

Divine Justice and the Library of Babel: Or, Was Al Capone Really Punished for Tax Evasion?

Gabriel S. Mendlow*

Abstract

A criminal defendant enjoys an array of legal rights. These include the right not to be punished for an offense unless charged, tried, and proved guilty beyond a reasonable doubt; the right not to be punished disproportionately; and the right not to be punished for the same offense more than once. I contend that the design of our criminal legal system imperils these rights in ways few observers appreciate. Because criminal codes describe misconduct imprecisely and prohibit more misconduct than any legislature actually aspires to punish, prosecutors decide which violations of the code merit punishment, and judges decide how much punishment specific violations warrant. In making these decisions, prosecutors and judges rely on evidence beside that which is necessary to sustain a conviction, including evidence of an offender's extraneous transgressions. This practice calls into question whether offenders are being punished for the offenses of which they're formally convicted or instead for the extraneous transgressions that inform the exercise of official discretion. As I argue, theory and common sense alike suggest that the offense an offender is punished for is determined less by the formal features of the criminal process than by that process's social meaning, which itself is determined at least in part by the motivations of the participating legal actors. Because we lack a sound basis to resolve these matters, we may not know how often our legal system dishonors the rights it proclaims sacrosanct.

* Professor of Law, University of Michigan. I'm grateful to audiences at all the places where I presented at least some of the material in this paper: Dartmouth, the Harvard/Yale/Stanford Junior Faculty Forum, Hebrew University, Michigan, Northwestern, Tel Aviv University, UCLA, and York University. Special thanks to Russell Christopher, Rich Friedman, Stephen Galoob, Sam Gross, Monica Hakimi, Alon Harel, Scott Hershovitz, Don Herzog, Doug Husak, Dan Kahan, Joshua Kleinfeld, Jae Lee, Kyle Logue, Bill Miller, Eric Miller, J.J. Prescott, Eve Primus, Richard Primus, Don Regan, Adina Roskies, Scott Shapiro, Seana Shiffrin, Sonja Starr, Eric Swanson, and Peter Westen.

I.

A criminal defendant has the right not to be punished for an offense unless charged, tried, and proved guilty beyond a reasonable doubt. If convicted, he has the right not to be punished disproportionately, or to be punished for the same offense more than once. He also has the right not to be punished for his thoughts, for his character, or for his race, ethnicity, or religion. I contend that the design of our legal system imperils these rights in ways few observers appreciate.

Our system is a hybrid of two widely divergent ideal types. One is the *Library of Babel model*,¹ which employs a criminal code of infinitely fine texture, a code that specifies with complete precision every conceivable wrong that a just state would want to punish. The Code of Babel obviates the need for prosecutorial discretion because every offense definition includes a description of the circumstances that would warrant bringing a charge, as well as those that would militate against it. Every offense definition in effect contains myriad “unless” clauses describing circumstances that would negate the occurrence of a wrong worth punishing. By design, these circumstances are the very ones that would lead a just prosecutor not to bring charges in the exercise of her sound discretion. Thus, the Code is crafted such that a prosecutor who pursues every provable violation will bring no unwarranted charges. The Code of Babel also obviates the need for sentencing discretion. Because the Code specifies every offense precisely—incorporating within each offense definition every circumstance that might have a just bearing on the sentence an offender should receive—the Code associates each offense with a specific penalty.

The other ideal type of criminal legal system is the *Divine Justice model*, which employs a criminal code of the coarsest texture and relies entirely on the discretion of prosecutors and judges. The Code of Divine Justice contains but one offense: bad conduct. Virtually every adult is guilty of this offense at one time or another, so prosecutors must scan the populace and select the worst offenders for prosecution—anyone who in some way merits state punishment. Unlike the provisions of the Code of Babel, which identify and denounce perfectly-specified acts of wrongdoing, the single offense set out in the Code of Divine Justice serves simply as an intake valve. Its function is to bring persons before a court that acts as an arbiter of divine justice, handing down sentences that impose punishment in one fell swoop for a lifetime of transgressions, both discrete and episodic, as well as for bad character and bad thoughts, all discounted by good deeds, lousy childhoods, and every other conceivable mitigating factor.

Our criminal legal system resembles the Divine Justice model more closely than we acknowledge and therefore imperils defendants’ rights more often than we appreciate. Because our criminal codes describe misconduct imprecisely and prohibit more misconduct than any legislature actually aspires to punish, prosecutors

¹ See Jorge Luis Borges’ short story “La Biblioteca de Babel” (“The Library of Babel”) about an infinite library containing every possible book.

decide which violations of a code merit punishment, and judges decide how much punishment specific violations warrant. In making these decisions, prosecutors and judges rely on evidence beside that which is necessary to sustain a conviction, including evidence of an offender's extraneous transgressions. This practice calls into question whether offenders are being punished for the offenses of which they're formally convicted or instead for the extraneous transgressions that inform the exercise of official discretion.

As this essay argues, theory and common sense alike suggest that the offense an offender is punished for is determined less by the formal features of the criminal process than by that process's social meaning, which itself is determined at least in part by the motivations of the participating legal actors. Because we lack a sound basis to resolve these matters, we may not know how often our legal system dishonors the rights it proclaims sacrosanct. When we don't know which offense an offender is being punished for, we can't know whether he is being punished disproportionately, whether he is being punished more than once for the same offense, whether the offense for which he is being punished is the offense of which he was charged, tried, and proved guilty beyond a reasonable doubt, and whether he is being punished for his thoughts, for his character, or for his race, ethnicity, or religion.

II.

Punishment by its nature takes an object.² Actually, punishment takes two objects: the transgression *for which* it's imposed and the person *on whom* it's imposed. Punishment's logical structure therefore mirrors that of its psychological cousins, anger and resentment. In the prototypical case, you are resentful *of* some person *for* some transgression (real or perceived). Often the transgression is temporally bounded, being a discrete act or omission to act. But you can resent a person for a series of misdeeds, or for a lifetime of inaction, or for a failure to love or care or connect, or simply for being the kind of person she is. Even in the last case, your resentment has two objects: the person and her transgression (namely, being that kind of person). Resentment's clearest counterpart—gratitude—also takes two objects: in the prototypical case, you are grateful *to* a person *for* some good deed.³ So too does punishment's clearest counterpart: the practice of bestowing prizes or awards. Prizes always are bestowed *on* someone *for* something, even when the thing they're bestowed for is as diffuse and unbounded as a lifetime of achievement.

² See Gabriel S. Mendlow, *The Elusive Object of Punishment*, LEGAL THEORY (Forthcoming 2019).

³ Other attitudes have a simpler logical structure. Love and hatred, affection and dislike—these attitudes prototypically take a single object: a person or other entity. Attitudes like happiness and sadness may take no object at all. You might just *be* happy or sad, but not about anything in particular.

The logical structure of punishment, like that of any social institution, is a human artefact and therefore impermanent. If we wanted to, we could discard punishment in favor of an outwardly similar institution that does nothing but incapacitate the dangerous. The institution would mete out treatment with the look and feel of punishment but with a different logical structure, being imposed for the sake of future consequences rather than for past transgressions. Some theorists would like to see an institution of this sort supplant legal punishment,⁴ and many officials have sought unabashedly to administer our existing penal institutions so as to achieve preventive ends. Yet they've done so almost always without betraying punishment's logical structure: they've pursued preventive ends *by* punishing offenders for supposed transgressions. The internal logic of punishment has held firm. Even if punishment is often imposed for the sake of future consequences, it's always imposed *for* past transgressions.⁵

But for *which* transgressions? Since the publication of George Fletcher's 1994 essay "What Is Punishment Imposed For?,"⁶ virtually no English-language scholarship has addressed Fletcher's titular question⁷—possibly because no scholar has seen reason to dissent from Fletcher's anodyne answer: "[Punishment] is imposed for the act of wrongdoing, *the unjustified violation of a statutory norm.*"⁸ This conception of the object of punishment might appear unassailable, as it seems to follow from an incontrovertible proposition of criminal justice: that a person may be punished only if proved guilty of violating a criminal statute. From this banal truth, it seems to follow that violating a statute is what an offender is punished *for*. But this way of thinking is fallacious, because it neglects at least two important sources of uncertainty about what an offender's punishment is being imposed for: the discretion to charge and the discretion to sentence. Both sorts of discretion create the possibility that a given offender is being punished for something other than, or something in addition to, his "unjustified violation of a statutory norm."

In what follows, I'll focus at first on the discretion to charge. Once the relevant concerns are in view, it won't take long to raise similar concerns, perhaps even graver ones, about the discretion to sentence. The basic problem is this: just as I might purport to resent you for one transgression while actually (if obliquely)

⁴ See, e.g., BARBARA WOOTTON, *SOCIAL SCIENCE AND SOCIAL PATHOLOGY* (George Allen & Unwin Ltd., 1st ed.1959).

⁵ John Gardner, *What is Tort Law For? Part 1. The Place of Corrective Justice*, 30 *LAW & PHIL.* 1, 7 (2011) ("[A]nyone who begins their evening out by saying 'let's punish some people tonight' is making no sense until we get an answer to the question 'Punish them for what? What are they supposed to have done?'").

⁶ George P. Fletcher, *What is Punishment Imposed For?*, 1994 *J. CONTEMP. LEGAL ISSUES* 101.

⁷ I'm aware of but one publication on Fletcher's question: Douglas Husak's short essay, *The Importance of Asking the Right Question: What is Punishment Imposed For?*, in *ESSAYS IN HONOR OF GEORGE FLETCHER* (Russell Christopher ed., 2012).

⁸ Fletcher, *supra* note 6, at 105 (emphasis added).

expressing resentment for another, so might the state purport to punish an offender for his nominal offense of conviction while actually censuring and sanctioning him for something else.

III.

Because our criminal codes are coarse-grained, not every instance of every offense can or ought to be prosecuted. Because not every instance of every offense can or ought to be prosecuted, prosecutors must exercise discretion in selecting whom to charge for what. Because they exercise such discretion abundantly, the offense of which a defendant formally stands charged may give a false impression of why the state has targeted him for prosecution. In many instances, the formal charge simply is the most substantial crime the state thinks it can prove. The state's true reason for targeting the defendant is that it suspects him of a more serious offense.⁹ This prosecutorial model—classic *pretextual* prosecution—represents one, but only one, scenario in which the state's reason for prosecuting a defendant is something other than a desire to see him brought to justice on the formal charge.

A second is where the state's objective is to incapacitate the defendant by hook or by crook, not because it suspects him of a crime it can't prove but because it thinks he's extremely dangerous. This prosecutorial model is a mainstay of federal criminal prosecution, receiving official sanction not only from the culture¹⁰ and policy¹¹ of the Department of Justice but also from Congress, through statutes like the Armed Career Criminal Act.¹² This statute mandates a prison sentence of at least fifteen years for anyone caught possessing a gun after sustaining three convictions for a violent felony or a serious drug offense.¹³ In promulgating the Armed Career Criminal Act, Congress made no claim that fifteen years in prison for gun possession was a punishment warranted by considerations of desert alone. The avowed

⁹ See generally Harry Litman, *Pretextual Prosecution*, 92 GEO. L.J. 1135 (2004); Daniel C. Richman & William J. Stuntz, *Al Capone's Revenge: An Essay on the Political Economy of Pretextual Prosecution*, 105 COLUM. L. REV. 583, 639 (2005).

¹⁰ See Litman, *supra* note 9, at 1135–36 (“Federal prosecutors take for granted that it is appropriate to prosecute those offenders whose cases will provide important collateral benefits, such as the incapacitation of a repeat violent offender.”).

¹¹ Consider the Department of Justice's Project Safe Neighborhoods and its predecessor, Project Triggerlock, both of which aim for the longest possible sentences for certain criminals with guns. See Bonita R. Gardner, *Separate and Unequal: Federal Tough-on-Guns Program Targets Minority Communities for Selective Enforcement*, 12 MICH. J. RACE & L. 305, 307 (2007).

¹² Armed Career Criminal Act, 18 U.S.C. § 924(e) (2018).

¹³ *Id.*

Congressional purpose was to give prosecutors a tool by which to incapacitate repeat offenders.¹⁴

A third situation in which the formal charge against a defendant may give a false impression of the state's reason for prosecuting him is where the state's true motivation is antipathy to the defendant's beliefs or desires. This scenario often coincides in practice with the second, because a common reason for the state's antipathy to a defendant's beliefs and desires is the perception that these mental states bespeak dangerousness.

Like the first model, the second and third prosecutorial models are in a sense pretextual, the formal charge serving as a pretext for a law enforcement maneuver carried out on an ulterior ground. All three models accordingly raise questions about what the defendant is really being punished for. Is it for the charged offense? Or is it instead for the ulterior ground—for the more serious crime of which the defendant is suspected, for the defendant's perceived dangerousness, or for the defendant's thoughts? I'll focus on pretextual prosecutions grounded in a defendant's thoughts,¹⁵ mindful of the fact that similar concerns could be raised about *any* ulterior ground for prosecution.

If the state prosecutes a defendant for spitting on the sidewalk but wouldn't have prosecuted him at all if it hadn't disliked his thoughts, are these thoughts what the defendant is really being punished for? Questions like this one arise often in connection with the fight against terrorism, where the suspicion that a defendant has been targeted primarily because of his thoughts may stem from multiple sources—including the government's official policy of prosecuting terrorists pretextually,¹⁶ its apparent tendency to view extremist religious speech as a proxy for terroristic

¹⁴ See H.R. REP. NO. 98-1073, at 1 (1984) (stating that the purpose of the Armed Career Criminal Act is “to increase the participation of the [f]ederal [l]aw enforcement system in efforts to curb armed, habitual (career) criminals.”). See also James G. Levine, *The Armed Career Criminal Act and the U.S. Sentencing Guidelines: Moving Toward Consistency*, 46 HARV. J. ON LEGIS. 537, 545–48 (2009).

¹⁵ In what follows, I assume that it's wrong to punish a person for his thoughts, although I don't think anyone has provided a fully satisfactory account of why that is. See Gabriel S. Mendlow, *Why Is It Wrong To Punish Thought?*, 127 YALE L.J. 2342 (2018).

¹⁶ See Attorney General John Ashcroft, Prepared Remarks for the US Mayors Conference (Oct. 25, 2001), https://www.justice.gov/archive/ag/speeches/2001/agerisisremarks10_25.htm (“Robert Kennedy's Justice Department, it is said, would arrest mobsters for ‘spitting on the sidewalk’ if it would help in the battle against organized crime. It has been and will be the policy of this Department of Justice to use the same aggressive arrest and detention tactics in the war on terror.”); Jeff Breinholt, *Seeking Synchronicity: Thoughts on the Role of Domestic Law Enforcement in Counterterrorism*, 21 AM. U. INT'L L. REV. 157 (2005) (defending pretextual prosecution as a method of incapacitating terrorists).

intent,¹⁷ and its apparent willingness to launch terrorism investigations seemingly on the basis of subjects' religion alone.¹⁸

The suspicion that the government has targeted a defendant because of his thoughts can linger even after the government has proved that the defendant engaged in genuinely disturbing conduct. A case in point is the prosecution of Hamid Hayat, an American citizen of Pakistani descent, who stood trial in April 2006 on charges of providing material support to terrorists.¹⁹ The government accused Hayat of traveling from his home in California to a jihadi training camp in Pakistan and then coming back to California about a month later with the intention of committing an act of terrorism.²⁰ When interrogated by the FBI, Hayat at first denied any connection to terrorism. After many hours of questioning, he conceded in the vaguest of terms that he might have been awaiting orders to launch an attack.²¹

¹⁷ See Aziz Z. Huq, *The Signaling Function of Religious Speech in Domestic Counterterrorism*, 89 TEX. L. REV. 833, 845 (2011) (discussing government's treatment of religious speech as a "signal of criminal intent").

¹⁸ Matt Apuzzo & Joseph Goldstein, *New York Drops Unit That Spied On Muslims*, N.Y. TIMES, Apr. 15, 2014, at A1 (describing New York Police Department's decision to "abandon[] a secretive program [called the Demographics Unit] that dispatched plainclothes detectives into Muslim neighborhoods to eavesdrop on conversations and built detailed files on where people ate, prayed and shopped . . .").

¹⁹ Rone Tempest, *In Lodi Terror Case, Intent Was the Clincher*, L.A. TIMES, May 1, 2006, at B1.

²⁰ See Second Superseding Indictment, *United States v. Hayat*, No. 2:05-CR-00240-GEB, at 2–4 (E.D. Cal. Jan. 26, 2006).

²¹ From two news reports, Professor Robert Chesney compiled the following exchange between Hamid Hayat and an FBI interrogator trying to determine whether Hayat was involved in a terrorist plot:

FBI: So jihad means that you fight and you assault something?

Hamid: Uh-huh.

FBI: Give me an example of a target. A building?

Hamid: I'll say no buildings. I'll say people.

FBI: OK, people. Yeah. Fair enough. People in buildings . . . I'm trying to get details about plans over here.

Hamid: They didn't give us no plans.

FBI: Did they give you money?

Hamid: No money.

FBI: Guns?

Hamid: No.

FBI: Targets in the U.S.?

Hamid: You mean like buildings?

FBI: Yeah, buildings. . . . Sacramento or San Francisco?

Hamid: I'll say Los Angeles and San Francisco.

FBI: Financial, commercial?

Hamid: I'll say finance and things like that.

FBI: Hospitals?

Hamid: Maybe, I'm sure. Stores.

Further investigation uncovered no evidence of any actual terror plot, but it did reveal facts that shed an eerie light on Hayat's mind: he admired the murder and beheading of Israeli-American journalist Daniel Pearl,²² he kept a scrapbook of articles about extremist Pakistani political parties, and he carried an Arabic prayer in his wallet that said, "Oh Allah, we place you at their throats, and we seek refuge in you from their evil."²³ The prosecutor emphasized these facts at Hayat's trial, urging the jury not to forget Hayat's "jihadi heart" and "jihadi mind."²⁴

After the jury found Hayat guilty, a reporter asked the U.S. attorney a pointed question: had the case against Hayat partaken of the methods used by the thought police in Philip K. Dick's short story "The Minority Report," where the police indefinitely detained would-be criminals based solely on information about their unexecuted criminal intentions?²⁵ The U.S. attorney rejected the comparison, saying that the government had prosecuted Hayat not for "what he thought" but for what he did, namely, "the overt physical act of attending a training camp and returning to commit jihad."²⁶ This answer did little to assuage the critics of Hayat's prosecution. Suspecting that the government had gone after Hayat more for his "jihadi heart" and "jihadi mind" than for his ambiguous actions,²⁷ critics vehemently accused the

FBI: What kind of stores?

Hamid: Food stores."

Robert M. Chesney, *Beyond Conspiracy? Anticipatory Prosecution and the Challenge of Unaffiliated Terrorism*, 80 S. CALIF. L. REV. 425, 489 (2007) (compiling the exchange from two sources: Mark Arax, *The Agent Who Might Have Saved Hamid Hayat*, L.A. TIMES MAG., May 28, 2006, and Tempest, *supra* note 19).

²² To an FBI informant whom Hayat thought was his best friend, Hayat said, "They killed him [Daniel Pearl]. So, I'm pleased about that. They cut him into pieces and sent him back. That was a good job they did. Now they can't send one Jewish person to Pakistan." Arax, *supra* note 21, at 16.

²³ Tempest, *supra* note 19.

²⁴ *Id.*

²⁵ *Id.* See Philip K. Dick, *The Minority Report*, in THE MINORITY REPORT AND OTHER CLASSIC STORIES 71 (2002).

²⁶ Tempest, *supra* note 19.

²⁷ See *Punished for a Thought: Law Perverts Justice with Charges for Acts*, DAYTONA BEACH NEWS-JOURNAL, Aug. 2, 2007, at 04A ("The government is arresting individuals on terrorism charges based on what individuals have said or thought—not on actual, concrete plans."); Glenn Greenwald, *The FBI's Anticipatory Prosecution of Muslims to Criminalize Speech*, THE GUARDIAN, Mar. 19, 2013, (quoting Professor Shirin Simar as saying that Hayat's case "rais[es] the haunting prospect that a man who had done nothing was convicted for a violent state of mind."); Eric Umansky, *Department of Pre-Crime: Why Are Citizens Being Locked Up for "Un-American" Thoughts?*, MOTHER JONES, Feb. 29, 2008 (asserting that Hayat effectively was charged with having a "jihadi mind"); Cf. Peter W. Beauchamp, *Misinterpreted Justice: Problems with the Use of Islamic Legal Experts in U.S. Trial Courts*, 55 N.Y.L. SCH. L. REV. 1097, 1106 (2011) (implying that defendants like Hamid Hayat are "prosecut[ed] . . . on the basis of what they or their support may someday accomplish, rather than for what has already been done.").

government of punishing Hayat for thinking bad thoughts²⁸ and for having a “violent state of mind.”²⁹

Because Hayat nevertheless had performed actions that a reasonable observer could characterize as furthering the cause of jihad, the government plausibly could claim that its motivation really was to bring Hayat to justice for the crime of which he was formally charged—placing himself at the disposal of a terrorist organization³⁰—rather than to punish Hayat for his beliefs and desires. With other terrorism prosecutions, however, we may have a harder time excluding the possibility that the government has gone after the defendant primarily because of his beliefs and desires. Often the charged offense bears no connection to terrorism.³¹ As the Chief of the Justice Department’s Criminal Division explained in Congressional testimony,

[t]here are a number of terrorism investigations where the decision . . . made at the charging stage [is] to charge the defendant with a non-terrorism crime in order to protect . . . national security and classified information that may be exposed, sources and methods and that sort of thing, that may be jeopardized by the criminal discovery that would ensue if we were to charge [the defendant with a] terrorism offense.³²

As a result of charging decisions made in secret, sometimes no one but the prosecutor knows that the government is going after the defendant because he’s suspected of being a terrorist or terrorist sympathizer—not the jury, not the judge, not even the defendant himself.³³

²⁸ Umansky, *supra* note 27.

²⁹ Greenwald, *supra* note 27 (quoting Professor Shirin Sinnar).

³⁰ See *United States v. Hayat*, 710 F.3d 875, 883 (9th Cir. 2013) (“The statute [under which Hayat was charged, namely, 18 U.S.C. § 2339A(a) (providing material support to a terrorist organization),] defines ‘material support or resources’ as [inter alia] ‘any . . . personnel (1 or more individuals who may be or include oneself). . . . The prosecution’s case was that Hayat had provided his personal services to a terrorist organization by attending the training camps in Pakistan and returning with the intent to carry out acts of terrorism when directed to do so.”).

³¹ Todd Lochner, *Sound and Fury: Pretextual Prosecution and Department of Justice Antiterrorism Efforts*, 30 *LAW & POL’Y* 168, 185 (2008) (claiming that “the crimes most commonly associated with pretextual antiterrorism prosecutions [are] immigration, false statements, fraud, and counterfeiting violations . . .”).

³² Daniel C. Richman & William J. Stuntz, *supra* note 9, at 621 (quoting Criminal Terrorism Investigations and Prosecutions: Hearing Before the S. Comm. on the Judiciary, Oct. 21, 2003, (testimony of Christopher Wray), LEXIS, Federal News Service Transcript).

³³ In 2008, I heard federal judge Kenneth M. Karas (formerly an anti-terrorism prosecutor in the United States Attorney’s Office for the Southern District of New York) tell an audience at Yale Law School that his old office routinely prosecuted suspected Hezbollah operatives for low-level infractions in order to procure their deportation—these defendants being oblivious to the fact that the

Utterly in the dark about the number and nature of these cases, we can only guess how many terrorism prosecutions are motivated by the government's antipathy to—or, more exactly, its fear of—the defendant's beliefs and desires, rather than by intelligence about incipient terror plots. The number of these prosecutions almost certainly is greater than zero, given the apparent frequency with which the government bases antiterrorism investigations and prosecutions on subjects' religious speech.³⁴ In all probability, the government sometimes prosecutes a defendant at least in part because of his "jihadi mind." Does it follow that the crime of which the defendant is formally charged and convicted isn't really what he's being punished for—that what he's really being punished for is having a "jihadi mind"? (Again, we could ask a similar question about any ulterior ground for prosecution.)

A *formalistic* answer (if true) would resolve all uncertainty about the object of a defendant's punishment and remove any doubt about whether the defendant is being punished improperly. That a defendant is prosecuted and punished *because* of his thoughts, a formalist will insist, does not entail that he is punished *for* his thoughts. Legal punishment is a condemnatory sanction imposed by the state upon a supposed offender for an offense,³⁵ and that offense (the formalist will say) is none other than the offense of which the defendant is formally charged and convicted. That offense is the offense the state publicly accuses the defendant of having committed, the offense governing the nature and scope of the evidence that the prosecution presents at the defendant's trial, the offense that the jury must conclude the defendant committed before the state may punish him, the offense determining the range of possible sentences, and the offense for which the court publicly claims to be punishing the defendant when it sentences him. The formalist will insist that it is *these* features of a prosecution—not the state's ulterior motivation for initiating it—that determine which wrong the defendant is really being punished for.

government was going after them because it suspected they were terrorists. *See also* Richman & Stuntz, *supra* note 9, at 639; Lochner, *supra* note 31, at 168 ("One federal district court judge [in Iowa] claimed that 'if there have been terrorism-related arrests in Iowa, I haven't heard about them' even though he presided over six of the criminal cases that the Department [of Justice] had catalogued as terrorism convictions.").

³⁴ *See* Huq, *supra* note 17, at 845 ("The incidence of religious speech at the prosecution stage as a signal of criminal intent is suggestive of a greater reliance on the same kind of evidence upstream—in investigations. Even setting aside those investigations that do not end in charges, many terrorism investigations (perhaps a majority) end in 'pretextual' charges, from wire fraud to immigration crimes. Such charges are unconnected with terrorism but form the possible basis for a less costly type of punitive action. In those cases, the state's upstream reliance on religious speech for singling out a suspect is never revealed. At the very least, therefore, any estimation of the use of religious speech as a signal in counterterrorism that relies on reported trials is likely to yield an undercount, and probably a substantial one."); Apuzzo & Goldstein, *supra* note 18.

³⁵ *See* H. L. A. HART, *Prolegomenon to the Principles of Punishment*, in PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 1, 4–5 (2d ed. 2008).

If the formalist is right, then even Al Capone, the most famous subject of pretextual prosecution in American history, really was punished for the crime of which he was formally charged and convicted—tax evasion—rather than for the more serious crimes that the government openly ascribed to him and that lay behind its decision to prosecute him in the first place. Tax evasion was the crime of which Capone was publicly accused, the crime of which the government presented evidence at his trial, the crime of which the jury found him guilty, and the crime for which the court claimed to be punishing him when it sentenced him to eleven years in prison.³⁶

If the formalist is correct—if Capone was punished for tax evasion and not, say, for the Saint Valentine’s Day Massacre—then Hamid Hayat was punished for providing material support to terrorists and not for his jihadi heart and mind; indeed, all defendants are punished for the crimes of which they’re formally charged and convicted and not for the activities or characteristics that might have kindled the government’s interest in prosecuting them.

If these claims come off as unsophisticated or naïve, the formalist might defend them by insisting there is no viable alternative. What is an offender really punished for, if not for the offense of conviction? For whatever made the prosecutor want to prosecute him? Why not for whatever made the police want to arrest him? Or whatever made the jury want to convict him? Or whatever made the judge want to sentence him? And when the motives of these actors diverge, as they frequently do, whose motive determines what the offender is really being punished for? How often can we know the true motives of these actors anyway? If only rarely—so the formalist will continue in a mocking tone—then only rarely can we know what an offender is really being punished for, and only rarely can we determine whether an offender’s sentence is truly proportionate to his offense, or whether that offense is something the state is permitted to punish in the first place.

To follow the formalist this far, we must accept that the motivations of the participants in the criminal process play no role whatever in fixing the object of an offender’s punishment. Should we accept this? The formalist’s rhetorical questions suggest that, if the motivations of the participants determined (or helped determine) what an offender is being punished for, then our ability to know whether an offender is being punished justly could be as limited as our ability to see inside the participants’ heads, which is to say, very limited indeed. But this isn’t a demonstration that formalism is true; it’s a description of the widespread uncertainty we’d face if formalism were false. As I’ll presently explain, we may have good reason to conclude that it is.

³⁶ LAURENCE BERGREEN, *CAPONE: THE MAN AND THE ERA* 440–87 (1994) (describing prosecution of Al Capone).

The root of the problem is that legal punishment is an act of collective expression³⁷ and its object is therefore a function of its social meaning. Just as a person can purport to express disapproval of one thing (by saying certain words) while actually expressing disapproval of another (with a wink and a nod, or a smirk and a shrug), so might the state purport to disapprove of one wrong while actually disapproving of another. The Soviet show trials of the 1930s seem a clear example of the phenomenon, perhaps more so than the prosecution of Al Capone. When the motivations of the participants diverge grossly from the participants' formal account of what they're up to, it's hard to deny that the participants are up to something other than what they say they are.

How gross must the divergence be in order for us to conclude that the defendant in a criminal case is being punished for something other than the nominal offense of conviction? If we don't know where to draw the line, then we ought to consider that we often may not know what offenders are being punished for. And if we often don't know what offenders are being punished for, then we don't know how often they're being punished disproportionately, how often they're being punished for an offense the law may not punish, and how often they're being punished in contravention of due process.

IV.

In a moment, I'll elucidate and deepen these uncertainties using principles of philosophy of language. But first I want to show that similar uncertainties—perhaps greater ones—arise from a court's discretion to sentence.

Because our criminal codes are coarse-grained, not every instance of a given offense merits the same punishment. Some burglaries, assaults, and kidnappings are worse than others. That's why the law authorizes a court to exercise discretion when it fashions a sentence.³⁸ To exercise its discretion rationally, a court must consider all relevant evidence—and the range of potentially relevant evidence is vast. It includes at least the following: evidence of details about the crime besides those needed to sustain a conviction; evidence of actions the defendant performed in preparation for the crime; evidence of actions the defendant performed to avoid detection afterward; evidence of related but distinct transgressions; and, of course,

³⁷ JOEL FEINBERG, *The Expressive Function of Punishment*, in *DOING AND DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY* 95, 98 (1970) (“[P]unishment is a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation, on the part either of the punishing authority himself or of those ‘in whose name’ the punishment is inflicted.”).

³⁸ See John Gardner, *Wrongs and Faults*, 59 *REV. METAPHYSICS* 95, 122 (2005) (“The Rule of Law, by insisting on clarity, stability, prospectivity, and so on, often prevents the law from showing full sensitivity to the differential moral merits of every wrong that is committed and every credible justification or excuse for its commission. Some questions bearing on fault, normally but not always relating to the fine-tuning of fault. . . are inevitably left over to be dealt with at the sentencing stage.”).

evidence of the defendant's prior misconduct.³⁹ The Federal Sentencing Guidelines direct courts to sentence defendants on the basis of all "relevant conduct,"⁴⁰ a category encompassing a wide range of uncharged misdeeds,⁴¹ even including misdeeds of which the defendant has been acquitted.⁴² Most state judges enjoy comparable discretion, sentencing offenders based on their uncharged offenses, history, and other personal characteristics.⁴³

Now it's often entirely appropriate for a judge to consider these facts at a defendant's sentencing hearing. As Julie O'Sullivan observes, "[i]nformation about the true scope of the defendant's and his accomplices' related criminal activity informs our assessment of the defendant's just deserts for the offense of conviction in so far as it illuminates the defendant's motivation and purposefulness in engaging in this criminal act."⁴⁴ Because a defendant's "motivation and purposefulness" may make his offense of conviction a *worse* offense—one meriting harsher punishment—the fact that a defendant's motivation increases his sentence doesn't necessarily mean that what he's really being punished for is his motivation in itself, a mental transgression distinct from the offense of conviction.

³⁹ See, e.g., *Williams v. New York*, 337 U.S. 241, 244–45 (1949) ("Within limits fixed by statutes, New York judges are given a broad discretion to decide the type and extent of punishment for convicted defendants. . . . To aid a judge in exercising this discretion intelligently the New York procedural policy encourages him to consider information about the convicted person's past life, health, habits, conduct, and mental and moral propensities.").

⁴⁰ U.S. SENTENCING GUIDELINES MANUAL § 1B1.3 (U.S. SENTENCING COMM'N 2015) (Relevant Conduct). See *United States v. Galloway*, 976 F.2d 414, 419 (8th Cir. 1992) (en banc) ("In promulgating the relevant conduct guideline [§ 1B1.3], the Commission adopted a so-called 'real offense' philosophy, since section 1B1.3 calls for sentencing based upon the actual conduct engaged in, rather than only upon the offense for which the defendant was convicted. . . . To this end, subsection (a)(2) of the relevant conduct guideline requires courts, in determining an offender's base offense level, to include 'all such acts and omissions that were part of the same course of conduct or common scheme or plan as the offense of conviction.'" (quoting U.S. SENTENCING GUIDELINES MANUAL § 1B1.3(a)(2))).

⁴¹ See, e.g., *Galloway*, 976 F.2d at 414 (upholding defendant's sentence on one count of theft from interstate shipment of goods where sentence was based on eight separate thefts, including seven that were not charged).

⁴² *United States v. Watts*, 519 U.S. 148, 157 (1997) (holding that "a jury's verdict of acquittal does not prevent the sentencing court from considering conduct underlying the acquitted charge, so long as that conduct has been proved by a preponderance of the evidence.").

⁴³ See Kevin R. Reitz, *Sentencing Facts: Travesties of Real-Offense Sentencing*, 45 STAN. L. REV. 523, 528 (1993) ("Nearly every state allows sentencing courts to engage freely in real-offense sentencing as a matter of discretion"); *id.* at 528 n.22 (collecting cases in support of same); *id.* at 534 ("Nearly all jurisdictions allow courts to consider . . . nonconviction charges when sentencing."); *id.* at 534 n.70 (collecting cases in support of same).

⁴⁴ Julie R. O'Sullivan, *In Defense of the U.S. Sentencing Guidelines' Modified Real-Offense System*, 91 NW. U. L. REV. 1342, 1370 (1997).

Still, whenever a sentencing court considers non-offense evidence, it's natural for us to wonder whether the court really is punishing the defendant just for the offense of conviction—as the law requires—or is punishing the offender instead (or in addition) for the other transgressions that the non-offense evidence may reveal. This is a question we need to be able to answer: if the true object of an offender's punishment is something other than the offense of conviction, then the offender is being punished in contravention of a supposed precept of our legal system: that no one may be punished for an offense unless charged, tried, and proved guilty of *that* offense.⁴⁵

A striking illustration of this possibility is the case of Dennis Hastert, a former speaker of the U.S. House of Representatives convicted of “currency structuring.” (The crime of currency structuring involves breaking up a banking transaction into amounts less than \$10,000 in order to avoid triggering the bank's duty to file a report.) Investigators examining Hastert's suspicious banking practices discovered not only that he'd been engaging in currency structuring but that he'd been doing it with a disturbing motive: to conceal the fact that he was buying the silence of a man he'd sexually abused as a boy decades earlier. Investigators eventually learned that Hastert had abused at least three other boys while serving as their high school wrestling coach. All of this abuse factored into Hastert's ultimate sentence for currency structuring. The sentencing judge offered this comment:

Had this [sexual abuse] been uncovered near the time when it occurred, a grand jury . . . would have indicted you, a jury likely would have convicted you, and you likely would have been sentenced to decades in a state prison. . . . But because the statute of limitations for your child molestation ran out many years ago, you can't be charged for that. And I can't sentence you as a child molester. It's not what you were charged with, it's not what you've pled guilty to, and any sentence I give you today will pale in comparison to what you would have faced in state court. But this conduct is relevant to your history and characteristics no matter how old it is. Some conduct is unforgivable no matter how old it is. If the juvenile victim of sex abuse can't forget decades later what happened, then neither can I as a judge nor can we as a society. The abuse was 40 years ago, but the damage lasts today. . . . My sentence today can't legally or properly be a sentence for child molestation, and I don't want it in any way to be perceived that the sentence here measures the harm caused by the child molestation. In the end, that would have to be a state court judge

⁴⁵ A large part of what gives point and urgency to the question of whether “repeat offenders [are] more culpable than first-time offenders” is our need to assure ourselves that recidivist sentencing practices don't amount to punishing recidivists more than once for the same offense. Youngjae Lee, *Repeat Offenders and the Question of Desert*, in *PREVIOUS CONVICTIONS AT SENTENCING: THEORETICAL AND APPLIED PERSPECTIVES* 49 (Julian V. Roberts & Andrew von Hirsch eds., 2010).

sentencing you for a conviction of child molestation, and the sentence in this case can never be as long as the time the victims and their families have suffered.⁴⁶

Although the judge acknowledged that Hastert's sentence "[couldn't] legally or properly be a sentence for child molestation," Hastert's child molestation was the focal point of the sentencing hearing. The judge questioned Hastert about each of the four men he'd abused as boys, rather than just about the man whose victimization Hastert had tried to conceal. The judge allowed testimony from one of Hastert's victims and also from the sister of another victim who had died of AIDS two decades earlier. The judge then imposed a sentence of fifteen months—two-and-a-half times the maximum recommended by the Federal Sentencing Guidelines. Commenting on the sentence, a former federal prosecutor said, "It's extraordinary that the case was on its face a cut-and-dry financial structuring case . . . but the sentencing [hearing] was about everything, essentially, but the structuring."⁴⁷

None of this was improper *per se*. Not every fact that increases an offender's sentence necessarily forms part of the object of the offender's punishment: to punish Hastert in the light of his child molestation wasn't necessarily to punish him *for* it. If Hastert's currency structuring truly warranted a sentence two-and-a-half times the recommended maximum because of his illicit motive, then the judge was right to consider the child molestation at the sentencing hearing.

But the hearing itself told a different story. Much of the child molestation recounted at the hearing wasn't molestation that Hastert had intended his currency structuring to conceal. The judge evidently believed that if he *said* he was punishing Hastert for currency structuring and not for child abuse, then that's really what he was doing. But why should we accept what the judge said? Why should we accept that a judge is doing whatever he says he's doing, and not anything else (or anything more)?

Generally, we don't believe that an entity's self-description determines the proper characterization of its conduct. Suppose Country A bans the import of meat from Country B after Country B launches a cyber-attack on Country A. Country A claims that the import ban is motivated by a concern for food safety, when in fact it's motivated by a desire to retaliate against Country B for the cyber-attack. Does it follow that the import ban isn't in fact an act of retaliation? Is the import ban a food-safety measure simply because that's what Country A calls it? If that were true, clandestine retaliation would be a conceptual impossibility. Often more important than what a person or entity says they're doing is what they *think* they're doing. If Country A thinks it's engaged in retaliation—if it enacts an import-ban for

⁴⁶ Transcript of Sentencing Hearing at 67–68, *United States v. Hastert*, (No. 15 CR 315) (N.D. Ill. Apr. 27, 2016).

⁴⁷ Monica Davey, Julie Bosman & Mitch Smith, *Dennis Hastert Sentenced to 15 Months, and Apologizes for Sex Abuse*, N.Y. TIMES, Apr. 27, 2016.

the purpose of getting back at Country B—that seems sufficient to make the import ban an act of retaliation.

Why accept that punishment is different from geopolitical retaliation? Why accept that, unlike an act of geopolitical retaliation, an act of punishment is the precise act that it is (i.e., an act of punishment for x rather than an act of punishment for y) simply because that's what the punisher *says*? We lack any argument, let alone a conclusive one, for this sort of formalistic, motivation-independent account of the object of punishment.

For my part, I find it hard to imagine that the object of punishment isn't fixed at least some of the time by the motivations of the relevant legal actors. I don't claim that a prosecutor or judge's motivations necessarily play a determinative role in fixing the object of punishment. I claim merely that officials' motivations probably play at least a partial role. It's consistent with this claim that Al Capone really was punished for tax evasion. My point is that the question whether Capone was or wasn't punished for tax evasion is an *open* question. And that's because the object of an offender's punishment is in part a function of his prosecution's social meaning, and that social meaning is in part a function of the participants' motivations, rather than solely a function of the prosecution's formal features. That the formal features all point in the same direction isn't enough to render the question closed. If the participants in Hastert's case had each believed that they were censuring him for his child molestation—if they'd had the child molestation foremost in their minds, if they'd aimed to craft a punishment befitting it, and if they'd hoped he'd feel the sting of his sentence as a just rebuke for what they saw as his true misdeed—I don't see how we could maintain with any confidence that Hastert was punished for currency structuring and not for child molestation.

But if the participants' beliefs and motivations help determine the object of punishment in the not-so-extreme hypothetical circumstances I just described, then why not in more moderate circumstances, too? The question, it seems, isn't *whether* participants' mental states can determine the object of punishment—it's *when*. Legal punishment is a kind of collective expression, a kind of collective utterance. As with all utterances, its meaning seems unlikely to be entirely independent of the beliefs and intentions of the speaker.

Our (potential) ignorance about these matters is on two levels. One level is practical: we may not know the specific facts (e.g., the specific collective intentions) that determine the object of a particular offender's punishment. The other is theoretical: we may not know the more general *type* of facts that determine the object of any offender's punishment. In other words, we might not know what a given offender is punished for because we don't even know where to look for an answer. We lack a consensus theory of the object of punishment. Indeed, we lack any theory at all. The only thing beyond reasonable dispute is that punishment essentially involves the censure of an offender for a supposed wrong.

We should start by asking what (or who) is the punitive agent, the agent that does the censuring. I've referred to it as the state, but the state is an abstraction, not a concrete actor. The state acts only through its officials. Is the punitive agent

therefore some particular official—the sentencing judge, perhaps, or the prosecutor, a more central figure in our system? Or is the punitive agent a collective agent—the criminal legal system as a whole? The question is which agent’s mental state we must ascertain in order to identify the object of punishment. We simply don’t know. Nor do we know which types of mental state play the determinative role.

We won’t be in a position to know until we’ve resolved a further issue: how should we understand the act of censuring? Is the act of censuring the *communication* of censure to some intended audience, or is it instead the mere *expression* of censure? “To *express* a mental state requires only that one manifest it in speech or action,” explain Elizabeth Anderson and Richard Pildes.⁴⁸ “To *communicate* a mental state [by contrast] requires that one express it with the intent that others recognize that state by recognizing that very communicative intention.”⁴⁹

If the censure in punishment is expressive rather than communicative,⁵⁰ the punitive agent can censure an offender for a given wrong without ever meaning to communicate that censure. As Anderson and Pildes observe, “[o]ne can express a mental state without intending to communicate it. The shoplifter may express her intention to get away with stealing a purse in her furtive glances. But she hardly intends to communicate this intention.”⁵¹ Likewise, the punitive agent may express its disapproval of an offender’s extraneous (i.e., non-offense) misconduct through its guiding preoccupation with that misconduct, whether or not it intends to express its disapproval, and even if it intends not to. If the censure in punishment is purely expressive, then what matters is the object of the reprobative attitude that the punitive agent’s preoccupation in fact expresses, not the attitude (if any) that the agent intends to express. This is true of expression generally. In complimenting the appearance of a female co-worker, a businessman may not intend to express disrespect—he may intend the opposite—but his compliment will express disrespect nonetheless if it evinces the belief that women in a business setting are, as Anderson and Pildes put it, “sexual or aesthetic adornments.”⁵² In cases of mere expression, the mental state that determines an act’s expressive content may be a mental state of which the agent is unaware, or even a mental state that the agent intends to conceal.

If by contrast the censure in punishment is communicative rather than merely expressive, then the punitive agent’s reprobative attitude isn’t what determines the object of censure. What determines the object of censure is instead either the

⁴⁸ Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1508 (2000) (emphasis added).

⁴⁹ *Id.* (emphasis added).

⁵⁰ Feinberg defines punishment as “a conventional device for the expression of attitudes of resentment and indignation, and of judgments of disapproval and reprobation.” Feinberg, *supra* note 37, at 98.

⁵¹ Anderson & Pildes, *supra* note 48, at 1508.

⁵² *Id.* at 1525.

punitive agent’s “communicative intention” (roughly, what the speaker intends his audience to understand by a given utterance), or the audience’s reasonable understanding of the speaker’s communicative intention—or perhaps some combination of the two.⁵³ Seemingly all theorists agree that the content of what a speaker communicates by an utterance is never just a matter of the words he uses,⁵⁴ just as the content of what an actor expresses through an action isn’t a matter of what he *says* while doing it. No one is a formalist about the communicative content of assertive utterances, as no one is a formalist about acts’ expressive meanings. Nor should anyone be a formalist about the communicative or expressive content of criminal prosecutions. What legal officials say when they prosecute someone doesn’t determine what they communicate or express by prosecuting him.

While there’s no consensus theory of the nature of the punitive agent or of the nature of the act of censure, all extant theories of expression or communication indicate uncertainty about the object of punishment. All such theories entail that the content of the act of censure hinges on more than its formal features—on more than the punitive agent’s explicit self-description. And nearly every theory of expression or communication also entails that the precise meaning of the act of censure hinges at least partly on the punitive agent’s mental states. I can imagine only one

⁵³ See Mark Greenberg, *Legislation as Communication? Legal Interpretation and the Study of Linguistic Communication*, in *PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW* 217, 230–31 (Andrei Marmor & Scott Soames eds., 2011) (“There are different notions of communicative content. One notion of communicative content—the *neo-Gricean* one, as I will call it—derives from the seminal work of Paul Grice. According to the neo-Gricean notion, the communicative content of an utterance is the content of a certain kind of complex intention [that I’ll call a *communicative intention*]. Roughly speaking, for a speaker’s utterance of a sentence to have the communicative content that P is for the speaker to utter the sentence intending his or her hearers to come to recognize that the speaker is communicating P, in part by their recognition of this very intention. . . . By contrast, according to what I will call an *objective* notion of communicative content, the communicative content of an utterance is what a member of the audience would reasonably take a speaker who had uttered the relevant sentence under specified conditions to have intended to communicate. . . . There are also hybrid notions. For example, according to one notion, the communicative content of an utterance is that part of what the speaker intended to communicate for which uptake by the audience could reasonably be expected.”).

⁵⁴ Noting the “now widely recognized fact that what speakers mean generally goes beyond sentence meaning,” Kent Bach observes that “it is now a platitude that linguistic meaning generally underdetermines speaker meaning. That is, generally what a speaker means in uttering a sentence, even if the sentence is devoid of ambiguity, vagueness, or indexicality, goes beyond what the sentence means.” Kent Bach, *Context ex Machina*, in *SEMANTICS VERSUS PRAGMATICS* 15–16 (Zoltan Gendler Szabo ed., 2015). See also *id.* at 26–27 (“[T]he semantic content of a sentence . . . always underdetermines what a speaker means in uttering it. Here’s why. Even if what a speaker means consists precisely in the semantic content of the sentence he utters and that content is precise, this fact is not determined by the semantic content of the sentence. The reason for this claim is very simple: no sentence has to be used in accordance with its semantic content. Any sentence can be used in a nonliteral or indirect way. A speaker can always mean something distinct from the semantic content of the sentence he is uttering. That he is attempting to communicate something, and what that is, is a matter of his communicative intention, if indeed he has one. If he is speaking literally and means precisely what his words mean, even that fact depends on his communicative intention.”).

circumstance in which the punitive agent's mental states play no role in fixing the object of punishment. That's the circumstance in which no legal actor exercises any discretion about which individuals to charge, what crimes to charge them with, and what sentences to impose. This circumstance exists only under the Code of Babel.⁵⁵

V.

I've argued that our uncertainty about the object of punishment stems from the fact that legal punishment is an act of collective expression. But if legal punishment is an act of collective expression, it's one that's drawn out and disjointed, with no member of the collective clearly authorized to speak on its behalf. This complexity might lead us to suspect that legal punishment *qua* expressive act isn't really a single act at all, but is instead a series of expressive acts that, taken together, lack a coherent meaning.⁵⁶ If that's right, then we theoretically could conclude that no one is ever punished for anything. The state imposes punishment only when it expresses condemnation of a particular person for a particular wrong. But if the state is too disjointed a collection of entities ever to function as a single collective agent expressing a coherent message of condemnation, then it never expresses condemnation of particular persons for particular wrongs: it never imposes punishment.

That's an absurd conclusion. So we should consider rejecting its unstated premise: that, in order to impose punishment, the state must function as a single collective agent expressing a coherent message of condemnation. The state in a criminal prosecution is perhaps not one agent but a series of them, some individual, some collective: the police officer who makes the arrest, the prosecutor and her advisors and supervisors, the jury (if any), and the judge. From this conception of the state *qua* punisher we might infer that a given prosecution involves two distinct

⁵⁵ Against all this, one might assert the unorthodox claim that a punishment's meaning is entirely a function not of the tribunal's intent but of the actual or expected understanding of that intent. The claim is unorthodox because philosophers of language widely agree that an expressive act's meaning is at least sometimes a function of what's going on in the speaker's head. But what if collective utterances were different? What if the meaning of a collective utterance, unlike that of an individual one, were a function entirely of the audience's understanding? It still would remain to be determined how broadly we must define the relevant audience. It's not obvious that the relevant audience would be the public at large. If the relevant audience were instead the set of officials connected in some fashion to the collective act, then the act's meaning might remain elusive. But what if the relevant audience were indeed the public at large? In that case, the object of an offender's punishment still might be something other than the nominal offense of conviction, but any such discrepancy would be public rather than hidden. So we'd still be right to worry about whether our legal system violates the rights it proclaims sacrosanct—and arousing this concern is this essay's primary goal—but we'd be wrong to fear that these violations occur more frequently than we recognize.

⁵⁶ See Ryan D. Doerfler, *Who Cares How Congress Really Works?*, 66 DUKE L.J. 979, 979 (2017) ("Congress has no collective intention, not because of difficulties in aggregating the intentions of individual members, but rather because Congress lacks the sort of delegatory structure that one finds in, for example, a corporation.").

objects of criminal liability: one at the guilt/innocence stage and another at the sentencing stage. If the object of punishment is determined partly by the mental states of the relevant legal officials, then the wrong of which an offender is convicted might differ from the wrong for which he's sentenced.

If we embraced this bifurcation of the object of punishment, we could avoid some of the concerns I raised earlier about how often defendants' procedural rights are given their due—but only if we were willing to understand the rights in question as relativized to different stages of a prosecution. Consider the right not to be punished for an offense unless charged, tried, and proved guilty beyond a reasonable doubt. If we understood this right as applying only at the guilt/innocence stage, we'd conceive it as a right not to be *deemed guilty* of some offense unless charged with and proved guilty of the offense beyond a reasonable doubt. So conceived, the right wouldn't ground a claim not to be *sentenced* for some offense unless charged with and proved guilty of the offense beyond a reasonable doubt. A defendant properly proved guilty of the charged offense would have no cognizable complaint if subsequently punished both for the offense of conviction and for some extraneous transgression proved (only) at the sentencing hearing (only) by a preponderance of the evidence. We could understand the right to a proportionate punishment similarly: not as the right to be punished in proportion to the gravity of the offense of conviction, but (merely) as the right to be punished in proportion to the gravity of the wrongdoing proved at the sentencing stage.

If we divorced “sentencing wrongs” from “guilt/innocence wrongs” in this way, we'd reduce but not eliminate our uncertainty about whether defendants' procedural rights are given their due. We'd no longer have reason to worry about whether the wrong targeted at the sentencing stage was the same as the wrong targeted at the guilt/innocence stage: we'd accept from the outset that these wrongs often differ and that their differing offends no right. But, thanks to prosecutorial discretion, we'd still have reason to worry about whether the wrong targeted at the guilt/innocence stage really was the wrong of which the defendant was charged and for which he was tried. And, thanks to judicial discretion, we'd still have reason to worry about whether the wrong targeted at the sentencing stage diverged from the nominal object of sentencing.

Our biggest concern, however, would be the fact that divorcing “sentencing wrongs” from “guilt/innocence wrongs” would portray the law of criminal sentencing, if not the entire criminal process, as suffering from a kind of false consciousness. On its face, sentencing law strives to ensure that penal sentences punish offenders (only) for the wrongs of which they're convicted: to ensure that “sentencing wrongs” and “guilt/innocence wrongs” are identical. Consider, for example, that criminal codes associate particular offenses with specific sentencing ranges. This feature implies that the guilt/innocence stage of a prosecution isn't an intake valve, as the Divine Justice model supposes, but that the guilt/innocence stage

serves instead to establish which wrong (or which general type of wrong⁵⁷) a given offender may be punished for.

Equally revealing is the line of Supreme Court cases requiring prosecutors to prove at the guilt/innocence stage any fact that raises the maximum sentence or triggers a mandatory minimum sentence.⁵⁸ The first of these cases, *Apprendi v. New Jersey*, concerned a man who fired several shots into the house of his African-American neighbors because he didn't want black people in his neighborhood.⁵⁹ Apprendi pleaded guilty to a gun possession offense carrying a sentence of five to ten years in prison.⁶⁰ The trial court then sentenced Apprendi to *twelve* years, invoking a hate crime law that permitted enhanced sentences for crimes done with a purpose to intimidate an individual or group because of their race.⁶¹ Apprendi appealed his sentence and ultimately persuaded the U.S. Supreme Court that the sentencing enhancement violated his right to a jury trial.⁶² The Court held that the New Jersey hate crime law had punished Apprendi for a crime of which he hadn't been duly convicted:

New Jersey threatened Apprendi with [a] certain [punishment] if he unlawfully possessed a weapon and with additional [punishment] if he selected his victims with a purpose to intimidate them because of their race. As a matter of simple justice, it seems obvious that the procedural

⁵⁷ The guilt/innocence stage of a prosecution serves to identify which type of wrong the defendant subsequently may be punished for, rather identifying which token of that type the defendant committed. That's why the "jury need not always decide unanimously which of several possible sets of underlying brute facts make up a particular element, say, which of several possible means the defendant used to commit an element of the crime." *Richardson v. United States*, 526 U.S. 813, 817 (1999). Indeed, with respect to a crime involving more than one participant, the jury need not even agree on whether the defendant was the principal or an accomplice. *See Rosemond v. United States*, 572 U.S. 65, 69 (2014) (noting without disapproval that "[t]he verdict form was general: It did not reveal whether the jury found that [defendant] himself had used the gun or instead had aided and abetted a confederate's use during the marijuana deal."). Determining pertinent details of the specific *token* wrong the defendant committed is one function of a sentencing hearing. (Thanks to Eric Swanson for help with this point.)

⁵⁸ *Apprendi v. New Jersey*, 530 U.S. 466 (2000) (holding that any fact that increases the penalty for a crime beyond the prescribed statutory maximum, other than the fact of a prior conviction, must be submitted to a jury and proved beyond a reasonable doubt); *Alleyne v. United States*, 570 U.S. 99 (2013) (holding that any fact that increases mandatory minimum sentence for a crime is an "element" of the crime, not a "sentencing factor," and therefore must be submitted to jury and proved beyond a reasonable doubt).

⁵⁹ *Apprendi*, 530 U.S. at 469.

⁶⁰ *Id.* at 469–70.

⁶¹ *Id.* at 470–71.

⁶² *Id.* at 476.

safeguards designed to protect *Apprendi* from unwarranted [punishment] should apply equally to the two acts that New Jersey has singled out for punishment. Merely using the label “sentence enhancement” to describe the latter surely does not provide a principled basis for treating them differently.⁶³

In the Court’s view, New Jersey’s “sentencing enhancement” law employed the nomenclature of a deceptive formalism, one that masked the state’s effort to punish offenders for wrongs it hadn’t proved at the guilt/innocence stage. The Court seemingly took for granted, as a bedrock principle, that the wrong for which a defendant is sentenced must be the wrong of which he was convicted.

In theory, we could turn the tables on the Court’s view by branding its *Apprendi* decision as nothing but a piece of deceptive formalism itself. We could insist that *all* principles of sentencing law exist to mask the fact that our criminal legal system actually operates on the Divine Justice model, the law of criminal liability serving as little more than an intake valve. If that’s our criminal legal system, then the real function of sentencing law is to put a filter over the valve—not for the purpose of ensuring that offenders are punished only for the wrongs of which they’re convicted, but instead for the purpose of bridling the state’s otherwise largely unbridled discretion to act as an arbiter of divine justice. Many imaginable variations on our existing sentencing law could serve this discretion-bridling purpose equally well. If the point is to bridle discretion, what’s important is that there be *some* set of regulatory principles. Their particular content is secondary.

If our criminal legal system covertly operates on the Divine Justice model, then it’s not just sentencing law but criminal law across the board that suffers from a kind of false consciousness. Our law of criminal liability *appears* to play what R.A. Duff calls “a declaratory rather than [merely] prohibitory [role],”⁶⁴ the law of crimes serving “to declare that [certain] pre-legal wrongs are public wrongs: to declare, that is, not merely that they are wrongs . . . but that they are wrongs that properly concern the whole polity, which should call their perpetrators to public account through the criminal courts.”⁶⁵ But if our system covertly operates on the Divine Justice model, then the penal law’s declaratory character is mere pretense. Albeit *useful* pretense: there’s utility in the pretense that the penal law serves to mark out pre-legal wrongs as worthy of collective condemnation. A criminal legal system speaking the language of moral wrongs commands allegiance more cheaply than one relying entirely on threats.

That could be the system we have. But I doubt it’s the system most legal officials think we have, and that’s pretty solid evidence it’s not the system we have

⁶³ *Id.*

⁶⁴ R.A. DUFF, ANSWERING FOR CRIME: RESPONSIBILITY AND LIABILITY IN THE CRIMINAL LAW 87 (2007).

⁶⁵ *Id.* at 86.

in reality. Officials' beliefs about these matters are to a certain extent self-justifying, provided it's true (as I've argued) that the beliefs and other attitudes of the participants in the criminal process help determine that process's basic nature. Most legal officials appear to believe that the penal law marks out certain pre-legal wrongs as worthy of punishment, that the penal law consequently ought to constrain which wrongs offenders are punished for, and that sentencing doctrines like the *Apprendi* principle serve to ensure that these constraints function effectively.

At the same time, most legal officials also appear to believe that they've got every right to deploy the penal law in service of goals that aren't distinctively punitive, such as deterrence and incapacitation. And most legal officials also appear to believe that they've got every right to deploy the penal law for the sake of punitive goals other than those nominally espoused by the specific laws they're prosecuting people under. Legal officials widely endorse maneuvers like prosecuting murderers for tax evasion.

I don't mean to imply that instrumental deployments of the penal law are necessarily unjust. There's no injustice in punishing a murderer for tax evasion because he's a murderer. What's unjust is punishing a murderer for murder by convicting him of tax evasion. And our system makes it difficult to tell these two scenarios apart. That's because our system is a hybrid: part Babel, part Divine Justice. We enjoy the advantages of flexibility, but the discretion that makes flexibility possible comes at a price: we don't know—maybe can't know—how often our system dishonors the rights it proclaims sacrosanct.

VI.

An unapologetic and thoroughgoing instrumentalist about the penal law will call the uncertainty I've been insisting on irrelevant. All that matters to a thoroughgoing instrumentalist is whether an offender's punishment befits his wrongdoing.

But that isn't all that matters to the rest of us. Yes, it matters which punishment an offender gets, but it also matters how he gets it. Even an avowed instrumentalist about the penal law should blanch at framing a guilty man or punishing him extrajudicially. We don't just want an offender to receive a punishment that befits his wrongdoing. We want him to receive it through a proper legal mechanism. And that means a public process of adjudication culminating in proof beyond a reasonable doubt of the wrongdoing for which the offender's punishment is being imposed. Or at least that's what a proper legal mechanism means if we take the law's claims about procedural justice at face value.

Now some will concede the intrinsic significance of procedure but admonish us not to take the law's specific claims about procedural justice at face value. They'll tell us we should look beyond the ideals we preach to the methods we practice, and that we should acknowledge a principle that these methods reveal we accept covertly: that wrongdoing can be a proper object of punishment if proved beyond a reasonable doubt at a trial *or* if proved by a preponderance of the evidence at a sentencing hearing.

Yet no such over-lenient principle sits comfortably with our actual practices. Our actual practices don't sanction (even tacitly) the punishment of wrongdoing *unrelated* to the wrongdoing proved at a trial or admitted at a guilty plea hearing. In our system, a criminal conviction doesn't serve as a gateway to a forum of divine justice where an offender who's been proved guilty beyond a reasonable doubt of some offense or other may now be punished for any offense whatsoever. No such gates-wide-open principle is a norm of our actual legal practices—even an unacknowledged one.

If there's an unacknowledged norm of criminal sentencing—one more stringent than the gates-wide-open principle that says that a criminal conviction may serve as a gateway to a forum of divine justice but more lenient than the law's official view that the wrongdoing for which punishment is imposed must be proved at trial beyond a reasonable doubt—the unacknowledged norm of criminal sentencing is a *limited* gateway principle. It's a principle that says that wrongdoing proved only at the sentencing stage may indeed be an object of punishment *but only inasmuch as it relates to wrongdoing proved at trial*.

If we accepted this limited gateway principle as a true principle of political morality, we could concede a key claim of this essay—that for all we know our legal system often dishonors the adjudicative rights it proclaims sacrosanct—without thereby conceding that our legal system perpetrates injustice. We'd see these adjudicative rights as “merely” legal rights, rather than as moral ones. The only adjudicative right we'd recognize as a moral right is one our system dishonors less frequently: the right not to be punished for wrongdoing proved only at the sentencing stage which bears no connection to wrongdoing proved at trial.

But problems beset even the limited gateway principle. One problem is how to specify it: how must the wrongdoing proved at the sentencing stage relate to the wrongdoing proved at trial if the former is to be a proper object of punishment? A tempting answer is that the wrongdoing proved at the sentencing stage must relate to the wrongdoing proved at trial by revealing the latter as worse than it would be in isolation—as more wrongful, more harmful, or more culpable. But if we construe the relation this way, we construe it to require that the wrongdoing proved at the sentencing stage reveal the wrongdoing proved at trial as itself worthy of severer punishment. And this construal points us back to the law's official view, that the only proper object of an offender's punishment is the wrongdoing proved at trial.

Suppose for a moment that we manage to do something I suspect is impossible: articulate a cogent sense of “relates” less restrictive than “reveals as worse” but restrictive enough to impose a meaningful constraint on the kind of wrongdoing punishable at the sentencing stage. Suppose, in other words, that we manage to construct an intelligible and morally defensible limited gateway principle, a principle midway between the law's official view, that the only proper object of an offender's punishment is the wrongdoing proved at trial, and the gates-wide-open principle, that a criminal conviction is a gateway to a forum of divine justice. What will be our reward for constructing this principle? Our reward will be a difficult and embarrassing pair of questions: Why isn't our moderate and morally defensible

principle *the law*? Why is this principle instead something legal actors conspire to suppress?

These questions won't arise if we take the law's official view at face value—if we accept that punishing someone for wrongdoing proved only at the sentencing stage really is unlawful and unjust. Taking the law's official view of these matters at face value isn't the same as taking a legal official's self-description as accurate or sincere. Nor is it the same as regarding the legal process's formal features as determining the true object of punishment. It's instead a matter of regarding the law's avowed commitment to adjudicative rights as expressing a sound political morality.