state itself, using what is traditionally the most oppressive tool in its arsenal—the criminal instrument. Liberalism is also suspicious of the repressive inclinations of popular publics, with their seemingly endless appetite for limiting individual freedom when physical or property harm is *not* on the line but merely offense, self-regarding conduct, or consensual other-regarding conduct, as with crimes of speech, vice, or private sexuality. With risks on all sides, the traditional liberal approach is to bind the criminal instrument in strict constraints, limiting criminalization to material harm, minimizing punishment to the extent feasible, putting criminal justice officials within structures of nondiscretionary positive law and process, and resisting popular passions to punish differently or more.

---

162   See the discussion of the harm principle above, *supra* notes 48–49 and accompanying text.

All this is in deep tension with a metric based on social trust.[163] Social trust grows on communitarian soil, focuses on affective benefits and risks, valorizes popular preferences, assumes a fair amount of official discretion, and ropes officials' incentives to popular preferences. A social trust metric would in practice encourage criminal justice actors to enforce the law according to extant norms in the culture. None of that is strictly inconsistent with liberalism, but it comes from a different intellectual universe, and it is easy to imagine why liberals focused on individual autonomy would not want to create that pattern of law enforcement. Is there a bridge? Is there good reason even for a criminal justice liberal to accept a trust-based metric?

Our answer is that liberalism as commonly understood greatly underestimates the role social trust plays in enabling people to live a self-directed life. That, after all, is the ideal: to live freely in a world that respects my rights and in which I respect the rights of others. But to live that way is actually to live in very high-trust conditions. Think of a person in a high-crime society who actively fears violence from the people around her. It is ridiculous to characterize that person as free. She is not free until two conditions obtain: she must take herself to be substantially safe from predation and she must in fact be substantially safe from predation. It is only when both conditions hold that she can take her life's walk through society in a way properly characterized as self-determining. The first of those conditions is trust; the second is the rational basis for trust.

Now think of a person living under a despotic regime who actively fears oppression from the agents of the state. It is, again, ridiculous to characterize that person as free. Real freedom requires that he believes that no agent of the state will arbitrarily abuse him and that he is right—or, again, trust and the rational basis for trust. In an individual's relationship to the state, there is in fact a third dimension as well: the individual must be able to trust that the state will not ignore crimes by others against him, that it will actively do its duty in extending to him the protection of the law. This too is a kind of trust.

Thus the stereotype of a liberal as someone whose political outlook is characterized by *distrust*, someone who regards both the state and other people with suspicion, gets things almost backwards. Suspicion might motivate the view. But as beings situated in societies, we are only free when we don't *need* to regard our fellow human beings or our government with suspicion, when we can rationally engage with others in our society in a spirit of equanimity. The word for that

---

163   We are grateful to Brenner Fissell for drawing our attention to this problem.

equanimity is trust. The necessary precondition of an individually autonomous life in a societal setting is a high degree of justified social trust.

That argument is fairly abstract. Here is how we think it would work concretely: a metric based on social trust would discourage criminal justice officials from unreasonable or abusive uses of power. For example, police officers would pay a cost in the currency of trust for engaging in searches designed to show dominance rather than to accomplish legitimate police ends. Likewise, the metric would discourage enforcement of laws criminalizing socially normal misbehavior, since most people would find it trust-reducing to disturb the social settlements that have evolved around crimes of vice. At the same time, however, the metric would *encourage* criminal justice officials to take violent and property crime seriously, since one of the things that most diminishes trust is the sense that police and other criminal officials don't care and won't help if one is victimized or if crime is high in one's neighborhood. Thus a metric based on social trust will tend to encourage—imperfectly to be sure, but better than any obvious alternative—vigorous but nonabusive enforcement of core criminal law. That is desirable from a liberal perspective.

One more point is in order: the search for an overlapping consensus—with its concomitant recognition of inevitable disagreement as to first principles and effort to find common ground despite such disagreement—is itself a liberal project.[164] To the extent a metric based on social trust does, as we argue, represent a site of overlapping consensus, that is an independent consideration recommending it to political liberals. As a matter of liberal principles themselves, they should recognize that social trust is a site of shallow buy-in for diverse points of view, including their own, and support it for that reason.

D.   Incentives

Three principles, we submit, are part of structuring incentives well in criminal justice. First, as in agency law generally, incentives should align the goals of institutional agents with the goals of the institution. Doctors should be incentivized to protect health, not to perform unnecessary procedures, because protecting health is the reason to have a medical system in the first place. In criminal justice, this means aligning prosecutors', police officers', and other criminal justice officials' incentives with the goals of the criminal system as a whole.

---

164   *See supra* notes 12–13 and accompanying text.

Second, in a context of multiple and conflicting goals, as in criminal justice, officials should be incentivized to care about all of them rather than to value some and ignore others. As discussed above, we want our criminal system to hold people responsible and show mercy, to protect the public and give offenders a second chance, and much more besides. Incentives shouldn't squeeze some of these goals out of the picture.

Third, incentives should incline officials to resist likely and destructive forms of capture. What those forms of capture are will depend on the activity. Bank regulators should have good employment options other than working for banks when they finish their terms of office. In criminal justice, this means resisting the tendency for prosecutors and, to a lesser extent, prison and police unions, to take over the field. We think it also means resisting ideological fads, which have been peculiarly powerful in the history of criminal law, and which have shown an ugly tendency to oscillate between extreme and unreasonable forms of harshness and extreme and unreasonable forms of mercy.[165]

With these principles in mind, compare the incentives facing, say, a traffic cop on a social trust metric versus a quota-based metric. Consider two scenarios: in the first, the officer observes someone driving over the speed limit but consistent with the speed of traffic; in the second, the officer observes someone driving over the speed limit to an extent that typical drivers would consider dangerous. On a quota-based metric, the officer is incentivized to ticket in both cases if he has *not* yet met his quota, and not to ticket in either case if he *has* met his quota; the two cases are, from the standpoint of the officer's incentives, exactly the same. But on a social trust metric, the officer has good reason not to ticket in the first case and to do so energetically in the second. Ticketing someone who is driving in socially normal ways would probably diminish trust for the ticketed person, other drivers on the road (who would have reason to think they too might be subject to tickets that seem arbitrary from a safety point of view), and even the officer himself (as he would have reason to see himself as acting cynically rather than protecting people's safety). Ticketing someone who is driving dangerously would likely increase trust among other drivers on the road, the officer himself, and plausibly even the offender.

Note that, in the example, the social trust metric gives police officers reason to focus on the public safety considerations for which traffic law exists in the first place. It encourages them to make distinctions

---

165   Observe the swing from extreme forgiveness in the 1960s to extreme antiforgiveness in the 1970s. *See generally* FRANCIS A. ALLEN, THE DECLINE OF THE REHABILITATIVE IDEAL: PENAL POLICY AND SOCIAL PURPOSE (1981).

between levels of wrongdoing and danger that are sensible from the point of view of social norms, even where the formal legal violation is identical. (We consider that a good thing, though we recognize some might not.) And it ties police officers' incentives to the sensibilities of the people participating in a given social sphere (other drivers on the road), who have an interest in upholding the norms governing that sphere. If there is some ideological movement among criminal justice professors about the goods or ills of ticketing for speeding, or among safety specialists about the dangers of fast driving, or within some other splinter group apart from drivers themselves, that group can only prevail over officers' incentives if they can first persuade the public to agree. Reformers might be able to do so (in fact, have done so) with respect to driving drunk or without a seatbelt. But there is an interface of public acceptance between the ideological movement and the control of incentives, which resists ideological capture. Again, we consider this a good thing, though we recognize that others might disagree.

The example generalizes. Imagine a much more serious case: a neighborhood in which drug use and low-level drug-dealing is common, but some major drug-dealers lead gangs and are sources of significant social harm. And let us this time compare the social trust metric to arrest- and clearance-rate metrics, and do so from the standpoint of a police department deciding how to deploy resources, rather than from the standpoint of an individual officer. On an arrest- or clearance-rate metric, the department would have reason to maximize the quantity of drug busts and therefore pursue lots of easy-to-prove cases against low-level offenders. It would have reason *not* to pursue the major, gang-associated dealer if, as is likely, that case would require comparatively expensive, sustained, and uncertain investigation. On a social trust metric, the department's incentives would be reversed: it would go after the major dealer and, assuming the low-level offenders are not causing community alarm, would not go after the minor offenders. The result would be a smaller quantity of higher-quality arrests—which strikes us as good in both directions, reducing the number of people incarcerated or going about life with criminal records (spreading criminal convictions around widely is not a good thing), and increasing the seriousness of the offenders who are arrested. The structure is the same as the traffic example: police departments would have reason to focus on the considerations of safety, wrongdoing, and community well-being for which criminal justice exists in the first place.

Transition now to the perspective of a prosecutor. Imagine two cases. In one, the crime was horrible, the offender is dangerous, and the evidence is good enough to satisfy the reasonable doubt standard but not a slam dunk. In the other, the crime was minor, the offender

salvageable outside the criminal process, but the evidence is undeniable. A prosecutor measured by her conviction rate has reason not to risk bringing charges in the first case and not to miss the opportunity to secure an easy conviction in the second. But that is exactly the opposite of what we should want from a societal standpoint: assuming constitutionally adequate evidence, we want prosecutors to pursue socially serious cases, not easy-to-prove ones, and to decline slam-dunk cases where punishing the offender would be unnecessary given the purposes for which punishment exists. The conviction rate metric also incentivizes the prosecutor to plead every case she can, since plea bargains are the one certain path to conviction, and, further, to overcharge or threaten excessive sentences in order to bring the defendant to the table and undercharge or offer excessively lenient sentences in order to get the defendant to agree to the deal. But from a societal standpoint, we should want offenders to be charged and sentenced based on what they actually did.

Now, what are that same prosecutor's incentives if measured by whether her decisions would build or diminish trust among stakeholders in the case and other members of the community? Assuming that people are upset by the serious case and not, or not equally, by the less serious one, her incentives are reversed: she would have reason to bring charges in the serious case and to decline the less serious case (if prosecuting would offend ordinary sensibilities) or at least to focus preferentially on the more serious one (if people would care more about the first but want to see both prosecuted if possible). She might still want to plea bargain and avoid trial, but not necessarily, and not 100% of the time: she might judge that a trial would show the public that justice is being done in the cases they care most about—a trust-building move. Even if she does plea bargain, she would have reason to keep the threats and offers within the bounds of what fully informed members of the public would consider reasonable given the nature of the crime: if someone commits what is in substance involuntary manslaughter, presumably members of the public would not trust a system that threatens first-degree murder and agrees to second-degree assault.

Thus, again, a trust-based metric aligns official incentives with the higher-order purposes for which criminal law exists in the first place—particularly if one has a solidaristic conception of criminal law's purposes, in which case the alignment becomes almost a necessary truth. It encourages moral nuance in officials' thinking, reflecting the multiplicity of goals in criminal justice. And it anchors official incentives to community sensibilities, which act as a guardrail against ideological or insider capture.

E.   Objections

1.   Bad Social Norms

A metric based on social trust links questions of whether a criminal system is successful to how people in the community *regard* it, which anchors one's measure of success, relativistically, to the social norms prevalent in a given society at a given time. But what if a society's norms are just *wrong*? Is it our claim that a criminal system that accords with bad norms is more successful than one that insists on better norms? In the Jim Crow South, a criminal system might inspire trust by reflecting values of white supremacy. Perhaps our former response about the distribution of trust is adequate to that case, but then change the example. What if a society's norms are too tolerant toward something that we have come to regard as wrong even if it doesn't particularly alienate a minority group (e.g., brawling, dueling, or public corruption), or too intolerant toward something that we have come to regard as not wrong or not so wrong as to warrant major punishment (e.g., vagrancy or gambling)? Isn't a well-functioning criminal system one that gets these fundamental moral questions *right* and, where social norms deviate from justice, pushes back against society's bad norms?

Our response is that criminal law should be a site of moderate relativism—relativism with justice-based side constraints. It is also that the objection misunderstands the special character of criminal law and exemplifies a fallacy known as the "nirvana" fallacy.[166]

Start with the last of these. The nirvana fallacy is the fallacy of constructing a choice such that an actual, imperfect institutional arrangement is implicitly measured against a hypothetical, idealized arrangement[167]—like criticizing the policy pitfalls of democratic government without ever asking whether undemocratic governments are better. The "bad social norms" objection does just that: it implicitly imagines that there are true and just positions on matters of criminal justice policy, that they are knowable and known, and that the relevant policy choice is therefore between a criminal system that fits the norms of its society and one that reflects perfect justice. If that were the choice, it probably would be unreasonable to opt for anything but perfect justice.

But that is not the choice. The choice is between a criminal system that reflects extant social norms and one that reflects an ideology of right and wrong contrary to extant social norms. Or still worse: the

---

166   Harold Demsetz coined the term in *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 1 (1969).

167   *Id.*

choice is between a criminal system that reflects extant norms and one that reflects whatever ideology contrary to extant norms is likely to get political control of the criminal system at any given time. But there is no reason to think the best of the available ideologies will prevail in real political competition: even if we could know to a moral certainty that, say, utilitarianism is enlightened and retributivism benighted, retributivists may defeat utilitarians in the struggle for control of criminal justice institutions. And we never actually know to a moral certainty which ideology is right. To say, "What if a society has bad social norms?" is implicitly to say, "What if a society's social norms violate my favored ideology?" The choice, in other words, is not between extant norms and perfect justice but between extant norms and an advocate's conception of better norms. Extant norms have at least the benefit of being culturally functional, securing societal adherence, and resisting the totalizing and faddish tendencies of academic orthodoxy. Really, what the objection betrays is a deep failure of self-skepticism.

The objection also fails to appreciate why a constrained relativism might make sense given the distinctive social function of criminal law. In assuming that criminal law should transform societies in the direction of justice, the objection neglects a set of costs associated with keeping the peace, maintaining social solidarity, and otherwise enabling societies to function. "Justice is the first virtue of social institutions, as truth is of systems of thought," John Rawls wrote.[168] "A theory . . . must be rejected or revised if it is untrue; likewise laws and institutions . . . must be reformed or abolished if they are unjust."[169] But as one of us has written in response: "[T]o think that is to undervalue the *sharing* of ideas, regardless of whether they are wholly right or just, for it is only with shared ideas that we can function socially."[170]

Imagine that a colonial power, like Britain in India, imposed a liberal criminal law on a caste-based society. Would that be an unqualified good—not just good on balance, but *unqualifiedly* good? Surely not; surely that law, whatever its benefits, would exact costs in the currencies of legitimacy and peace, pitting the members of an inegalitarian culture against their government and, potentially, against one another. Are those benefits worth it? Perhaps in extreme cases, they are worth it, but that is the wrong way to think about criminal law generally. Criminal law as we see it is distinctively equipped to play a vital social role: maintaining social solidarity and keeping the peace on the basis of shared norms, and it cannot perform that function unless it is

---

168   RAWLS, *supra* note 161, at 3.

169   *Id.*

170   Kleinfeld, *supra* note 52, at 1494.

responsive to shared norms.[171] Unless shared norms happen to accord with perfect justice (which is to say, never), a criminal law that embodies perfect justice therefore cannot perform the function for which criminal law exists.

Note that, while this implies a certain relativism *within* criminal law, it does not imply relativism generally. One can have nonrelativistic views about what justice requires, as we do, and advocate changing society to make it more just, as we would, while still thinking criminal law has a special function to perform that properly ties it to extant social norms. That is not a contradiction. It just means one should fight for a more just society while using criminal law minimally to enforce norms that are already widely established.[172] As one of us has argued elsewhere, criminal law must be a relativistic instrument "nested within a larger commitment to justice."[173]

But what about radically unjust societies? Moderate or local relativism of the kind we propose presents difficult questions in societies with imperfect norms; it presents profound and perhaps insurmountable problems in evil societies. Imagine a violently theocratic society whose criminal law inspires trust by committing abominations on heretics. Should we even then call the criminal system a success?

We acknowledge the limits of the social trust-based metric in such limit cases. If one is not a complete relativist, and we are not, the metric must be used with justice-based side constraints, like "no increasing trust by punishing the innocent" and "no increasing trust by engaging in brute violence toward the unpopular." Further, we acknowledge that a trust-based metric doesn't make sense at all in a society so morally deficient that its norms are not worth preserving. A metric based on trust makes sense on certain assumptions: the society's norms must be modestly decent and certain constraints of justice must be accepted regardless of their effect in a particular case on people's levels of trust. It might be that some radical critics of American criminal justice, particularly some criminal law abolitionists, do not share those assumptions. But that, we submit, is not a problem with the social trust metric. Were we proposing a metric by which to measure the performance of an oil company, there would be a lack of joinder with those who think fossil fuel companies should not exist. It means only that our argument rests, as every argument does, on its assumptions.

---

171   This view of criminal law is defended in the theory of criminal law known as "reconstructivism." *See id.* at 1486, 1506.

172   Again, this view is defended at length in *Reconstructivism: The Place of Criminal Law in Ethical Law. Id.*

173   Joshua Kleinfeld, *Why the Mind Matters in Criminal Law*, 53 ARIZ. ST. L.J. 539, 544 (2021) (emphasis omitted).

2. Popularity Contests and Noble Lies

A metric based on social trust treats a *popular* decision as a *good* decision, and, by the same token, incentivizes criminal justice officials to do what is popular regardless of what they might otherwise think is right. But isn't it often the duty of a civil servant—especially a prosecutor, and even more a judge—to *resist* what is popular, to be independent, to treat criminal defendants with dispassionate fairness even when the public is enraged, to do what is right regardless of public pressure? It has always been a judge's courage to uphold the rights of criminal defendants even when those defendants are hated.[174] Neutral professionalism, unmoved by public sentiment, is a common ideal among police and prosecutors.[175] Does a trust-based metric imply rejecting those ideals? Would it in practice undo them?

There are variations on the theme of this objection. A metric based on social trust might incentivize criminal justice officials to use media techniques to manipulate perceptions. It might encourage officials to neglect the constraints of good evidence and due process. It might reward them for using their discretion to aggressively target, scapegoat, and excessively punish whatever class of offenders is out of favor given the politics of the moment (the ideologically unpopular, the rich, a religious or racial minority, drug-dealers during a drug epidemic, bankers after a financial meltdown—there is always someone). Or it might simply favor style over content, being appealing over doing a good job. Procedural and restorative justice approaches to criminal law have been criticized on similar grounds: justice and effectiveness are not just matters of making people feel good.

A last variation on the objection—the problem taken to its logical extreme—is the noble lie. If the goal is simply to increase public trust, why not convict the innocent, so long as the conviction will, say, make people feel safe, or satisfy their need for accountability? It is trust-reducing when police officers can't solve widely known crimes. Should we therefore keep them from being widely known? Should we pretend to have solved them? Is the only constraint on these practices the risk that the secret might get out?

There is no easy answer to this line of objection. It draws blood. But we must set our expectations appropriately. Most major

---

174   *See* Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175, 1180 (1989) ("Judges are sometimes called upon to be courageous . . . . Their most significant roles, in our system, are to protect the individual criminal defendant . . . .").

175   *See Legitimacy Policing in Depth*, RAND CORP., https://www.rand.org/pubs/tools /TL261/better-policing-toolkit/all-strategies/legitimacy-policing/in-depth.html [https:// perma.cc/N5MP-9X5V]. *See generally* Bruce A. Green & Fred C. Zacharias, *Prosecutorial Neutrality*, 2004 WIS. L. REV. 837, 840 (2004).

governmental arrangements carry risks and costs; most, if presented as proposals, would be subject to good objections for which there is no easy answer. Separation of powers comes at the cost of inefficiency. Judicial review comes with the risk of judicial abuse. Security forces can protect but also oppress. That is the way of the world. We do not abandon separation of powers, judicial review, and security forces because they are subject to objection, but strive to find ways to contain and mitigate their risks and costs, and then accept or reject them relative to alternatives and all things considered.

There are ways to contain and mitigate the risks and costs of the popularity contest. The first is to recognize, as with the previous objection, that a metric based on social trust must operate in a context of justice-based side constraints instantiated as rules. In effect, we maximize social trust within a framework of rules. For example, police cannot lawfully arrest someone merely for doing something unpopular or offensive; the trigger for an arrest is a crime. Thus the penal code creates a system of rules; police incentivized to make socially acceptable decisions can (or should) only be able to exercise their power within the context of those rules. Likewise, prosecutors can properly indict someone only if the evidence shows guilt beyond reasonable doubt. Judges can sentence only as provided by law. If the framework of rules is sound, and adequately enforced, a metric based on social trust will allow some aggression or leniency around the edges, but will not allow terrible abuse.

Second, considerations of *long-run* trust and trust *on the whole* can restrain abuse. In the short run, a show trial of an innocent may satisfy the public—but secrets have a way of getting out, and secrets like that shatter trust in the future. Media manipulation, style without content, scapegoating, and other abuses likewise tend to erode trust over time. A judge who resists public pressure may be unpopular in the short term and individually, but, in the long run, judges as a class are trusted precisely because they show such courage. In the short run, the public might want police and prosecutors to brutalize defendants as an expression of public rage; in the long run, police and prosecutors are trusted, when they are trusted, because they maintain professionalism in the face of such feelings.

The real challenge is to get officeholders to focus on the long-term. There may be technical solutions to that challenge. Perhaps empirical measures of trust should not be considered valid, or even measured, with respect to individual officeholders or decisions, but only with respect to institutions as a whole and over time (not "do you think indictment X was fair?" but "do you think prosecutors in your town are generally fair?"). Perhaps criminal justice officials should have safe harbor from trust-based measures to the extent they credibly

show their unpopular decisions were grounded in law. Perhaps high-ranking managers could be encouraged to take a long run and on-the-whole perspective even if line officers are not. In any case, it is striking that social trust is not without responses to the popularity objections even without turning to values outside of social trust. There are *internal* grounds—considerations of social trust itself—for resisting abuses.

Another internal protection against the popularity contest is that the social trust metric takes into account the views and interests of so many different stakeholders. Victims' opinions matter, but so do the opinions of offenders and their families. Perceptions within the general public matter, but so do the views of minority groups. Consider the impact of these different stakeholders from the standpoint of, say, a police chief. Perhaps the public is upset about gang violence. A police chief would be incentivized to focus on making gang-related arrests. Yet, in doing so, he also has reason to think about the views of the people he is arresting and their families. He has reason to consider whether his tactics might offend people in the neighborhood of the gang arrests. He has reason to consider whether his arrests might carry racial overtones. Trust in criminal justice is a zone of conflicting values. The yield of an internal balancing of views and interests is a form of checks and balances, which complicates and moderates officials' incentives.

We should also bear in mind the virtues as well as the vices of trying to appeal to the public, and the vices as well as the virtues of insulated, neutral professionalism. Consider again the example of a public upset by gang violence. Is it really better for a police chief not to care about that demand? To focus on easy-to-prove arrests that keep arrest rates high but don't matter for gang violence? The professional and dispassionate prosecutor or police officer is not the only thing one gets when the pressure to care about what the public thinks is lifted. Just as likely is an indifferent one, or a lazy one, or one motivated by an ideology foreign to the community he is supposed to protect. A judge totally independent of the need to answer to the public is not necessarily a brave defender of rights. She could just as easily be harsh (a hanging judge), or so mild as to exhibit a distorted sense of justice, or legalistic in ways that needlessly obscure questions of guilt and innocence, or enthralled to an ideology foreign to the law itself. Probably the most likely consequence of immunity from public pressure is just inefficiency.

Finally, if the effort to promote social trust means that the emotions felt by the public have some impact on how the criminal justice system is run, that is not a bad thing. Emotions matter in criminal justice—more, and more legitimately, than in many other areas of law. The call-and-response of wrongdoing and condemnation, punishment

and expiation is an emotional exchange, and it is not possible for the system to perform its necessary social functions if the officials of the system are indifferent to those emotions. The influence of those emotions should be constrained by law. But within those constraints, they have a place.

CONCLUSION: TOWARD AN INSTRUMENT

Our goal in this Article has been to explain *why* a metric based on social trust should take pride of place in assessing criminal systems. We plan to explore questions of how to design and implement the metric in work to come. [176] Yet with a view to that future work, we would like here at the conclusion to preview some of the relevant design and implementation considerations at least sufficiently to allow one to imagine how the instrument might work.

A first question is whether to assess trust by means of surveys (essentially, asking people questions to determine how they felt about a contact with the criminal system) or to find some objective behavioral correlate of trust (perhaps people's revealed willingness to call 911 in conditions of distress). Now, the two are not mutually exclusive; we would favor doing both. Yet, with all due acknowledgement of how fallible people's self-reporting can be, we think surveys should take center stage in this context. For one thing, surveys are standard fare in the existing social trust literature (research on social trust's relationship to economic productivity depends on surveys, for example).[177] It is difficult to think of objective behavioral correlates that would, in the criminal context, be adequate measures of trust. It does not seem likely that people would be mistaken or dishonest, even to themselves, about the effect of a criminal justice contact on their levels of trust. And, finally, there is no need in this context to avoid the messiness of thoughts, feelings, and perceptions. Thoughts, feelings, and perceptions are the very thing we are interested in establishing. The subjectivity is the point.

---

176 Recently, the Organisation for Economic Co-operation and Development (OECD) published detailed Guidelines on Measuring Trust, as part of the work program of the OECD Statistics Committee and the Public Governance Committee. *See* ORG. FOR ECON. COOP. & DEV., OECD GUIDELINES ON MEASURING TRUST (2017). The background for formulating these guidelines was the need pointed out by several recent policy initiatives in Europe for having better measures of trust. *Id.* at 11–12. The OECD guidelines, which do not focus on measuring trust in the criminal context specifically, aim "to support data producers in their own initiatives to measuring trust." *Id.* at 3. These guidelines synthesize what is known today, based on existing literature, about "good practice on how trust can, and should, be measured." *Id.*

177 *See supra* note 14.

A next question is *whose* attitudes to measure. Crimes commonly generate concentric circles of stakeholders: suspects, defendants, and their families; victims and their families; witnesses; criminal justice officials themselves (for example, the police and lawyers who work on the case); community members with relatively direct connections to the case (for example, coworkers in the victim's or offender's workplace); and community members with relatively indirect connections to the case (for example, the victim's or offender's neighborhood or city). Should we measure all of them? Should we weight them all equally (thus giving trust-building among offenders and trust-building among victims equal weight)? Once one gets away from the immediate parties to the case, people might not know about the crime or the response to it; should researchers inform them before asking them whether they think the system behaved properly in that instance? We acknowledge that we have not settled all these issues yet. However, we suspect that surveying all of the different stakeholders throughout the concentric circles is worth doing, precisely because they have such different interests, perspectives, and information. The multiplicity could be illuminating. It would be interesting to know, for example, if lawyers believe the system is working and victims do not, or vice versa. And it might be possible to build trust within different groups of stakeholders in ways that do not reduce it among other groups. Perhaps defendants' levels of trust might be affected by whether they are treated with procedural fairness and respect, while victims care about perpetrators being caught and prosecuted, and officials about whether the case was handled equitably in comparison to past cases. None of those are mutually exclusive; it would be possible to build trust for all. As to randomly selected citizens, their perspective seems central to assessing ambient levels of societal trust toward the criminal system and to establishing appropriate incentives for criminal justice officials (it might be helpful from an incentive standpoint for police and prosecutors to know that any particular decision they make might later become the subject of a trust-oriented survey with the public).

Concretely, then, our plan would be to survey those who have direct contact with the criminal justice system, other stakeholders, and randomly selected members of the public. Questions might include:

1. To what extent do you feel that justice was done in this case?

2. How fairly do you feel the officials involved in this case behaved?

3. How respectfully do you feel the criminal justice officials involved with this case behaved?

4. Do you feel more or less trust in your government now that this case has concluded?

    5.   Do you feel more or less trust in your fellow citizens now that this case has concluded?

    6.   How much would you trust officials in a future criminal case in which you had a stake?

Much remains to be done to make a survey instrument of this kind a reality. But we hope this Article has demonstrated the need for such an instrument and why social trust is the right thing to measure.