2006

Purposes and Functions of Sentencing

Michael Tonry

University of Minnesota Law School, tonry001@umn.edu

Follow this and additional works at: http://scholarship.law.umn.edu/faculty_articles

Part of the Law Commons

Recommended Citation


This Article is brought to you for free and open access by the University of Minnesota Law School. It has been accepted for inclusion in the Faculty Scholarship collection by an authorized administrator of the Scholarship Repository. For more information, please contact lenzx009@umn.edu.
ABSTRACT

American sentencing systems have fragmented since the modern sentencing reform movement began in the 1970s and predominant retributive theories of punishment have become obsolete. Indeterminate sentencing was ubiquitous, but when it lost credibility, no single approach replaced it. Every American jurisdiction has a crazy quilt of diverse and, in principle if not practice, irreconcilable elements. Influential retributive theories of punishment were a 1970s response to 1970s problems that no longer galvanize opinion or reform. They are incompatible with the recent growth of restorative and community justice, therapeutic jurisprudence, prisoner reentry programs, and knowledge about the effectiveness of treatment programs. Theories of punishment can provide normative criteria by which policies and practices can be measured, but to be used in that way they must speak to their own times. New sentencing systems more effective and more just than current fragmented systems can be developed, and they will be better if new normative theories germane to them develop in tandem.

Sentencing policy in the United States has fragmented. There is no overriding theory or model. There are no widely shared understandings about what sentencing can or should accomplish or about conceptions of justice it should incorporate or reflect. Painting with a broad brush, utilitarian ideas held sway in the United States from 1900 through the mid-1970s (evidenced most strongly by indeterminate sentencing and extensive parole release), followed in succession by retributive ideas (evidenced most strongly by desert-based guidelines systems

Michael Tonry is Sonosky Professor of Law and Public Policy and director of the Institute on Crime and Public Policy at the University of Minnesota Law School, and senior fellow in the Netherlands Institute for the Study of Crime and Law Enforcement, Leiden. I am grateful to Marc Miller, who first proposed I write on the subjects of this essay. An earlier version of this essay appeared in the Stanford Law Review (Tonry 2005a).
in Minnesota, Oregon, and Washington) in the 1970s and 1980s, deterrent ideas (evidenced most strongly by federal and state mandatory minimum sentence and truth-in-sentencing laws) in the 1980s, incapacitative ideas (evidenced most strongly by three-strikes, dangerous offender/sexual psychopath, and sex offender registration and notification laws) in the 1990s, and a muddle in the early years of the twenty-first century. The muddle is exemplified by greatly reinvigorated interest in rehabilitative programs, such as drug, mental health, and domestic violence courts, reentry programs, and a plethora of new community-based and institutional treatment programs, combined with the continuing operation of laws enacted in all the earlier periods but especially the 1980s and 1990s. Ideas about restorative and community justice and therapeutic jurisprudence percolate through many of these initiatives.

Muddles are not necessarily bad. The absence of a stifling orthodoxy and a limiting conventional wisdom allows for creativity and experimentation, both of which have been evident for a decade. In the longer term, though, new practices and prevailing views will emerge and they will be incorporated into new sentencing systems and laws. As that begins to happen, it will be done better or worse. Thinking through the purposes and functions of sentencing in liberal democratic societies may improve chances it will be done better.

American governments did not do a very good job responding to the collapse of indeterminate sentencing. The developments underlying the collapse—declining confidence in rehabilitative programs, concerns about disparities and racial bias, and lack of procedural fairness—are well known (Blumstein, Cohen, Martin, and Tonry 1983, chap. 1). For a short time sentencing policy changes in the forms of sentencing and parole guidelines and abolition of parole release addressed those problems in some places. From the early 1980s through the late 1990s, though, policy makers lost their way. Some of the policies they adopted, notably mandatory minimums for drug and violent crimes, truth-in-sentencing laws, and three-strikes laws, worsened the problems the sentencing reform movement initially sought to address. Sentencing disparities generally and racial disparities in particular worsened, and proportionality links between the seriousness of crimes and the severity of punishment were broken (Tonry 2005b).

The winds, however, began changing direction again in the mid-1990s. Renewed possibilities exist for resuming the original efforts to
achieve fair and effective systems of punishment while taking account of changing conceptions of justice. Recent Supreme Court decisions and changes in political and public attitudes toward punishment have focused attention on sentencing policy to a degree not matched since the 1980s. The Court decisions fundamentally altered the federal sentencing guidelines, changed the legal context in which state guidelines operate, and led some states to reconsider the desirability of major elements of their sentencing systems. Evolving attitudes have produced significant policy changes in many jurisdictions.

Two decisions of the U.S. Supreme Court significantly altered the legal environment for guidelines systems. *Blakely* (124 S. Ct. 2531 [2004]) held that Washington State’s system of presumptive sentencing guidelines was unconstitutional to the extent that it gave judges discretion to impose prison sentences longer than the maximum authorized by the applicable guidelines. Such “upward departures” may be imposed, the Court asserted, only following a determination, beyond a reasonable doubt, of the facts on which the longer sentence was predicated. The Court’s rationale was that the upper limits of the ranges of presumptive sentencing guidelines are equivalent to statutory sentence maximums; just as a person cannot be sentenced more severely than the maximum authorized by statute for the offense of which he was convicted, a person cannot be sentenced more severely than the maximum provided in the applicable guidelines range for that offense. State sentencing commissions had to figure out how best to comply with *Blakely*. It was in some ways a harder problem in theory than in practice; upward departures are rare in every guidelines system.²

*Booker* (125 S. Ct. 738 [2005]) applied *Blakely* to the federal guidelines. *Booker*’s implications were substantially more radical than *Blakely*’s because the federal guidelines are much more detailed and prescriptive than any state’s and because their central feature is the “relevant conduct” provision that bases sentences not on the offense of which the defendant was convicted but on the “actual offense be-

---

¹ One option is to apply it literally and compel prosecutors to charge and prove aggravating factors that might justify a harsher than normal sentence; prosecutorial failure to do so would preclude an above-guidelines sentence. A second is to create a postconviction mechanism for jury consideration of aggravating factors at the sentencing stage. A third is to replace presumptive guidelines with voluntary ones. Most states with presumptive guidelines adopted the second option (Reitz 2005b).

² In Washington State in 2005, e.g., only 0.6 percent of standard cases involved upward departures (Washington State Sentencing Guidelines Commission 2005, table 10).
behavior” determined by the judge to be present on the basis of a civil-law “preponderance of the evidence” standard. Booker in practice applied to most sentences imposed in federal courts. The Supreme Court had to decide whether its decision required that the entire guidelines system be struck down. It elected to leave the guidelines in effect but with “advisory” rather than “presumptive” or “mandatory” force. Judges must calculate the advisory sentence post-Booker exactly as they would have earlier but may reject the guidelines’ “advice” and impose some other sentence as long as they give reasons for doing so. By making the guidelines advisory, the Court empowered judges more often and more easily to impose sentences less harsh than those prescribed by the pre-Booker guidelines.

Independently of the guidelines developments, public and political attitudes toward criminal justice policy have been changing for a decade. The Supreme Court decisions themselves, prohibiting sentencing policies in effect for two decades, are one sign of this. So is an earlier decision of the California Supreme Court, the People v. Alvarez (60 Cal. Rep. 2d 93 [1997]), which greatly weakened California’s fearsome three-strikes law. Another is the proliferation of drug, gun, domestic violence, and mental health courts, all premised on taking problem-solving rather than solely punitive approaches to offending (Nolan 2001; Casey and Rottman 2003; Redlich et al. 2005). A third is the passage of California’s Proposition 36, and similar referenda in other states and the District of Columbia, mandating diversion from prosecution and into treatment of many drug-abusing offenders (Tonry 2004b, chap. 1). A fourth is the adoption of laws in many states intended to reduce prison populations and costs by altering mandatory minimum sentencing laws, creating new alternatives to imprisonment, and authorizing new and broader programs of early release (Butterfield 2003). A fifth is the proliferation of new programs, usually at grassroots levels, based on ideas about restorative and community justice (Clear and Karp 1999; Braithwaite 2002). A sixth is the rapid expansion of the “prisoner reentry” movement predicated on the belief that we know

1 At least some California and U.S. Supreme Court justices would deny that the decisions occurred because the elected branches of government were politically immobilized from repealing or fundamentally altering sentencing systems that nearly all informed people considered deeply unjust and the courts decided they had to step in; it is hard, however, to imagine that the justices were unaware of the systems’ fundamental problems and of the intractable political difficulties that precluded legislative solutions and that their decisions would address the problems while sidestepping the politics.
how to help offenders avoid reoffending (Travis 2005; Travis and Visher 2005).

Sentencing reform seemed a much simpler subject in the 1970s than it does now. The widely shared and recognized aims then were to reduce sentencing disparities generally and in relation to race, to make judges more accountable for their decisions, and to make the sentencing process more transparent (Frankel 1972; Morris 1974). Few people saw sentencing reform as a way to prevent crime or as the proper subject of partisan politics. As a result, the proponents of change even in a state as traditionally partisan as California included liberal and conservative legislators, police unions and prisoners' rights advocates, and defense lawyers and prosecutors (Messinger and Johnson 1978). Sentencing reform seemed straightforward, raising practical and policy issues about which reasonable people disagreed, but, at least politically and ideologically, it appeared to be relatively noncontentious.4

The trial and error of the past three decades has left lessons to be learned that, if heeded, can help contemporary policy makers do better than their immediate predecessors. We know much more about sentencing and the effects of different policy approaches than we did thirty years ago. Here are the main lessons:

- Well-designed systems of presumptive sentencing guidelines can reduce disparities, make sentencing more consistent and predictable, attract support from most judges and prosecutors, and provide a policy tool for predicting and managing correctional resource needs (Tonry 1996, chaps. 2–5; Reitz 2001).
- Some systems of voluntary sentencing guidelines have been used successfully to predict and manage correctional resources, but there is no credible evidence that they reduce disparities (Tonry 1996; Frase 2005c).

4 That sentencing reform was widely seen as noncontentious in the 1970s seems politically naive in retrospect and therefore probably is not true. A more plausible view is that people recognized that various proponents of change had different policy and political agendas in mind but failed to anticipate how partisan and ideological crime control policy would later become and how starkly different the ramifications of the different agendas. When California adopted its Uniform Determinate Sentencing Law of 1976, more than a decade had passed since Barry Goldwater raised crime in the streets as a campaign issue in 1964, and President Richard Nixon had already launched wars on drugs and crime (Edsall and Edsall 1991). No one then would have predicted that Willie Horton, a black murderer and rapist, would play a major role in the 1988 presidential election (Anderson 1995) or that Democratic presidential candidate Bill Clinton in 1992 would campaign on the basis that he was tougher on crime than the Republicans (Gest 2001).
Mandatory minimum sentencing laws, including three-strikes laws, are inconsistently applied, provoke widespread circumvention, produce gross disparities in the treatment of similarly situated offenders, and sometimes result in sentences that all involved believe to be unjustly severe (Merritt, Fain, and Turner 2006; Tonry 2006).

Well-run and well-targeted treatment programs of a variety of sorts have been demonstrated to be capable of achieving significant reductions in reoffending (Gaes et al. 1999; Halliday 2001; Welsh and Farrington 2006).

Restorative justice programs of diverse kinds have been shown to be preferred to regular criminal justice processes by offenders and victims alike, more likely to result in offenders' successful completion of resulting obligations, and no more likely to result in reoffending (Braithwaite 2002).

With that knowledge, and the aspiration to do so, we could build sentencing systems that take account simultaneously of the state's responsibilities to respect its citizens' liberties and autonomy and to assure the stability and security that allow people to live their lives in peace.

Building sentencing systems that are rational, just, and effective starts with understanding what sentencing is for and what it can do. This essay therefore examines the purposes and functions of sentencing and sentencing reform. Unless we are clear about them, we will not have much success at designing sensible, effective, and workable systems for the twenty-first century.

"Sentencing reform" means different things to different people, depending on whether the proponent wants sentencing made softer, tougher, fairer, more consistent, more efficient, more economical, more transparent, or more effective at preventing crime. Whether a proposed change counts in the eyes of others as a reform depends on what the proponent wants to accomplish and whether others think that a good thing.

The same observations apply to sentencing. Whether a sentencing system or practice can be said to work well depends on its purposes, what it is supposed to do, and how well it does that. Generally when people discuss the "purposes" of punishment, they refer to normative rationales such as retribution or crime prevention through deterrence, incapacitation, rehabilitation, and moral education.

Examination of normative purposes, and how adequately and con-
sistently they are manifested in sentencing policies and practices, however, is only the beginning of analysis of the effectiveness of a sentencing system. Practitioners and policy makers want sentencing to accomplish other things as well. They include efficiency, cost-effectiveness, and public confidence. They also sometimes want sentencing systems, and legislation affecting them, to help them achieve personal goals, advance ideological agendas, and pursue partisan political objectives. Pursuit of all these aims can be thought of as a "function" of sentencing or sentencing reform.

Functions can be thought of instrumentally, as though sentencing systems and policies were machines. Whether a machine can be said to work well depends on what it is supposed to do, and typical occupants of various institutional roles sometimes differ about that. Distinctions need to be made among primary functions, the things we want to do directly; ancillary functions, other things related to primary functions that we also want to accomplish; and latent functions, collateral things some people want to accomplish that are unrelated to the aims of primary or ancillary functions.

Primary and ancillary functions are straightforward because they are instrumental. Certain things are done in pursuit of certain aims. Latent functions are more complex because, while they can be instrumental, they need not be. They can be thought of in terms of social functionalism, doing things whether we want them to or not, and whether we are aware they do them or not. Loic Wacquant has argued, for example, that American punishment policies since 1970 have served to maintain racial hierarchy in the United States, something achieved in earlier generations through slavery and then through Jim Crow laws and related social conventions (Wacquant 2001). Contemporary punishment policies need not consciously have been directed at that result, but that is what (Wacquant would say) they have done and what accordingly can be thought of as one of their latent functions. A more benign Durkheimian example of a not necessarily purposive latent function is the norm-reinforcing (or undermining) effects of imposition of particular punishments for particular crimes (Durkheim [1893] 1933). Some judges may have this in mind when they impose sentences, but it will happen whether they do or not.

For the rest of this essay, I set noninstrumental latent functions aside. Here I am interested in instrumental latent functions, things practitioners and policy makers want to accomplish that are unrelated to the
primary functions of punishment and prevention. The most common of these are pursuing the actor's personal self-interest, pursuing partisan political objectives, and bearing witness to ideological beliefs.

People setting sentencing policy or imposing sentences in individual cases have multiple motives and pursue diverse objectives, and these may be more or less legitimate. The primary functions of sentencing, most people would agree, are imposition of appropriate punishments and prevention of crime. People differ on what makes a punishment appropriate and how best to prevent crime, but few people would disagree that these are legitimate functions.

There are greater disagreements about the ancillary functions. Some people believe that defendants must be offered incentives to plead guilty in order to manage case loads; others believe that inappropriate. Some people believe that maintaining good personal and professional relations with others in the courtroom work group can justify plea bargains that might not otherwise be countenanced; others disagree. Some people, especially corrections officials, believe that judges should take resource considerations into account when setting sentences; others, especially judges, disagree. These disagreements are the products of different conceptions of justice but relate closely to considerations of appropriate punishments and effective crime prevention.

Few people would disagree that some latent functions are palpably illegitimate. For example, in order to win favorable publicity, a judge up for a vigorously contested reelection might impose a much harsher sentence in a particular case than can be justified in his own mind or than at other times he would. The latent function would be winning reelection, and few would regard that as a legitimate reason to punish an offender especially severely. Nor would most judges themselves regard it as legitimate: a judge would be little more likely publicly to justify a sentence “because that sentence will help me get reelected” than “because my daughter is the prosecutor,” “because I'm in a bad mood today,” or “because I always punish Hispanics more harshly than whites.” Policy decisions or individual sentences that result not from good-faith efforts to achieve the ends of appropriate punishment or effective crime prevention but from the decision maker's wish to achieve other unrelated personal, ideological, or political goals are illegitimate.

If future sentencing systems are to be more successful than their recent predecessors, they must attend to both purposes and functions.
Many initiatives of the past two decades failed to address purposes coherently. The best examples are broad three-strikes laws such as California’s (Zimring, Hawkins, and Kamin 2001), the federal 100-to-one crack/powder cocaine sentencing differential (McDonald and Carlson 1993), and many mandatory minimum sentence laws for drug offenses (Tonry 2006). These all resulted in individual sentences that those obliged to apply them often found unjustly severe, and all violated commonsense notions of the relative seriousness of the crimes they affected and other more serious crimes that they did not.

Some misconceived laws resulted from policy makers’ overreaction to emotional incidents, such as the death of college basketball star Len Bias, apparently from a crack cocaine overdose, commonly credited as driving the enactment of the federal 100-to-one law (Kennedy 1997). Others may have resulted from simple failures to understand the implications of seemingly straightforward changes. Many, however, resulted not from confused purposes or failures to forecast a law’s effects but from the use of sentencing “reforms” to achieve personal, partisan, and ideological goals. Put differently, illegitimate latent goals displaced legitimate primary and ancillary goals. The inevitable result is that individual human beings, like the hypothetical defendant sentenced more severely to help the judge get reelected, received unjust punishments.

Doing better requires two things: first, that both purposes and functions be considered in setting sentencing policy, so that primary and ancillary functions can be taken into account in ways that do not produce individual injustices; second, that illegitimate latent functions be recognized for what they are and eschewed. The second of these can never be completely successful since that would require that human nature be changed. Policy makers may be able, however, to do better than they recently have done if they come to understand that sentencing has multiple functions, that ethical policy making requires that they be recognized, and that many of the latent functions cannot be openly and honestly acknowledged without embarrassment.

The rest of this essay provides the analysis on which the preceding observations rest. Section I sets out a fuller framework for thinking about purposes and functions of sentencing. Sections II and III, applying that framework, examine current knowledge concerning the primary and ancillary functions. Section IV examines latent functions.
The last distills lessons honest policy makers might wish to consider as they devise the sentencing systems of the twenty-first century.

If only legitimate primary and ancillary functions were to be sought to be performed and account were taken of what has been learned from the last thirty years of changes in sentencing laws, here is what a rational, just, and effective sentencing system would look like:

- Only normative purposes and legitimate primary and ancillary functions of sentencing would be taken into account in setting sentencing policy or setting sentences in individual cases, and illegitimate latent functions would be eschewed in both contexts.

- Limiting retributivism, the idea that retributive concerns can set meaningful upper and lower bounds within which other considerations may validly be taken into account in setting sentences in ordinary cases, would be adopted as the system's primary normative rationale, and its scope would be extended to encompass both purposes and functions.

- Subject to that premise, guidelines would set out ranges of presumptive sentences from within which judges would normally select sentences in ordinary cases; but judges would have authority, subject to observance of the requirements of Booker and Blakely, to depart from guidelines to impose other sentences as long as they specified for the record their reasons for doing so.

- In setting presumptive sentencing ranges, policy makers would not take into account evidence on deterrent and incapacitative effects of sentencing but would authorize judges to take such evidence into account in setting individual sentences within (but not above) presumptive ranges.

- In promulgating sentencing guidelines, policy makers would authorize judges to take account of rehabilitative and restorative programs and possibilities to individualize sentences both within and below (but not above) presumptive ranges, in the latter case as long as they specified for the record their reasons for doing so.

- All mandatory minimum, including three-strikes, laws would be repealed.

I. A Framework
The fundamental purposes and primary functions of sentencing are clear, and are the same: to punish criminals and prevent crimes. There
are, however, other functions that concern officials and policy makers. They range from encouraging most defendants to plead guilty and managing criminal justice budgets to reassuring the public and getting reelected.

People concerned primarily with the word justice in "criminal justice" would say that the overriding purpose of sentencing is the imposition of punishments that are just relative to prevailing normative criteria; the overriding functional goal would be to ensure, or at least maximize the likelihood of, imposition of just punishments according to those criteria. People concerned primarily with the word "criminal" would say that the overriding purpose of sentencing is the prevention of crime and accordingly that the overriding functional goal is to minimize the incidence of crime and its consequences. To distinguish the question of what purposes should be pursued from the question of whether laws have been fairly and effectively applied relative to those purposes, I refer to the first as purposes and the second as distributive functions.

The distinction between purposes and functions may appear a bit obscure, but it should not be. Richard Frase and I were recently asked by the editors of the *Stanford Law Review* to write separate essays on purposes (Frase 2005a) and functions (Tonry 2005a) of sentencing. Initially we thought the distinction a bit artificial but decided it was not, for two main reasons. First, normative purposes provide the theoretical criteria for deciding whether sentences imposed on individual offenders are just or appropriate, whereas distributive functions relate to how well or poorly any particular sentencing system distributes sentences according to those criteria. Second, sentencers and designers of sentencing systems have multiple functions in mind when they make decisions, and some of them, such as managing resources effectively, are entirely independent of normative purposes of sentencing.

All mainstream contemporary theories include crime prevention among the purposes of punishment, though they vary widely in what ways, with what weight, and subject to what constraints prevention may be taken into account. Most people, however, believe that prevention and diminution of crime, fear of crime, and their consequences are important and legitimate functions of sentencing. These are the preventive functions.

Many practitioners would also include efficiency, cost-effectiveness, and resource management among the functional goals of an acceptable
sentencing system. Managers need to set and pursue substantive priorities, allocate personnel and resources, meet performance goals, and operate within their budgets. They also need to maintain good working relations with other agencies and officials who can make their work easier or harder, and thereby make achievement of other goals easier or harder. These are management functions.

Some policy makers and analysts urge that it is important that governments reassure the public, maintain confidence in the legal system, denounce wrongful behavior, and reinforce basic social norms. Contemporary analyses of “expressive” policies emphasize government officials’ wishes and needs to assure the public that things are being done to address subjects that trouble them (Garland 2001). Recent work on legitimacy and procedural justice emphasizes that justice needs to be seen to be done (Tyler 2003). Durkheimian analyses portray criminal law as a primary contributor to maintenance and refinement of basic social norms and its dramaturgical features as key sources of social cohesion (Durkheim [1893] 1933). These are communicative functions.

The first two of these, distribution and prevention, concern primary goals (what sentencing “is supposed to do”). The third, management, is an ancillary goal (how institutional imperatives can be acknowledged while pursuing primary goals).

The fourth, communication, is the most complex and multifaceted and operates at all three functional levels. Communication of legal threats and acknowledgment in court processes and outcomes of basic social norms are necessary to achieve the primary goals of distribution and prevention. Communication about court processes and procedures and likely punishments is key to achievement of a variety of management goals. These might, however, be thought of as incidental components of primary and ancillary functions. To avoid repetition, I discuss communicative functions among the ancillary functions, even though each of the three functional levels includes communicative elements.

The more controversial forms of communication relate to latent functions. Practitioners and policy makers are sometimes moved by personal, ideological, and partisan objectives. People sometimes make decisions or support policies because it is in their personal self-interest, because they want to bear witness to their ideological beliefs, or because they want to pursue partisan political advantage. Earlier I provided an example of a judge imposing a harsher than otherwise sen-
tence in order to help his reelection prospects. A parallel example for policy making is the legislator or governor who proposes policy changes he knows will not achieve their purported primary aims and will, if enacted, cause new injustices, because he hopes their enactment will help him get reelected (e.g., California Governor Pete Wilson’s three-strikes proposal: Zimring, Hawkins, and Kamin [2001]).

In the preceding paragraphs, I identified a number of purposes and functions:

A. Purposes (belaboredly, these could be called normative functions)
B. Primary functions
   1. Distribution (consistency, evenhandedness, fairness)
   2. Prevention (crime, fear of crime, costs and consequences of both)
   3. Communication (threat communication, denunciation of wrongful behavior, reinforcement of basic social norms)
C. Ancillary functions
   1. Management (efficiency, cost-effectiveness, resource management)
   2. Communication (procedural justice, legitimacy, public confidence)
D. Latent functions
   1. Self-interest
   2. Ideology
   3. Partisanship
   4. Communication

Normative purposes provide the ultimate criteria by which the justness of a punishment system is assessed. The most important of the remaining functions are the distributive functions of doing justice and the preventive functions of minimizing crime and its consequences. These can be pursued effectively only if account is taken of the implications and constraints associated with the ancillary functions and if ways can be devised to limit pernicious influences of the latent functions.

II. Purposes and Primary Functions
The pre-Booker federal sentencing guidelines are a case study in what can happen when policy makers fail to be clear about their purposes
and functional goals. Here are four examples. First, the U.S. Sentencing Commission did not make the guidelines’ normative purposes clear. It could have; the exceedingly detailed guidelines fine-tuned sentencing to circumstances (e.g., weapons, victim injury, property loss, role in the offense) that relate closely to culpability and thus were largely consonant with retributive purposes. Instead, the commission chose not to make the “profoundly difficult” choice of adopting explicit purposes, electing instead “an empirical approach that uses data estimating the existing sentencing system as the starting point” (1987, pp. 1.3–1.4). Guidelines for drug crimes were a major exception to the “empirical approach,” with the result that the guidelines prescribed much harsher punishments for drug crimes than for violent crimes that most people consider more serious and thereby violated widely held proportionality notions. Probably the single most recurrent criticism by practitioners was that the guidelines sentences for drug cases were unjustly severe. This led to widespread circumvention, which increased as time passed (Bowman and Heise 2001). Adopting explicit purposes, and then being guided by them, would have produced guidelines that were more internally consistent and normatively acceptable to practitioners and would more effectively have performed their distributive functions.

Second, the commission was worried that lawyers would manipulate the guidelines by means of charging and plea bargaining decisions and adopted the “relevant conduct” policy as a countermeasure. The logic was that if judges were required to impose sentences on the basis of the defendant’s “actual offense behavior,” charges to which the defendant pled guilty would generally not much matter (U.S. Sentencing Commission 1987, pp. 1.5–1.6). In practice, the system did not work that way. The sentencing guidelines were widely circumvented by prosecutors, defense lawyers, and judges (Nagel and Schulhofer 1992; Schulhofer and Nagel 1997; Bowman and Heise 2001). The notion that offenders are to be punished for what the judge believed they did, as opposed to what they were convicted of, offended widely held notions of procedural fairness and produced such unjust results that the Supreme Court intervened in Booker and demoted the guidelines from

---

5 A survey of federal appellate and trial judges conducted by the commission found that 73.7 percent of district court judges and 82.7 percent of circuit court judges reported that drug-trafficking sentences were “greater than appropriate” (U.S. Sentencing Commission 2003, exhibits II-4, III-4).

6 The major exception is that guidelines are trumped by plea bargains to offenses bearing lower statutory maximums than guidelines prescribe.
Purposes and Functions

strongly presumptive to advisory. The fundamental difficulty was that policy makers in the U.S. Sentencing Commission, the U.S. Department of Justice, and the U.S. Congress cared less about sentencing's distributive functions than did the judges and prosecutors whose job it was to apply them.

Third, by attempting to constrain prosecutorial choices radically, the guidelines failed to acknowledge practitioners' needs to achieve various management goals and thereby obstructed performance of ancillary functions practitioners believed important. Judges and prosecutors typically believe that the smooth operation of the courts depends on high rates of guilty pleas and that incentives are necessary to induce them. This is one of the foreseeable reasons why the guidelines often were circumvented.

Fourth, the guidelines failed to perform some of their communicative functions adequately; through their rigidity and severity and the relevant conduct rule, they undermined perceptions of their own legitimacy in the eyes of the judges and lawyers who were expected to implement them. Legal rules that possess moral legitimacy are much more likely to be followed than those that do not.

Judges and prosecutors see themselves as in the business of doing justice. Judges, prosecutors, defense counsel, and probation officers work together on an ongoing basis in most courtrooms. Shared understandings evolve about "going rates" and about reasonable and unreasonable sentences and modes of official behavior (Eisenstein, Fleming, and Nardulli 1988, p. 376). Practitioners in work groups believe that the standards they observe and apply are good ones, not merely cynically because observing those standards helps them accomplish what they want to do, but because the standards are normatively appropriate and result in reasonably just results. Arthur Rosett and Donald Cressey described this in their report on plea bargaining in California: "Even in the adversary world of law, men who work together and understand each other eventually develop shared conceptions of what are acceptable, right, and just ways of dealing with specific kinds

7The best illustrations of this are the Congress's enactment in 1986 of mandatory minimum sentences for drug crimes that required sentences much harsher than were typically imposed for violent crimes and the commission's decisions to incorporate even more severe drug crime penalties into the guidelines than the Congress prescribed (Tonry 1996, pp. 96-98). These decisions had the effect of making sentences much severer and the guidelines much more complicated than they needed to be to pursue sentencing's legitimate primary and ancillary functions.
of offenses, suspects, and defendants. These conceptions form the bases
for understandings, agreements, working arrangements, and coopera-
tive attitudes. Norms and values grow and become a frame of reference
which prosecutors, defense lawyers, judges, and experienced offenders
all use for deciding what is fair in each case" (1976, pp. 90–91).

When laws or guidelines prescribe sentences that are harsher than
most practitioners believe reasonable or just, there is a problem. When
laws require that sentences be calculated primarily by means of me-
chanical scoring systems, as the federal guidelines did, rather than by
looking closely at the circumstances of individual cases, as most prac-
titioners believe justice requires, there is a problem. Judges and lawyers
are socialized into both enforcing the law and doing justice. When
they cannot do both, doing justice often trumps enforcing the law,
especially when, as in the case of the federal guidelines, the law lacks
moral authority. The federal guidelines, however, were but an extreme
instance of a wider failure to take conscious account of sentencing’s
multiple and sometimes conflicting purposes and functions.

A. Purposes (Normative Functions)

The overriding normative function of a sentencing system in a so-
ciety committed to individual liberty, procedural fairness, and limited
powers of government is to assure that individuals convicted of crimes
receive the sentences that, in principle, they should. Those final words
are purposely indeterminate about the principles that should govern.
Preponderant views about this vary across time and space. For most of
the twentieth century through the 1970s, utilitarian ideas about crime
prevention through rehabilitation and incapacitation were predominant
(Rothman 1980); they have not gone away, but in more recent decades
they have been eclipsed by retributive and just deserts emphases on
the desirability of apportioning the severity of punishment to offend-
ers' blameworthiness (e.g., von Hirsch 1976; von Hirsch and Ashworth
2005). Interest in rehabilitation has revived in the early twenty-first
century (as evidenced by drug and similar courts offering individual-
ized, treatment-based dispositions, a plethora of new treatment pro-
grams in institutions and in the community, and the "prisoner reentry"

8 If one assumed that policy making since the mid-1980s was implicitly based on prin-
cipled ideas about purposes, they would be a not very coherent mixture of retributive
and preventive purposes; I know of no scholar who has attempted to delineate a normative
theory that would justify recent policies.
movement), and interest in restorative (Braithwaite 2002) and community justice (Clear and Karp 1999) and in therapeutic jurisprudence (Winick and Wexler 2003) has burgeoned. These developments may signal an immanent shift in preponderant views.

Immanuel Kant, the German idealist philosopher whose ideas are often regarded as marking the beginning of development of modern theories of retributive punishment, argued that the only morally appropriate justification for punishment is that it is deserved and is appropriately apportioned to the offender's wrongdoing (Kant [1796] 1887). Jeremy Bentham ([1789] 1970), the bellwether of utilitarianism, to the contrary, argued that only net positive preventive effects—achieved usually through deterrence, incapacitation, and rehabilitation—can justify the imposition of pain that punishment entails. A Benthamite corollary, commonly referred to as the principle of parsimony, forbids imposition of pain that is not outweighed by pains averted.

The normative choices involved in deciding what principles should govern sentencing have important ethical and moral dimensions about which reasonable people differ, sometimes fundamentally, and the choices sometimes have radically different implications for policy. A thoroughgoing Kantian theory of sentencing would predicate just punishments on the relative degree of blameworthiness imputable to the offender on the basis of his culpability and, perhaps, the harm he has caused. There would be no room in the calculus for incapacitative concerns based on predictions of dangerousness, deterrent concerns based on belief that prospective wrongdoers might be dissuaded, rehabilitative concerns aimed at enhancing the offender's prospects for law-abidingness, or restorative concerns aimed at healing broken relationships, unless they could be achieved within the scope of a sentence justified in retributive terms.

Conversely, a thoroughgoing Benthamite theory of sentencing would allow those various crime-preventive concerns to be acted on irrespective of the seriousness of the offender's culpability, constrained only by the parsimonious principle that no punishment cause greater harm, taking offenders into account when making the calculation, than

---

9 Or, some might say, the harm he foreseeably caused, but this is redundant since foreseeability of harm is an element of any coherent conception of culpability.
it prevents.\textsuperscript{10} Ultimately, a utilitarian calculus could justify imposition of a very serious punishment on an offender who committed a venial offense, or even on an innocent person whom everyone except the judge believed to be guilty.\textsuperscript{11}

No American jurisdiction has ever created a punishment system based on thoroughgoing retributivist or utilitarian premises, but several have come close. Washington State sentencing guidelines in the 1980s were primarily based on just-deserts reasoning, appellate courts enthusiastically monitored judicial compliance, parole release was abolished, and the only conditions to which probationers and parolees could be subjected related to their criminality (Boerner and Lieb 2001). The federal sentencing guidelines, even though the commission insisted it had not adopted normative premises, were rigid and prescriptive and closely apportioned sentences to detailed offense characteristics that most people would have little difficulty relating to relative blameworthiness. Until California abolished parole release in 1976, all prison sentences were for "one year to the statutory maximum," and in theory prisoners were to be released only when they had been rehabilitated or no longer required incapacitation. California’s parole system, however, is long since gone, Washington’s guidelines have changed in many respects, and the federal guidelines have been eviscerated.

There are good reasons why no relatively philosophically pure sentencing system exists or has been seriously proposed at a policy level. The Kantian and Benthamite visions offer models that identify important moral choices, trade-offs, and compromises. Neither, however, can be converted to the real world. The Kantian ideal has well-known problems. Its claim that respect for individuals' moral autonomy requires that we punish them "in like kind," perfectly commensurately with the wrongdoing embodied in their crimes (Pincoffs 1991, pp.

\textsuperscript{10} I ignore for purposes of this essay the complexity of distinguishing between rule and act utilitarianism and the practical problems of ever knowing enough to make the necessary utilitarian calculations.

\textsuperscript{11} The accusation that utilitarians must be prepared, if true to their principles, to punish innocent but apparently guilty people is a classic retributivist objection to utilitarianism. Utilitarians generally reply that information about such punishments of innocents would inevitably slip out and foster a general sense of insecurity ("it could happen to me") that would be more harmful than any preventive gains fleetingly gained from the undeserved punishment (Hart 1968, p. 25). It is an irresoluble argument but a red herring. Harsh punishments of people guilty of trifling crimes raise the same problem. No one argues that a utilitarian system would prohibit harsh punishments for minor offenders on the basis of deterrent, incapacitative, or rehabilitative considerations, as long as the parsimony principle is observed.
16–17), cannot be converted into practicable policy. Killing a murderer fits the model, but what about the punishment for an attempted murderer? How can a rape be punished in like kind or, harder, an attempted rape? Would anyone really want to torture or maim an assailant who caused serious pain or disfigurement? How can deserved punishments for property crimes be calculated, or “victimless” crimes, or inchoate crimes (conspiracy, attempts, aiding and abetting)?

People have tried to solve these problems, but the solutions necessarily are imperfect. Andrew von Hirsch devised a system of ordinal proportionality in which punishment in absolute like kind is not sought but punishment proportionate to the blameworthiness expressed in the crime (von Hirsch 1993). “Proportionality” here, however, is not relative to the nature of the crime but to punishments for other more, less, and comparably blameworthy crimes. The difficulties here, though less well known than those affecting a purer Kantian model, are also considerable (Tonry 2005b). There are three principal ones. First, the principled call for comparable punishments for comparably situated offenders is illusory. Widely divergent practices concerning charging, plea bargaining, and sentencing discounts for guilty pleas make seemingly comparable offenses of conviction widely diverse and seemingly comparable offenders substantially different in their blameworthiness.

Second, definitions of offenses in U.S. law are typically broad: “robbery” encompasses behavior ranging from gangland holdups to takings of basketballs on school playgrounds. Punishing every person convicted of robbery the same way would produce what Arie Freiberg (2001) called “unjustifiable parity.” The example I just gave is extreme, of course, but the problem is a general one. Most guidelines systems divide all felonies into a small number of categories of severity, and they inevitably include widely diverse kinds of crimes. Because of wide differences in any jurisdiction in charging and plea bargaining practices, substantially similar crimes result in convictions for different crimes, and conviction offenses obscure wide differences in underlying culpable behavior.

Third, most important, von Hirsch’s desert theory is subject to the “punishment of the innocents” charge that retributivists classically toss at utilitarians. The logic of the system requires that all offenders who fall within a category receive the same penalty, including when the severity of punishment for that category is increased for preventive
reasons relating to some but not all members of the category. What this means is that some, perhaps many or most, offenders in a category will be punished more severely than their offense warrants because the architecture of the system requires it. In the words of the classic retributive critique of utilitarian punishment theories, the offenders being punished more severely than they otherwise would be are being "used" as means to the end of maintaining the formal structure of the sentencing theory (Pincoffs 1991). As an ethical matter, it is difficult to distinguish between knowingly punishing an innocent person and knowingly punishing a guilty person more than he or she deserves.

A pure utilitarian model also has well-known practical limits, over and above the punishment of the innocents problem. The overriding one is that the necessary "felicific calculus," as Bentham called it, is impossible. Fundamentally contested literatures on economic costs of crime, especially pain and suffering (McDougall et al. 2006), deterrent and incapacitative effects (Nagin 1998), and crime prevention more broadly demonstrate that we are far away from a time when we know enough to take the Benthamite model seriously. Any decisions about the utilitarian gains and pains associated with punishment in an individual case or within broad-based general policies can be little better than guesswork.

On a continuum between those Kantian and Benthamite poles lie hybrid punishment theories that combine retributive and preventive considerations in diverse ways. One, proposed by Andrew von Hirsch (1993), is that culpability concerns predominate heavily but that deterrent or incapacitative concerns may be taken into account in setting the penalty scale. If longer sentences are wanted, they may be imposed collectively on all equally culpable people, by moving that offense up the scale or by extending the scale, and thereby punishing all offenders more severely, but they may not be imposed on an individualized basis. This is unsatisfactory because it forces a choice between forgoing

---

12 This sentence conflates two separate problems. First, insofar as overbroad offense definitions or charging and plea bargaining practices result in substantially different crimes bearing the same label, sentencing all bearing that label in the same way punishes those whose crimes were less serious more severely than they deserve. Second, as I explain below in the text, von Hirsch would forbid the use of crime preventive considerations (e.g., deterrence or incapacitation) to increase sentences for individuals and would instead allow them to be taken into account only to increase sentences for the entire category of offenses in which target individuals fall (von Hirsch and Ashworth 2005). The effect would be to increase sentences for all in the category so that some can be punished more severely for instrumental reasons.
seemingly realizable preventive goals and gratuitously punishing some, or many, in a class more severely than their crime would otherwise merit. The former is politically and possibly ethically impracticable. The latter violates the parsimony principle’s injunction against imposition of gratuitous suffering.\textsuperscript{13}

A second hybrid theory, associated with Andrew von Hirsch and Martin Wasik (Wasik and von Hirsch 1988), would permit preventive considerations to influence the type but not the severity of punishment. Punishments must be apportioned to blameworthiness; but if two punishments (e.g., $X$ days of imprisonment versus $Y$ days of community service) have a comparable punitive bite, a choice may be made between them on preventive bases. Their proposal assumes, however, that few penalties can be made comparably punitive and accordingly specifies that imprisonment is more severe than community-based punishments; they in turn are more severe than fines, allowing for interchangeability only at the margins between types of punishments. This is unsatisfactory generally because it is primarily a variation of the preceding hybrid theory and because it consigns preventive considerations to the narrowest margins.

A third hybrid theory, proposed by Paul Robinson (1987), calls for strict observance of culpability standards except in a small number of cases for which disproportionate punishments would prevent really serious harms. This is unsatisfactory for conceptual (how in principle justify the abandonment of retributive premises?), definitional (what harms count as really serious?), and empirical reasons (how can we know we’re additionally punishing the correct people? We’re not very good at predicting really serious future harms).

A fourth hybrid, proposed by former English civil servant John Hallday (2001, app. 8), is described as being a form of limiting retributivism concerned with “consistency of approach rather than uniformity of outcome” (p. 20). The notion is that the same set of punitive and preventive considerations should be taken into account in every case, but the seriousness of the offense would limit the maximum sentence that should be imposed. This is at best obscure and implies something closer to a theory of procedural than to a theory of substantive justice.

\textsuperscript{13} Von Hirsch’s response (1993) would be that parsimony should be sought at collective levels by setting the overall penalty scale as low as possible while furthering the aim of treating like cases alike. General parsimony-based objections could be raised to this position, but, importantly for the point made in the text, it does not speak to the special problems involved in raising the overall scale to pursue preventive goals.
It is in any case a very weak form of limiting retributivism in that the upper limits of aggregate punishment that may be imposed are vaguely expressed and are subject to major exceptions for serious offenses, repeat offenders, and offenders who fit into a very broad definition of dangerousness (Tonry 2004a).

The fifth and sixth hybrids are other versions of limiting retributivism. In “asymmetric proportionality,” culpability sets strict upper limits on the sentence that may be imposed, but there are no lower limits (Lappi-Seppälä 2001). Finnish law sets out fairly detailed statutory principles that govern sentencing; they give pride of place to culpability considerations but allow on a case-by-case basis for the operation of other considerations including preventive considerations in a narrowly defined set of cases. The Finnish system, however, operates in a country in which most sentences are measured in days or weeks, sentences longer than a year are uncommon, and the total imprisonment rate including pretrial detainees fluctuates between sixty and seventy per 100,000 population, less than a tenth the American rate. Finland has no sentencing guidelines as Americans understand the term because the latitude for unwarranted disparities is incomparably lower. Asymmetric proportionality on the Finnish model is probably unsatisfactory for American adoption because it depends on Finland’s very different punishment practices and penal culture and its application by career cadres of judges and prosecutors who subscribe to its values.

In Norval Morris’s limiting retributivism (1974), culpability sets upper limits and for some but not all crimes sets lower limits. Compared with von Hirsch’s proportionality doctrine, in which desert limits are artificial, set relative to sentences for lesser and greater crimes, Morris’s are cultural. He famously urged that “not undeserved” was a more useful conception of culpability constraints on punishment than “deserved” because it is far easier for people to agree that a sentence is undeservedly harsh or unduly lenient than that it is just right. Within those bounds, however, for Morris, Bentham’s parsimony principle, cause no unnecessary or gratuitous pain, required that the least severe “not undeserved” punishment be imposed.

Morris first set out his limiting retributivism argument in the early 1950s (Morris 1953), and it took its fullest form in his Future of Imprisonment (1974), before modern sentencing guidelines were more than a gleam in anyone’s eyes. There is no reason in Morris’s argument why the upper and lower bounds of allowable punishments cannot de-
rive from policy decisions that grid-style guidelines made possible (as long as they respect cultural limits), as his later writings showed (e.g., Morris and Tonry 1990, chap. 3).

It appeared clear to Morris in 1974 that preventive considerations would seldom justify imposition of sentences more severe than the parsimonious minimum. The most successful predictions of dangerousness were wrong at least two times out of three, too often to justify confining people longer because they were purportedly dangerous, and the research evidence concerning deterrence and rehabilitation gave no grounds that either was sufficiently often effective\(^4\) to justify deprivations of liberty on their basis.

The world has changed since 1974 in ways germane to Morris's analysis. Though the evidence concerning the effectiveness of deterrent and incapacitative penalties is not much stronger now than it was then, doubts about basing sentencing laws on them are much less influential. More significantly, the research evidence on rehabilitative programs has reversed, and both conventional wisdom and many meta-analyses and literature reviews agree that well-run, well-targeted programs can reduce reoffending rates (and drug abuse) (e.g., U.S. Surgeon General 2001; Farrington and Welsh, in this volume; Welsh and Farrington 2006).

This discussion of purposes supports two conclusions. First, as the discussion of the federal guidelines at this subsection's beginning showed, purposes have consequences and provide guidance that is necessary if policies are to be coherent, internally consistent, and perceived as just. Second, any coherent and practicable set of purposes will be embodied in a hybrid model.

In the sections that follow, and in the conclusion, I propose that both sentences imposed and legitimate functional goals pursued should be consistent with whatever normative purposes are set. For myself, and for most modern sentencing systems (Fraser 1997, 2004, 2005a), something like Morris's limiting retributivism appears best to express the optimal balance of competing moral intuitions about punishment and crime prevention, but reasonable people disagree about that.

\(^4\) In these views he was ahead of his time; Martinson's "What Works?" article, suggesting that existing evidence gave little basis for confidence in rehabilitative programs, did not appear until 1974. The National Academy of Sciences report on deterrence and incapacitation, whose conclusions Morris anticipated, did not appear until 1978 (Blumstein, Cohen, and Nagin 1978).
B. Distribution

Whatever the applicable normative purposes of a sentencing system may be, the principal distributive aims are consistency, evenhandedness, and fairness relative to those purposes. Imposition of punishment for crime is a paradigmatic instance of the conflict between individual liberty and collective interests. Sentencing is a process in which government actors are empowered to intrude on individuals' liberty and autonomy in order to prevent crimes and sanction wrongdoing. To be consistent with core values underlying constitutional notions of due process and equal protection, sentencing needs to respect offenders' rights to be treated as equals, to have decisions about them made deliberately and impartially, and to be dealt with under fair procedures.

I would be surprised if many people disagree. A major challenge for sentencing policy, however, is that consistency, evenhandedness, and fairness look different when viewed from different places. To legislators, a ten-year mandatory minimum sentence, three-strikes, or 100-to-one law is not necessarily inconsistent with those values. Nor, presumably, to members of the U.S. Sentencing Commission were the "actual-offense behavior" policies, drug-trafficking guidelines, or a decision to forbid judges to take account in sentencing of ethically important differences in individual offenders' circumstances. From the perspective of policy making at the wholesale level, the policies were for Congress and the commission to set and for judges and prosecutors, equably, evenhandedly, and fairly, to implement.

If literally applied in every case, however, those policies would have made practitioners complicit in imposing sentences in many cases that everyone directly involved believed to be substantively or procedurally unjust. Practitioners see cases firsthand. Problems appear simpler the farther away the point from which they are observed. District court judges, whether liberal or conservative, or appointed by Republican or Democratic administrations, for example, who see living, breathing defendants, are more hostile to the federal guidelines than are appellate judges who see only written appellate briefs. The same generalization applies to laypeople. The more detailed the information people are given about particular criminal cases, the less harsh and stereotyped

From the outset, the guidelines provided that such commonsense considerations as family responsibilities, drug dependence, and mental health problems were not "ordinarily relevant in determining whether the sentence should be outside the guidelines" (U.S. Sentencing Commission 1987, secs. 5H1.1–6).
the sentences they would impose become. As a result, as surveys in a number of countries show, the sentences laypeople would impose for particular kinds of cases are often less severe than the sentences judges do impose (Roberts et al. 2002).

Legislators and senior executive branch officials, by contrast, deal in stereotypes. Often, like citizens responding to survey questions, they think in terms of extreme cases, even when setting policies that will mostly be applied to run-of-the-mill ones. They are also more likely to think of sentencing policy in terms of its latent functions of self-interest, ideology, and partisanship, considerations that are much less likely to move practitioners when pondering an individual's fate. Nearly all judges and prosecutors—"officers of the court" in American legal systems—take their ethical obligation to do justice seriously. Policies that substantially obstruct their ability to do justice are likely to be resisted or undermined.

The federal guidelines often directed judges to impose sentences they believed to be unjustly severe (Zlotnick 2004). In such circumstances, judges had three options: compliance, resignation, and circumvention. They could apply the guidelines as written, but at the cost of behaving unjustly; most did this much of the time, occasionally after announcing that they felt they had no alternative but to impose an unjust sentence. They could announce that the guidelines required them to impose an unjust sentence and, refusing to do so, resign. A few did this. They could resolve neither to impose an unjust sentence nor to defy the law openly nor to resign, but together with counsel circumvent the guidelines to produce a sentence everyone involved considered appropriate. Many did this, and increasing numbers did so as time passed (Bowman and Heise 2001).

J. Lawrence Irving, the first federal district court judge to resign in protest over the federal guidelines, was appointed by President Ronald Reagan. On resigning in 1990, he said: "If I remain on the bench I have no choice but to follow the law. I just can't, in good conscience, continue to do this" (New York Times 1990, p. 2). Federal district court judge John S. Martin resigned in 2003 over what he called "a sentencing system that is unnecessarily cruel and rigid" (Martin 2004, p.

---

16 A U.S. Sentencing Commission survey in 2003 showed that 73.7 percent of district judges and 82.7 percent of circuit judges said that federal drug-trafficking sentences are inappropriately severe, which necessarily means that many judges appointed by Presidents Reagan, Bush I, and Bush II held this view (U.S. Sentencing Commission 2003, exhibits II-4, III-4).
Richard Boylan’s statistical analyses concluded that the guidelines’ unpopularity led judges on average to retire earlier than they otherwise would have (Boylan 2004).

My personal experience working with judges and prosecutors is that most practitioners most of the time want to do justice in individual cases, to do the right thing, to achieve a just and appropriate sentence. When applicable laws make that difficult or impossible, they still feel powerfully moved to do so anyway.

This is human nature, and there is a lot of evidence for it, going back centuries. Between 1714 and 1830, the British Parliament created 156 new capital crimes punishable by death, many of them property offenses, but the number of executions fell by three-fourths. Why? Because juries refused to convict offenders or convicted them for offenses not punishable by death, and judges created and narrowly interpreted technical rules that had to be followed if death was to be imposed, but seldom were (Stephen [1883] 1977; Select Committee on Capital Punishment 1930; Hay 1975).

The American Bar Foundation in the 1950s and 1960s carried out major studies of criminal court operations in several states (Newman 1966; Dawson 1969; Miller 1969). Prosecutorial and judicial circumvention of mandatory sentences was common in each court studied. Exemplifying this, Donald Newman described the application in Michigan of a fifteen-year maximum sentence for nighttime burglary (compared with a five-year maximum for daytime burglary): “The frequency of altering nighttime burglary to breaking and entering in the daytime led one prosecutor to remark: ‘You’d think all our burglaries occur at high noon’” (Newman 1966, p. 182). The maximum mattered so much then and there because, taking time off for good behavior into account, it set the date by which the parole board had to release a prisoner. A lower maximum was insurance against a much longer stay in prison.

As Frank Remington, executive director of the nearly two-decades-long series of projects, summarized, “Legislative prescription of a high mandatory sentence for certain offenders is likely to result in a reduction of charges at the prosecution stage, or if this is not done, by a refusal of the judge to convict at the adjudication stage. The issue . . . thus is not solely whether certain offenders should be dealt with severely, but also how the criminal justice system will accommodate to the legislative change” (Remington 1969, p. xvii).

Evaluations of the effects of mandatory minimums enacted in the
Purposes and Functions

1970s in Massachusetts (Rossman et al. 1979), Michigan (Heumann and Loftin 1979), and New York (Joint Committee 1978) reached the same conclusions. So did an evaluation of Oregon’s Measure 11, enacted by referendum in 1994 and requiring lengthy mandatory sentences for people convicted of sixteen serious crimes (Merritt, Fain, and Turner 2003, 2006). An evaluation of New Jersey’s 1994 “No Early Release Act,” requiring offenders to serve 85 percent of their sentences before release, hypothesized that changes in plea bargaining would largely nullify the act and result in times served little different from before the law took effect. That is what was found (McCoy and McManimon 2003).

Experience with the federal guidelines was the same. During their first decade, a host of studies, some sponsored by the U.S. Sentencing Commission, showed that circumvention of the guidelines’ harshest and most rigid provisions was commonplace (Nagel and Schulhofer 1992; Schulhofer and Nagel 1997). In recent years, Frank Bowman and Michael Heise (2001) showed that a wide range of circumvention techniques by prosecutors and judges resulted throughout the 1990s in continuously decreasing average sentence lengths in the federal courts.

The consistent evidence that practitioners often circumvent sentencing laws mandating sentences they believe to be unjust should not be surprising. James Eisenstein and his colleagues, reporting on research on how practitioners in felony courts in nine counties reached sentencing decisions, observe that sentences result from “accommodations among competing values and interests that support these values, accommodations that are superimposed on a common basic structure supported by broad consensus.” Of reforms imposed from outside, they conclude that “the more radical a proposed change, the less likely is its adoption” (Eisenstein, Flemming, and Nardulli 1988, p. 294).

C. Prevention

The preventive functions focus on crime, fear of crime, and reduction in their costs and consequences. For purposes of this essay, I assume that efforts to prevent crime will also lessen fear and reduce crime’s costs and consequences; that is a reasonable assumption, even though a finer-tuned analysis would discuss different policy mechanisms for achieving each outcome.\(^{17}\)

\(^{17}\) A harm-reduction approach to drug policy, e.g., might imply different drug policy
The preventive functions center on the collective interest in domestic tranquility, on enabling citizens to get on with their lives in security. These are core functions of the state and basic goals of criminal law and punishment. If they are to be pursued effectively, policy makers need to take account of the considerable body of knowledge on the effectiveness of sanctions.

Precisely how crime prevention, fear reduction, and cost savings should be sought involves policy decisions about which reasonable people differ and depends in part on what normative purposes are deemed applicable. I assume though that most policy makers regard crime prevention as among the justifiable normative purposes even if some sets of normative purposes that might be adopted might constrain the policy options available to pursue them. That being so, knowledge of what works, and what doesn’t, is essential.

Writing on this subject is a ticklish business because it is probably generally assumed that the writer has axes to grind that shape how the evidence is summarized. Perhaps surprisingly, however, there is broad agreement among academics of liberal (e.g., Ruth and Reitz 2003) and conservative (e.g., Wilson 2002) political inclinations about conclusions to be drawn from research on the preventive effects of criminal sanctions and criminal justice interventions.

1. Deterrence. Current knowledge concerning deterrence is little different than eighteenth-century theorists such as Beccaria ([1764] 1995) supposed it to be: certainty and promptness of punishment are more powerful deterrents than severity. This does not mean that punishments do not deter. No one doubts that having a system of punishment has crime-preventive effects. The important question is whether changes in punishments have marginal deterrent effects, that is, whether a new policy causes crime rates to fall from whatever level they would otherwise have been at. Modern deterrent strategies, through sentencing law changes, take two forms: increases in punishments for particular offenses and mandatory minimum sentence (including “three-strikes”) laws.

Imaginable increases in severity of punishments do not yield significant (if any) marginal deterrent effects. Three National Academy of...
Purposes and Functions

Sciences panels, all appointed by Republican presidents, reached that conclusion (Blumstein, Cohen, and Nagin 1978; Blumstein, Cohen, Roth, and Visher 1986; Reiss and Roth 1993), as has every major survey of the evidence (Cook 1980; Nagin 1998, 1999; von Hirsch et al. 1999; Doob and Webster 2003). This is also the belief, in my experience, of most experienced judges and prosecutors.

There are a number of good practical reasons why this widely reached conclusion makes sense. First, serious sexual and violent crimes are generally committed under circumstances of extreme emotion, often exacerbated by the influence of alcohol or drugs. Detached reflection on possible penalties or recent changes in penalties seldom if ever occurs in such circumstances. Second, most minor and middling and many serious crimes do not result in arrests or prosecutions; most offenders committing them, naively but realistically, do not expect to be caught. Third, those who are caught and prosecuted almost always are offered plea bargains that break the link between the crime and the prescribed punishment. Fourth, when penalties are especially severe, they are often, albeit inconsistently, circumvented by prosecutors and judges. Fifth, for many crimes including drug trafficking, prostitution, and much gang-related activity, removing individual offenders does not alter the structural circumstances conducing to the crime. Sixth, even when one ignores all those considerations, the idea that increased penalties have sizable marginal deterrent effects requires heroic and unrealistic assumptions about “threat communication,” the process by which would-be offenders learn that penalty increases have been legislated or are being implemented.

Mandatory minimum penalties including three-strikes laws are nothing more than efforts to deter crime through penalty increases. The clear weight of the evidence, not surprisingly given what we know about severity increases generally, is that they are seldom if ever crime preventatives (e.g., Tonry 1996, chap. 4; Zimring, Hawkins, and Kamin 2001). Besides the not inconsiderable problem that mandatory minimums often are circumvented by practitioners, and always will be, the best evidence suggests that they have no marginal deterrent effects or have only modest effects that quickly waste away.

There is no good evidence that justifies fine-tuning sentences in individual cases, or enacting mandatory minimum sentence laws, for deterrent reasons. Having a system of punishments has deterrent effects. Having punishments scaled to the severity of offenses should
provide rational incentives, as Beccaria and Bentham urged; to commit fewer rather than more serious crimes and, as Durkheim urged, to reinforce basic social norms.

The principal implications of current knowledge about deterrence are that certainty of punishment is more important than severity, that penalties that are too rigid or too severe result in widespread circumvention, and accordingly that mandatory minimums and three-strikes laws are ineffective, cause foreseeable injustices including gross disparities, and should be repealed.

2. Incapacitation. Incapacitative crime control strategies have some potential if used discriminatingly for people who commit very serious crimes at high rates (Monahan 2004). Offenders in confinement necessarily are disabled from committing offenses in the free community. We know a good bit more about incapacitation than we did thirty years ago. For a time, policies referred to as “collective incapacitation” were in vogue, influenced or at least evidenced by proposals in James Q. Wilson’s Thinking about Crime (1976): extend all sentences of people convicted of violent crimes such as robbery by several years. This was soon shown to be impracticable. Large numbers of people are convicted of such crimes, but the average likelihood of recidivism is too low for the resulting huge increases in imprisonment to be justified (Blumstein, Cohen, and Nagin 1978; Cohen 1983).

This was followed by proposals for “selective incapacitation,” catalyzed by a RAND Corporation report purporting to show the feasibility of such policies (Greenwood and Abrahamse 1982). This was soon discredited by a National Academy of Sciences report that showed that RAND’s proposals worked with hindsight—it was based on self-reported past crimes of imprisoned violent offenders and in retrospect seemed to characterize those prisoners—but not prospectively (Blumstein et al. 1986). The prediction methods proposed were so over-inclusive, producing so many “false positives,” that selective incapacitation was dismissed as impracticable.

After that, the theoretical writing and related empirical research on incapacitation pretty much came to a halt (Nagin 1998), but policy makers kept increasing lengths of sentences and accepted the imprisonment numbers that in the 1970s and early 1980s appeared unimaginary (Blumstein 1988). Collective incapacitation, in effect, became national policy, though not under that name. Studies in the late 1990s (Spelman 2000) and early in this century (Blumstein and Wallman
2000) concluded that mass imprisonment had reduced crime rates by 10–25 percent.

That last conclusion needs to be taken with a grain of salt for two reasons. First, comparable declines in crime rates occurred throughout the Western world, and no country other than the United States quintupled its imprisonment rate and experimented with mass imprisonment (Tonry 2004b, chap. 2). There is a reasonably good chance that American crime rates would have fallen like every other Western country’s. The contrast with Canada—a flat imprisonment rate of 100 per 100,000 population for forty years but crime rate trends closely paralleling those in the United States (Doob and Webster 2006)—is especially cautionary. Second, and no less important, the studies concluding that incapacitation caused significant declines in crime rates are based on exceedingly complex modeling, which like all such modeling is highly vulnerable to the assumptions on which it is based and the modeling technique used; even highly sophisticated modelers seldom agree about either (viz., compare Webster, Doob, and Zimring [2006] with Kessler and Levitt [1999] and Rosenfeld, Fornango, and Baumer [2005a, 2005b] with Berk [2005]).

Incapacitative considerations have little legitimate place in decisions about individual offenders. The principal difficulty is that current incapacitation policies often, and expensively, target the wrong people. There are five main problems. First, incapacitation strategies targeted on repeat serious offenders generally result in long sentences for older offenders, most of whom in any case would soon desist from crime. This is known as the problem of residual career length; most thirty- or forty-year-old offenders are unlikely to continue as active offenders for long. There may be other justifications for confining such offenders for extended periods; but from an incapacitative perspective, such sentences are expensive, inefficient, and largely ineffective.

Second, strategies that target repeat offenders, in earlier times generally called “habitual-offender” laws and now called “three-strikes” laws, tend to ensnare socially inadequate, high-volume property offenders for whose prevented crimes lengthy prison sentences costing $40,000–$50,000 per year are not a sensible investment of public resources.

Third, for many offenses, including particularly drug sales, drug trafficking, and some gang-related crime, there is the “replacement” problem: offenders taken off the streets are quickly replaced by willing suc-
cessors. Removing a seventeen-year-old drug dealer from an inner-city street corner is seldom likely to stop drug dealing at that corner; plenty of willing volunteers are usually waiting in the wings to step into what looks to be a lucrative opportunity. Similarly, unless the gang itself is completely destroyed, locking up one or several of its members will not stop its illegal activities. Other, and new, members will step in to replace their fallen comrades.

Fourth, there is the “false-positive” problem. For crimes of any significant seriousness, we are not very good at predicting which offenders will commit them in the future. For every “true positive” who will, three or four others predicted to do so will not. From a cost-benefit perspective, locking up all those people is not an obviously good investment of public resources. From a moral perspective, many people are troubled by the thought of lengthy confinement of individuals not because of their current crime but because they might commit another, especially when we know most will not.

Fifth, there is the “true-but-trifling-positive” problem. We are pretty good at predicting petty reoffending. David Farrington (1979) long ago demonstrated that an offender who has been arrested nine times has an 80 percent likelihood of being arrested again. This raises the cost-benefit question already mentioned: Is it worth $40,000–$50,000 per year to prevent shoplifting, minor thefts, acts of prostitution, or low-level drug dealing? It also raises the moral question whether the disproportionate character of a multiyear prison sentence for a person convicted of such crimes can be justified.

As with deterrence, the answer is not to forgo efforts to incapacitate offenders, but to do so intelligently. This means working hard at the policy level to develop strategies targeting particular kinds of serious offenders whose future behavior we can reasonably confidently predict and whose averted crimes save substantial suffering. At the individual level, however, it means limiting judges’ authority to predicate sentencing decisions on incapacitative ambitions.

3. Rehabilitation. Prevention through rehabilitation looks to be a considerably more viable strategy in the early twenty-first century than it did during the closing decades of the twentieth. The view that “nothing works” was an important backdrop to the shift toward determinate sentencing, the abolition of parole, and adoption of incapacitative and deterrent crime control strategies. If we do not know how to reduce
offenders' prospects for later offending, it is hard to justify giving judges and parole boards broad discretion to individualize sentencing.

The prospects for rehabilitation, however, have changed radically. Evidence is accumulating from many sources—individual evaluations, meta-analyses (e.g., Lipsey and Landenberger 2006; Mitchell, MacKenzie, and Wilson 2006), literature reviews (e.g., Gaes et al. 1999; U.S. Surgeon General 2001), and practical experience—that well-managed, well-targeted programs can reduce participants' probability of reoffending. A wide range of programs, including drug treatment, anger management, cognitive-skills programs, sex offender treatment, and various educational- and vocational-skills programs have been shown to reduce reoffending (Gaes et al. 1999). A report from the English Home Office, which underpinned a massive reorganization of the English criminal justice system mandated by the Criminal Justice Act of 2003, concluded that "a reasonable estimate at this stage is that, if the [treatment] programmes are developed and applied as intended, to the maximum extent possible, reconviction rates might be reduced by 5–25 percentage points (i.e. from the present level of 56 percent within two years to (perhaps) 40 percent)" (Halliday 2001, p. 7). The most recent meta-analysis of the effects of cognitive-behavioral programs concluded that, on average, they reduced reoffending by 27 percent (Lipsey and Landenberger 2006). The proliferation of drug courts and prisoner reentry programs in the United States bears witness to the widely shared perception that some things work.

That litany of positive findings does not mean that reducing reoffending rates is easy. The results obtained from a well-funded pilot project, led by motivated people, cannot automatically be obtained by institutionalizing a new program model throughout a jurisdiction. A recent survey of violence prevention programs by the U.S. Surgeon General (2001) concluded that many programs can reduce violent offending but that the challenge is broad-based implementation. This is nonetheless much better news than the state of the evidence twenty-five years ago. We now know what we do; we need to figure out how to do it on a large scale.

An important implication is that rigid sentencing policies obstruct efforts to prevent crime through rehabilitation of offenders. For drug and other treatment programs to work, they must be targeted to the characteristics and needs of particular offenders, and this requires sen-
sentences to be individualized. With the fall of the nothing works psychology goes much of the case for rigid sentencing standards.

4. Moral Education. Deterrence, incapacitation, and rehabilitation are not needed to restrain most adults from selling drugs, burglarizing houses, holding up convenience stores, or mugging passersby. Most people's personal norms and values make predatory crime almost unthinkable. European scholars and theorists have long observed that the criminal law's main function is "general prevention": reinforcement of basic social norms that are learned in the home, the church, the school, and the neighborhood (Lappi-Seppälä 2001). These primary socializing institutions must do the heavy lifting—what the criminal courts can do is too little and too late for the criminal justice system to serve as a primary socializing institution—but it is important that law and the legal system reinforce those norms and not undermine them.

One implication, consistent with research findings on deterrence, is that it is more important that crimes have consequences than that the consequences be severe. Another, consistent with the research findings on rehabilitation, is that the choice of sanctions should be tailored to the circumstances of particular cases.

The evidence on the preventive effects of punishment all points in the same direction. A sentencing system that sets firm upper limits on punishments, scaled to offenders' culpability or the severity of their crimes, but allows substantial individualization below those limits, promises greater preventive effectiveness than current laws in the federal system or most states.

III. Ancillary Functions
The criminal justice system has limited material and manpower resources. Prosecutors, presiding judges, and corrections managers need to organize their offices, courtrooms, and programs to get the job done as efficiently and cost-effectively as possible while trying to achieve appropriate outcomes in individual cases.

Sentencing guidelines in some states have proved to be useful tools in reconciling distributive, preventive, and management functions. These guidelines have achieved reasonable consistency in sentences imposed and have proved an effective tool for projecting the effects of alternate policy proposals and thereby managing corrections resources and controlling corrections budgets at a systems level. Although plea
bargaining occurs everywhere, in the states with well-designed guidelines, bargaining takes place in the shadow of the guidelines and has not been characterized by the widespread circumvention, and resulting disparities, that characterized the federal system (Tonry 1996; Reitz 2001, 2005a; Frase 2005c).

A. Efficiency and Cost-Effectiveness

Court dockets are crowded, treatment resources are scarce, and correctional programs including prisons are overcrowded and underresourced. Managers have to figure out how to juggle their budgets and manpower to keep cases flowing, hold backlogs down, and prioritize their resource allocations so that the most serious and important cases receive the attention they warrant. Most courts and prosecutors' offices have monthly case disposition and backlog targets; failure to meet them is interpreted as laziness or ineffectiveness. This is generally understood to mean that most cases must be resolved as quickly as possible. Trials should be avoided whenever possible, and bench trials are preferable to jury trials.

Plea bargaining is the mechanism for achieving most of those goals and requires that the state have something to offer the defendant to induce cooperation and that prosecutors, defense counsel, and judges get along. Generally, the parties do get along because many criminal court practitioners work for extended periods in the same setting and soon learn the prevailing conventions ("going rates") and get to know the relatively small number of people in the courtroom in which they work.

Peter Nardulli and his collaborators in the most extensive study to date of courtroom communities put it like this: "The web of interpersonal ties and mutual need that holds the [courtroom community] together also creates norms and understandings that guide and restrict behavior. In a sense court communities emerge because of the need to establish these understandings, as well as to meet other human needs . . . . It is in the interest of all to establish, abide by, and enforce a common set of understandings about how normal cases should be handled and how defendants should be punished" (Nardulli, Eisenstein, and Flemming 1988, p. 306).

People who are oppositional or uncooperative soon find that they receive little cooperation back and that they are felt by their colleagues to be unreasonable. Those are pretty effective sanctions for uncoop-
operative behavior: "The importance attached to conflict avoidance and conflict management is another important component of the court community's court culture. In some communities, a great deal of emphasis was placed upon going along and getting along, upon not rocking the boat" (Nardulli, Eisenstein, and Flemming 1988, p. 125).

Lawyers and judges become socialized into the local courtroom culture, which, given that felony courts are usually staffed by people from the communities they serve, generally reflects prevailing community notions about right and wrong and the severity of punishment. People working in big-city courts tend to have somewhat different local cultures than people working in suburban and small-town courts. Legal cultures vary from city to city and region to region (Eisenstein, Flemming, and Nardulli 1988). A U.S. Sentencing Commission–sponsored study of prosecutorial circumvention of guidelines documented wide variations in legal culture in different federal district courts (Nagel and Schulhofer 1992).

What practitioners do largely depends on the prevailing courtroom culture. If sentencing statutes and guidelines set standards that are incompatible with local traditions and beliefs, local traditions often win out, with the acquiescence of everyone involved. Practitioners' needs to achieve management goals and the importance of local courtroom cultures together mean that local courts will do what is needed to operate efficiently and cost-effectively and to achieve outcomes that those involved consider appropriate and just. Sentencing standards that are too harsh, too lenient, or too mechanistic often will be circumvented.

B. Resource Management

Governments and individual agencies need to get their business done within the resource constraints presented by their budgets. They need to be able to predict changes in their likely flow of work and to plan for resources they will need in the future. These things inevitably entail choices.

One of the great advantages of sentencing guidelines is that they are a proven tool for effective resource management (Boerner and Lieb 2001; Wright 2002; Frase 2005b). States that adopted sentencing guidelines as a resource management tool have successfully tailored their sentencing policies to their correctional budgets and programs. By making sentencing decisions predictable in the aggregate, guidelines
enable policy makers to project likely future resource needs. If, for example, a jurisdiction contemplates doubling sentence lengths for sex crimes, it can project how many more prison beds will be needed. If current and planned facilities cannot accommodate the increased numbers, policy makers can respond in a number of ways. They can appropriate funds to build new facilities, revise sentences for other offenses downward to free up the needed spaces, reconsider whether the sex offender proposal is a good idea, or do some combination of these things.

Sound resource management is a necessary characteristic of effective government and a primary rationale for sentencing guidelines systems. Its scope should extend to needs for nonincarcerative sanctions, availability of appropriately targeted rehabilitative programs, and effectuation of appropriately targeted incapacitation policies.

C. Communication

Communicative functions are described briefly, not because they are less important than the other functions but because the basic propositions can be simply stated. Communicative functions interact with normative, distributive, and preventive functions. Legitimacy and public confidence depend in part on whether the law is seen to operate reasonably and in ways that are consonant with widely held ideas about justice (including, e.g., that claims be fairly considered and that penalties be proportionate to the severity of crimes).

1. Legitimacy. Legitimacy is increasingly recognized as a key element in how people relate to government (Tyler 2003). Whether people have confidence in, support, and cooperate with public institutions is influenced by the institution’s legitimacy in their eyes. People are more likely to react positively when the police stop them, or to accept an adverse decision by a court, if they believe that they have been treated fairly, that decisions about them have been evenhanded and impartial, and that they have had a chance to have their say. Sentencing policies that respect offenders’ rights to be treated as equals, to have decisions about them made deliberately and impartially, and to be dealt with under fair procedures are more likely to be perceived as legitimate than policies that do not satisfy these requirements of procedural justice.

2. Public Reassurance. Sometimes laws are enacted or policies are adopted to “send a message” that the public’s views have been noted
or that their anxiety, apprehension, or unhappiness has been noted. Sociologist David Garland (2001) argues that much recent crime control policy in the United States and England was adopted primarily for "expressive" reasons, not because policy makers necessarily thought that new, tougher policies would reduce crime rates but because they wanted to reassure an anxious public that was fearful of rising crime rates and troubled by the uncertainties of a rapidly changing world.

Legislators have long adopted symbolic policies meant primarily to acknowledge public anxiety. Senator Alfonse D'Amato of New York, for example, according to the New York Times, conceded that two amendments he successfully offered to federal crime legislation, "which Justice Department officials said would have little practical effect on the prosecution of crimes, might not solve the problem. 'But,' he said, 'it does bring about a sense that we're serious'" (Ifill 1991, p. A6). The seeming eighteenth-century English paradox that many more crimes were made punishable by death during a century in which the number of executions rapidly and substantially declined is generally explained in expressive terms; legislators were trying to stress the seriousness of the new death-eligible offenses, generally property crimes, even though there was little appetite for executing property offenders (Hay 1975).

3. **Public Confidence.** For nearly a decade the English government has explained major changes in criminal justice policy as parts of a larger effort to increase public confidence in the legal system (e.g., Home Office 2002). The criminal justice system is believed by the English government to lack credibility in the eyes of voters, and measures meant to make punishments harsher and more certain and convictions more likely are portrayed as parts of a larger campaign to increase public confidence in the legal system and generally. This is in effect a collective, bystander’s parallel to legitimacy. Public confidence in the legal system is likely to increase if it is seen to be addressing problems that trouble citizens effectively.

4. **Reinforcement of Basic Social Norms.** This is the corollary of the preventive function of "moral education." That is why Scandinavian countries attach great significance to proportionality in punishment (so that the severity of sanctions acknowledges the comparative seriousness of crimes) and more frequently impose prison sentences (though shorter ones) and their equivalents than most other countries do (so criminal behavior has consequences) (Lappi-Seppälä 2001).

Conversely, if the law’s treatment of crimes is incompatible with
basic social norms, the effect will be to undermine the legitimacy of
the legal system. At times, for example, when cannabis is treated as a
dangerous drug on par with heroin or crack cocaine, enforcement be-
comes lax and inconsistent, and the legal system is seen by many as
being out of step with prevailing behavioral norms. Or when drug
trafficking in small amounts is treated as being more serious than rape,
manslaughter, or serious assault, it is clear that the legislature is out of
step. Recent prosecutions and severe punishments in white-collar crime
cases offer an example of the system responding to a widespread public
perception that it was unjust that corporate malefactors in crimes in-
volving hundreds of millions of dollars often received milder punish-
ments than less privileged citizens convicted of much less serious
crimes.

A few strong inferences can be drawn from these observations about
communicative functions. First, they fit together: people expect legal
institutions to be fair, consistent, rational, and in step with prevailing
social norms. Second, public confidence and reassurance are likelier to
be achieved by fair, consistent, and rational policies and practices than
by expressive and symbolic policies that lack those qualities. It is hard
to imagine, for example, that statutes that call for five-, ten-, and
twenty-year minimum sentences for drug crimes, which are routinely
circumvented and thereby lack legitimacy and fail to mirror prevailing
social norms, have greater positive effects on public reassurance and
confidence than more moderate laws would.

IV. Latent Functions
The latent functions of sentencing and sentencing policy—using sen-
tencing to achieve personal, ideological, or partisan political goals—
sometimes fundamentally obstruct performance of sentencing’s pri-
mary and ancillary functions. Many practitioners and policy makers
hanker for more lucrative or powerful jobs. Elected officials want to
be reelected and aspirants want to be elected for the first time. Pros-
ectors and judges sometimes call for or impose unusually severe pun-
ishments to win favor with voters or party leaders. Legislators propose
and vote for new laws for those reasons. Many practitioners and policy
makers have strong ideological beliefs and want to bear witness to their
beliefs. Policy makers sometimes want to posture before ideological
soul mates and make symbolic statements. Policy makers sometimes
subordinate their personal beliefs to achievement of partisan political goals.

A. Self-Interest

Purely personal ambitions and self-serving motives are illegitimate considerations in decision making by prosecutors and judges about individual cases. Performance of the distributive, preventive, and communicative functions is undermined when cases are dealt with in a particular way because a practitioner believes she is likelier to gain a nomination, win an election, or obtain a new job. Retributive theories of punishment call for punishments to reflect the personal culpability and harm associated with an offense. Utilitarian and mixed theories call for punishments to be based at least in part on good-faith assessments of their likely crime-preventive effects. No theory justifies punishments based on substantively unrelated ulterior considerations. Decisions based on practitioners' personal self-interest are in effect whimsical or arbitrary and irreconcilable with any prevailing theory of punishment.

The analysis is somewhat more complicated for legislators and other policy makers, primarily because the issues of self-interest are sometimes more opaque. When a prosecutor or judge treats an offender in a particular way to realize a purely personal goal, there is no arguable public benefit that can be said to outweigh or counterbalance the harm done the offender. When an entirely self-motivated legislator votes to support a particular policy for which there is little substantive justification, but claims in good faith to be trying to reassure the public or, more complicatedly, claims a need to vote in a particular way on this subject in order to win others' support for a vote on a more important subject, it is almost impossible to assess what the real motives are. The governing ethical premise, however, should be the same as for a prosecutor or a judge: only disinterested motives are legitimate. Observers may have difficulty knowing policy makers' motives, but they themselves know whether their motives are legitimate or ignoble. A good ethical litmus test is whether a legislator at a press conference could comfortably and completely truthfully explain his or her motives.

In the introduction I described a judge up for a vigorously contested reelection who in a particular case, to win favorable publicity, might impose a much harsher sentence than in his own mind can be justified, and I suggested that few observers who knew why he had done so
would approve. The analysis of policy making predicated on pursuit of instrumental latent goals is the same as that which applies to a judge who bases the sentence he imposes on his electoral hopes. A legislator or governor who proposes or enacts policy changes he knows will not achieve their purported aims and will, if enacted, cause new injustices, because he hopes his support for the new policy will help him get reelected, is behaving unethically.

The entanglement of sentencing policy in individuals' self-interest has been especially acute at the federal level. Repeatedly in this essay I have used examples from the federal sentencing guidelines. A major reason why they were so inadequate and disliked was that the U.S. Sentencing Commission's policy making in its formative years became entangled in the personal ambitions of William Wilkins, the first chair (initially, unsuccessfully, to become FBI director, then successfully to become a federal court of appeals judge), and Stephen Breyer (successfully, to get to the Supreme Court). Wilkins had been a senior staffer to Senator Strom Thurmond and Breyer had been chief of staff to Senator Edward Kennedy when the two senators were successively chairmen of the Senate Judiciary Committee. Breyer's cross-aisle political credibility was such that Senate confirmation of his appointment as a federal court of appeals judge occurred in October 1980, just before the presidential election, at a time when Republicans otherwise were refusing to act on President Carter's judicial nominations. Wilkins's and Breyer's judicial ambitions meant that a major function of the guidelines was to win favor with Strom Thurmond and his personal staff in the Senate in particular, with congressional Republicans more generally, and with Ed Meese's Justice Department (von Hirsch 1988; Tonry 1996, p. 84). Tens of thousands of federal defendants paid a high price for Wilkins's and Breyer's judicial appointments.

Sometimes penal policies are proposed for reasons having nothing to do with the goals of punishment per se. The prison guards union in California, for example, is often said to promote harsher policies as means to the ends of job creation and maintenance (Petersililia 2003, pp. 237–38). Many communities have sought placement of prisons within their boundaries and, implicitly, increased use of imprisonment, as a form of economic development (p. 239). On the face of it these considerations are entirely unrelated to punishment and prevention.

All the things I have described and decried happen and no doubt will continue to. Anyone who reads newspapers knows that. That is
the world we will continue to live in, but practitioners and policy makers can be more self-aware, intellectually honest, and transparent, and the rest of us can try to hold them to account.

B. Ideological Expression

Crime control policy in our time has become entangled in ideological conflict. Recent conflicts over medical use of marijuana and Oregon's assisted-suicide law offer front-page examples. Voters initially in California and eventually in more than a half dozen states approved referenda authorizing medical use of marijuana. Voters in Oregon, initially by a narrow margin and later by a wide margin, approved a law allowing physicians in narrowly defined circumstances to provide drugs to mentally competent, terminally ill patients who want to die. In general under American constitutional doctrines of federalism, criminal justice policy is within the purview of the states (Richman, in this volume).

Public health and the regulation of medical practice also are generally recognized as matters of state law, but that has not stopped ideologically motivated federal officials from trying to control state policies. Successive U.S. Attorneys General John Ashcroft and Alberto Gonzales have attempted to use tortured readings of federal drug laws to overrule the decisions by Oregon voters to allow doctors to assist suicides and by California voters to allow doctors to prescribe marijuana for health care reasons. In each case, the federal government threatened to act under federal drug laws against doctors who complied with the state laws. The U.S. Supreme Court repudiated the effort in the Oregon case, in Gonzales v. Oregon (126 S. Ct. 904 [2006]), but, in Gonzales v. Raich (125 Supreme Court 2195 [2005]), supported the federal government's position concerning medical use of marijuana.

Ideology is at least as powerful an influence on sentencing policy. Drug policy offers stark examples. The federal 100-to-one law punishes people, mostly black, convicted of crack cocaine offenses as severely as people, many white, convicted of powder cocaine offenses involving amounts that are 100 times larger. The law was enacted in 1986, after the much-publicized death, generally attributed to a crack overdose, of Len Bias, a University of Maryland basketball player forecast to become a National Basketball Association superstar (Kennedy 1997). The law's role in exacerbating racial disparities in federal prisons soon became clear (McDonald and Carlson 1993), and the U.S. Sentencing
Commission proposed that the differential be eliminated. Both the Clinton White House and the Congress opposed any change, and none was made. In later years Attorney General Janet Reno, Drug Czar Barry McCaffrey, and the U.S. Sentencing Commission proposed that the differential be reduced. The Clinton and Bush II White Houses opposed all changes, and the differential remains. No one presumably wants federal sentencing laws to worsen racial disparities, but neither successive administration nor congressional leaders have been willing to risk being accused of condoning drug use or trafficking (Tonry 2004b, chap. 1).

Mandatory minimum and three-strikes laws, for other examples, often are primarily based on ideology. Research findings discussed in earlier sections make it clear, and for at least a century have made it clear, that such laws seldom achieve their putative goals and always produce undesirable and unwanted consequences. Ideological posturing is no better an explanation for why an offender is punished unjustly than is a judge's hope for reelection.

C. Partisan Advantage

Pursuit of partisan advantage is the most cynical of the latent functions. Ideological influences grow out of deeply held beliefs; those beliefs may sometimes be blinding. Self-interested acts can, of course, be criticized, but as long as men aren't angels they will remain at least psychologically understandable. Decisions made for partisan reasons, by contrast, are made in cold calculation. Partisan influences often result in the passage of laws that cannot be justified on the substantive merits and foreseeably produce unjust results. Republican Senator Orrin Hatch, according to a member of his staff, for example, long believed reducing sentences for drug offenders "was the right thing to do, but he couldn't do it for political reasons" (Gest 2001, p. 214).

California's three-strikes law, according to Frank Zimring's account, resulted from politicians' competing attempts to use punishment policies to pursue partisan advantage. It was enacted not because thoughtful policy makers really believed that people who stole pizza slices in schoolyards or handfuls of compact discs from Wal-Mart deserved decades-long prison sentences, but because Republican Governor Pete Wilson and California Assembly leader Willie Brown played a game of chicken in which, in the end, neither backed down. Democratic legislators agreed among themselves to pass any proposal Governor
Wilson offered, in hopes "that he would back down from an unquali-
fied 'get tough' stand or be politically neutralized if he persisted." Wil-
son did not blink. Nor did the Democrats. The law was passed as 
proposed because both sides were "unwilling to concede the ground 
on 'getting tough' to the other side in the political campaign to come." 
Neither Wilson nor Brown was willing to propose refinements to Wil-
son's extreme initial proposal and thereby expose himself and his party 
to the other's accusation of softness. As a result California adopted the 
most far-reaching, rigid, and unjust three-strikes law in the country 

Former president Bill Clinton famously decided in the 1980s never 
to let the Republicans get to his right on controversial crime control 
issues, no matter how unjust, wasteful, or ineffective the policies under 
consideration (Edsall and Edsall 1991; Friedman 1993). Clinton him-
self in 1994 admitted that "I can be nicked on a lot but nobody can 
say I'm soft on crime" (Tonry 2004b, p. 8). It worked for Clinton but 
not for the tens of thousands of people whose lives were changed for 
the worse.

Britain's Labour Party, after a trip to Washington by Tony Blair to 
meet with Clinton's political advisors, did the same thing in 1993 
(Downes and Morgan 2002). Since then England has adopted sen-
tencing laws of unprecedented severity, and the prison population has 
reached historic highs and has increased more rapidly since 1994 than 
in any other Western country (Tonry 2004a).

Latent functions and to a lesser extent communicative functions pose 
a formidable problem for principled policy making. They are not pri-
marily aimed at what goes on in courtrooms or on punishments im-
posed, crimes prevented, or efficiency enhanced, but on reputations 
developed, political goals achieved, and elections won.

V. Reconciling Purposes and Functions
In principle, only normative purposes and primary and ancillary func-
tions should count. Thought of as machinery meant to produce some-
thing of value, a sentencing system's most important functions, most 
people would agree, are the imposition of deserved punishments and 
the prevention of crime. Political maneuvers and passions sometimes 
result in the passage of laws that produce injustices, and practitioners 
may sometimes impose unjust sentences for idiosyncratic (e.g., an emo-
tional reaction to a particular offense or offender) or corrupt (e.g., to gain reelection or curry favor) reasons. But it is hard to imagine that many people think that is a good way to make policy or that unjust outcomes in individual cases are a good thing. Some of the main ancillary functions—addressing administrative concerns for efficiency, cost-effectiveness, and resource management—are regularly challenged on normative or policy grounds. In practice, though, the latent functions always powerfully shape policy and sometimes shape practice.

I do not kid myself that my or any other academic’s disparagement of policy makers’ motives is important or likely to influence policy makers to behave differently. It does, however, seem to me important that academics describe what they see and speak truth to power when opportunity arises.

Researchers and theorists can, however, work to develop better ways of understanding sentencing and sentencing policy. We need, for example, to devise new normative theories to encompass ways the world has changed since desert theories took shape in the 1970s (Tonry 2005b). In the early years of the twenty-first century, many people believe that efforts should be made to rehabilitate offenders, and a substantial body of credible evidence shows that sometimes, in well-targeted, well-run programs, this can be done. New restorative and communitarian conceptions of justice are taking shape and programs based on them are proliferating. Recent initiatives such as drug, family violence, and mental health courts are being established across the country as are new programs to divert many offenders from prosecution or prison into treatment. Most of these developments are happening and will continue to happen even though they are hard to reconcile with prevailing retributive theories of punishment. If theorists want to try to influence policy for the good and to be able at least to offer models and ways of thinking that will help policy makers make better choices, the work of developing new theories needs to be done.

That theorizing needs, however, to take account of all normative purposes and primary and ancillary functions. Practitioners want to act justly and to prevent crime effectively, if they can, but they also need to manage their offices efficiently, work within resource constraints, and allocate their efforts among multiple objectives. Most sentencing theories address the primary distributive and preventive functions of punishment, but I know of none that attempt also to encompass the major ancillary functions.
Ignoring the ancillary functions forces their influence into the shadows and often, as happened with the federal guidelines, produces outcomes of the sort Judge Marvin Frankel (1972) referred to as "lawless." By that he meant that sentencing in the 1970s was not subject to rules and was vulnerable to idiosyncratic and arbitrary decision making, and the resulting sentences often said more about the judge than about the offender. Bringing the ancillary functions within the frameworks for thinking about principles governing punishment and prevention will make explicit much that has been invisible and, maybe, achieve greater fairness and justice overall.

Much empirical evidence shows that well-managed guidelines systems can make sentencing more predictable and consistent. A quarter century's experience shows how guidelines' multiple functions can be reconciled in ways that achieve broad support from practitioners and reasonably consistent patterns of sentences imposed while allowing jurisdictions to take account of valid management concerns including resource allocation. The key elements are the development of guidelines that classify offenses and offenders in reasonable ways, that authorize sentences that accord with the sensibilities of the judges and prosecutors charged to apply them, and that allow sufficient flexibility for the individualization of sentences to take account of special circumstances.

The goal is a sentencing system that is fair, evenhanded, and consistent, that takes realistic account of key management interests, and that optimizes legitimacy, public reassurance, and public confidence. We know how to do that.

REFERENCES


Michael Tonry


Farrington, David, and Brandon C. Welsh. In this volume. "A Half Century of Randomized Experiments."


Redlich, Allison D., Henry J. Steadman, John Monahan, John Petrila, and


Michael Tonry


Purposes and Functions


