New members of the United States Sentencing Commission should begin their task by reviewing the history and origins of the guidelines. It is easy for regulators who inherit an established administrative system to presume that it was constructed on a solid foundation. After the new commissioners review the history of the current federal guidelines, however, they will not be able to apply a presumption of steady rational progress to the work of the last fifteen years. Instead, they will have to build more selectively on pieces of the current system and rethink major parts of the structure.

This essay has two short parts. The first part argues that two key guidelines concepts—the assertion that the original guidelines “mirrored” past practice and the claim that the guidelines encompass the “heartland” of offenses—are both false. Yet these two concepts remain central to operation of the current guidelines. They are poor methods for applying the guidelines today.

If the mirroring and heartland concepts are both false as history and destructive as practice, then the new commission must begin not at the present, but back at the beginning. The major shortcoming of the mirroring and heartland concepts is the failure of the commission to explain the real basis and scope of its decisions. As the commissioners start to rebuild, they must explain their choices more completely than their predecessors.

The second part suggests a brief reading list for new commissioners. The list is designed to enlighten the commissioners about the checkered past of the current guidelines. The readings also highlight the most important choices they will face during the reconstruction of the guidelines. There is a rich and vibrant corpus of books and articles written before and after the implementation of the federal sentencing guidelines, and we encourage broad reading and use of this sentencing literature and the related research and data. Nevertheless, we believe there are a handful of articles that make up an essential core. These core writings are the absolute minimum knowledge base, in our view, for understanding the system that now exists, and the paths to further and much-needed reform.

I. Shattered Mirror

The original Commission claimed that its original guidelines “mirrored” the prior sentencing practices in the federal system, and were therefore presumptively reasonable. Mirroring made the original guidelines presumptively reasonable since the guidelines were built on the average of what presumptively reasonable judges had done. While there were exceptions and “adjustments” to the mirroring, the Commission said, the original guidelines were designed in general to produce sentences like those that federal judges selected before the guidelines.

The mirroring claim, however, was not a useful account of how the Commission created the initial guidelines. The Commission never explained how it translated its various sources of data into a starting point for the guidelines. The Commission conceded that the guidelines did not precisely mirror past practice, but it never explained the relative importance of the mirrored and non-mirrored components in the guidelines. It suggested only that mirroring was the rule rather than the exception.

If the Commission had in fact mirrored its guidelines, then sentences under the initial guidelines would look similar to sentences before the guidelines. But looking at the role of probation suggests just how distorted the “mirror” was. Before guidelines, almost fifty percent of federal sentences were to straight probation. Under the initial guidelines, that figure dropped to around fifteen percent.

If the guidelines mirrored past sentencing practices in the federal system, they did so in a fun-house mirror whose angles, curves and twists were unexplained. The Commission, from all appearances, did not attempt to capture many important aspects of past practices in its summary data. For instance, the Commission did not bother to measure the impact of most offender characteristics on past sentences. It also departed from the summary data drawn from prior practice in large groups of cases at the center of federal practice. To the extent that the Commission did acknowledge it was modifying prior sentencing practices, it gave no explanation for the amount or degree of departure from past practice. Nor did the Commission explain in any useful detail its reasons for choosing particular areas for modifying past practice. As a result, judges and scholars today have no meaningful way to evaluate the connection between guideline sentences and pre-guideline sentences.

The claim of mirroring, then, became an oversimplification for a complex blend of strategies that produced the initial guidelines. But the mirroring concept served a purpose: it made the guidelines more acceptable to judges and other constituencies. The rules were said to be translations of the judges’ own collective prior work rather than a new creation.
By alleging that it based its guidelines on prior sentencing practices, the first commission also made a strong claim to legitimacy with Congress. Such a system would address Congress' express concern with disparity by eliminating or restricting outlier sentences, both high and low. Even if the average prior sentence was unjust—too severe or too lenient or the wrong kind of sentence—a system based on prior sentencing practices would address the "uncertainty" and "honesty" concerns with the prior system. It would give the commission a starting point for creating a more just set of guidelines.

A defender of the first Commission might reply that the Commission never claimed to be mirroring past practices on a consistent basis. Language from the 1987 Supplemental Report and the Introduction to the 1987 guidelines acknowledged that the Commission raised sentences and sometimes departed from past practices.

But even if changes in prior practice were acknowledged in this abstract and uninformative way, it has been impossible for scholars to recreate, and therefore to test and assess, the original commission's decisions. If the federal sentencing system were a true administrative system, where courts reviewed the commission's decisions under traditional administrative law principles, this failure of explanation would have invalidated the guidelines from the get-go. Instead, the weakened administrative system of the guidelines has left this utter failure of explanation utterly unexamined.

II. Heartless Heartland
The false claim of mirroring works in tandem with a second central guidelines concept, used to describe the ordinary case under the guidelines. The Commission asserted that sentences should remain within the guidelines whenever a case fell within the "heartland"; a sentencing judge should depart from the guidelines only in unusual cases lying outside of this "heartland." Since the ordinary case was said to mirror past practice, the two concepts were subtly linked. A judge should only rarely sentence outside the guidelines (that is, the heartland should be presumed to be large), because the guidelines by and large identified the most relevant factors from prior practice, and then regularized what judges had been doing before the guidelines appeared.

The 1987 version of the departure statute allowed a district court to depart whenever a case presented "aggravating or mitigating circumstances of a kind or to a degree not adequately taken into consideration by the commission." The legislative history of the SRA suggested that departures should be unusual, the exception rather than the rule. But the Commission refined this standard with its geographic analogy. According to the Introduction to the guidelines, in language that has remained unchanged:

The Commission intends the sentencing courts to treat each guideline as carving out a "heartland," a set of typical cases embodying the conduct that each guideline describes. When a court finds an atypical case, one to which a particular guideline linguistically applies but where conduct significantly differs from the norm, a court may consider whether departure is warranted.

Where did the Commission come up with the "heartland" concept? It was not mentioned in the statute or the legislature history. Nor did the Commission define the term, other than to assert that the heartland is intended to reflect "a set of typical cases embodying the conduct that each guideline describes." But the guidelines do not include any description of the set of typical cases to which they apply, and many guidelines apply to a host of possible statutes. None of the linked concepts of "heartland" or "typicality" or "atypicality" or "unusualness" or "the norm" are self-defining.

Anyone who doubts the power of rhetoric in general, or the power of metaphor in particular, should study the impact of the heartland idea on federal sentencing. Federal courts, from the district courts to the Supreme Court, have given the heartland concept undeserved respect. The vast majority of federal courts, in literally thousands of published decisions, have simply regurgitated the statement from the guidelines introduction, and then announced that the facts and factors before the court either are or are not within the guidelines "heartland." A smaller number of courts, including the United States Supreme Court in United States v. Koon, have discussed heartland, but all federal courts have assumed that the concept makes sense. We believe they are wrong.

Why doesn't the heartland concept make sense? Why isn't the heartland concept simply—and coherently—a reflection of the most important factors in sentencing, translated from prior practice to the guidelines through the Commission's empirical analysis? This is exactly the claim the Commission makes: that sanctions are chosen based on some average, or some accepted "middle" of descriptions of criminal behavior (whether "real" behavior, or charged or convicted statutory offense behavior, or guideline-defined behavior) and of criminals.

The original Commission explained in the guideline introduction that it expected departures to be rare, because the guidelines, offense by offense, seek to take account of those factors that the Commission's data indicate made a significant difference in pre-guidelines sentencing practice. Thus, for example, where the presence of physical injury made an important difference in pre-guidelines sentencing practice (as in the case of robbery or assault), the guidelines specifically include this fac-
tor to enhance the sentence. When the guidelines do not specify an augmentation or diminution, this is generally because the sentencing data did not permit the Commission to conclude that the factor was empirically important in relation to the particular offense.  

There are many puzzling aspects to this statement. First, the guidelines are not organized by statutory offenses. Instead, they are organized around a radically simplified set of "guideline offenses." Thus, at a minimum, it would be essential for the Commission to show how it mapped its limited available information about prior statutory crimes onto the guideline categories. A second problem should be clear from the discussion of mirroring: if the Commission did not in fact mirror, then any reliance on mirroring to validate heartland must be rejected.  

But the essential problem with the heartland concept was its high level of generality. For any given crime or type of offender, there was no way to know what exact factors brought a case within the heartland. Judges could only guess about what sort of case fell within the heart, and presumed that most guideline sentences mirrored prior judicial practice. As a result, judges never became seriously involved in developing a common law of sentencing. They never played an important role in improving the supposedly evolutionary guidelines.  

The heartland concept can only make sense if typical crimes and criminals are expressly linked to one or more guidelines or, in the reverse, if each guideline is expressly linked to a typical crime and criminal. Without some explicit and concrete description of what is usual, there is no way to know what is unusual. Without such a description, the heartland is heartless.  

A judge or lawyer working in the federal system might believe that the heartland concept has more validity than this critique of the concept suggests. The judge or lawyer might say "I know what the heartland of acceptance of responsibility is, or the heartland for drug crimes." But what these actors in the system are asserting is not that the heartland concept explains what the Commission had in mind, but that the working groups in the court -- the prosecutors, judges, probation officer and defense attorneys -- have established settled patterns and rates for some kinds of offenses and offenders.  

It is natural for actors in any system to want to give meaning to that system. Scholars have long recognized the tendency of actors in a legal system to work out functional accommodations. But post hoc rationalizations of unprincipled guidelines by courtroom workgroups cannot give meaning to heartland unless heartland is defined to mean "what workgroups did before guidelines." The analysis of the mirroring claim in this article has revealed that assertion to be suspect. The enormous variation in guideline practice among the 94 federal districts reaffirms that the heartland is a formless lump that can be shaped to each district's habits and beliefs.  

Courts have treated the heartland notion more kindly than it deserves. Federal courts have made assumptions, right and left, about what the Commission "must" have considered to make up the heartland for each guideline -- in the absence of any evidence of what the Commission in fact considered. Until the 1996 decision of the United States Supreme Court in United States v. Koon, this approach led almost every federal court in the United States to ignore the critical statutory terms "adequately" and "considered" and "kind or degree" in the departure standard (18 U.S.C. § 3553(b)). It led most federal circuit courts, before United States v. Koon, to categorically reject district court reliance on a large number of factors because the Commission "must have" or "certainly" or "obviously" or "surely" or "of course" considered the factor at issue.  

Koon has sparked something of a heartland revival.  

In some of these recent decisions, the heartland concept has been used to justify far greater discretion for departures by district courts. Under an aggressive reading of Koon, a few federal courts have begun to take the view that only factors explicitly precluded by statute or the Commission can be categorically rejected from trial judge consideration. As provocative as these new cases are, however, they ultimately leave the notion of heartland as bereft of heart as it ever was. 

This liberalizing trend is by no means universal. Many courts after Koon continue to use the heartland concept as a drag on the departure power, and many continue to find categorical barriers to departure, the language in Koon notwithstanding. Either way, the heartland concept has not helped to produce the rational, just system that Congress intended.  

III. Vibrant Sentencing Corpus  

What should a new commissioner do when presented with a body without a heart? We think commissions need to return to the sentencing literature. Sentencing has emerged as a distinct field of study and practice in the last thirty years. It is feasible for newcomers to read enough for themselves to gain a fairly comprehensive view of the early developments in the field. This history can then inform choices about the most important aspects of current practice to study, and the most likely strategies for change.  

The guidelines emerged from concerns about the indeterminate sentencing systems that defined all U.S. justice systems for most of this century. The reform purposes most often mentioned now as the goals of Congress in the Sentencing Reform Act are greater
"honesty" in sentencing and the avoidance of "unwar-
ranted disparity." The history of the guidelines move-
ment, and the federal guidelines experiment, is much
richer than these shorthand phrases suggest.

Even if the indeterminate sentencing system was
more equitable and more honest than current observers
assume, it would still have been a lawless system, at
least to the extent that "law" requires explicit rules and
standard, public processes. There is a distinct virtue of
bringing law and process to sentencing even for those
who dislike the law and process that has been chosen.

Another lost theme is the desire of reformers, and
of Congress in the sra, to bring purpose or purposes — a
philosophy of punishment — to sentencing." A final
theme, lost to the over-simple refrain of "honesty" and
"avoiding unwarranted disparity" is the desire to bring
science to sentencing — to make sentencing not only an
area of law but an area of law guided by knowledge and
research.

While the sentencing literature is large, it is easy to
search. Most sentencing articles are available on-line, in
Westlaw or Lexis. The principal materials not available
on-line are also readily available in libraries, with most of
the leading texts still in print.

What are the essential readings on the origins of
the guidelines movement? We believe they would
include Marvin Frankel’s classic, short volume, CRIMI-
NAL SENTENCES: LAW WITHOUT ORDER (1975). While
Frankel was only one among several leading voices
against the old order, it was his lectures at the Univer-
sity of Cincinnati Law School, printed in this volume,
that captured the widest attention." Frankel’s critique
continues to offer a fresh and powerful view of what
makes a sentencing system fundamentally wrong and
right, even 25 years after the points were published.
Commissioners today might profitably ask how many of
Frankel’s basic challenges to the old system were
answered by the federal guidelines.

There is surprisingly little literature on how the
pre-guidelines federal system actually worked. The best
single work may be the volume by Stanton Wheeler,
Kenneth Mann, and Austin Sarat, titled SITTING IN
JUDGMENT: THE SENTENCING OF WHITE-COLLAR CRIM-
INALS (1988). Though limited to a study of white collar
cases in a handful of districts, this study brilliantly
opens new and sometimes contrasting views of the pre-
guidelines system.

Judge Frankel’s book is remembered in part
because it proposed and defended the idea of a rela-
tively non-political administrative agency for the devel-
opment of sentencing rules. The best book detailing
the issues that such an agency might face is Pierce O’Don-
nell, Michael J. Churgin & Dennis E. Curtis, TOWARD A
JUST AND EFFECTIVE SENTENCING SYSTEM (1977). This
volume was the product of a series of seminars at Yale
Law School, and served as a blueprint for the initial fed-
eral legislative proposals that eventually became the
Sentencing Reform Act of 1984 (sra).

It may sound obvious, but the Sentencing Reform
Act itself is essential reading. The sra is a fascinating
and complex piece of legislation: key provisions, includ-
ing the 18 U.S.C. § 3553(a) directive to judges and the 28
U.S.C. § 994(g) capacity constraint provision, seem to
have been ignored by courts and the Commission.

Most commentators have treated the legislative his-
tory as if it is captured in a single document — Senate
Report 98—225 (1984). Forgotten, wrongly, is the parallel
Act passed both houses. The final version of the sra
reflected distinctive concerns of both houses; both
reports illuminate the meaning of the language ulte-
rately adopted and the purposes of the act as a whole."

There are also some classic readings dealing with
the Commission’s early development of the federal
guidelines under the new statute. The first document
is, of course, the guidelines manual itself, and espe-
cially the introduction. Another essential early Com-
mission document is the Supplemementary Report on the
Initial Sentencing Guidelines and Policy Statements (June
18, 1987). Then-commissioner Stephen Breyer pub-
lished an insightful defense of the basic drafting
choices encountered during construction of the guide-
lines, in The Federal Sentencing Guidelines and the Key
Compromises Upon Which They Rest, 17 HOFSTRA L. REV.
1 (1988). The dissenting views of Commissioner Paul
Robinson, published at 52 Fed. Reg. 18,133 (1987) also
highlight important features of the guidelines.

The literature about the federal guidelines in oper-
ation has become substantial over the years. Particular
topics, such as guidelines for organizational defen-
dants, the sentencing of drug crimes, and the treatment
of "loss" in fraud or theft cases, have received much
attention from time to time. But there are at least three
themes that cut across all discussions of the federal
guidelines, themes that have showed staying power in
the literature.

The first of these recurring themes is the tension
between "flexibility" and "disparity" in sentencing
rules. Often this debate centers on the proper role of
the judge at sentencing and the proper scope of the depar-
ture power. Sometimes it addresses the proper level of
"complexity" or "simplicity" for guidelines. An early
and comprehensive critique of the system, empha-
sizing its complexity, inflexibility, and restricted role
for judges, appears in Daniel Freed, Federal Sentencing in
the Wake of Guidelines: Unacceptable Limits on the Discre-
the definitive critique from this point of view appeared in
print recently, in the volume by Professor Kate Stith and


A second persistent theme cutting across the federal guidelines literature has been the controversy about "relevant conduct" and corresponding issues of prosecutorial influence over the sentence imposed. A trio of leading articles on relevant conduct by Professors Elizabeth Lear, Kevin Reitz, and David Yellen set out the case against the use of relevant conduct.7 The most important defenses of the concept come from the commission itself and from Professor Julie R. O’Sullivan.8


The third overarching theme we glean from the federal guidelines literature (although far less well-developed than the other two) is the question of the severity of the guidelines, and the related concepts of the proportionality of guideline sentences and the guidelines' lack of attention to alternative (non-prison) sanctions. Some of the literature here discusses the role of Congress in promoting more severe sanctions on a selective basis. Among the most important literature are two reports to Congress from the Commission: *Mandatory Minimum Penalties in the Federal Criminal Justice System* (1991) and *Cocaine and Federal Sentencing Policy* (1995).

A focus solely on the federal guidelines system would produce a blurred image of guideline sentencing in the year 2000. Various guideline structures have been introduced in about half of U.S. states, and structured sentencing systems are used in many other countries. Both U.S. state experience and foreign experience with structured sentencing appears in most ways to have been more successful than the federal guidelines experiment. There are many lessons here for the federal system to learn.

What are the essential readings on other sentencing systems? The single best volume on the design of a U.S. state guideline system is Dale Parent’s book about Minnesota, titled *Structuring Criminal Sentences: The Evolution of Minnesota’s Sentencing Guidelines* (1988). An earlier text dealing with the role of sentencing purposes in the design of sentencing struc-
understand the guidelines we have, and help the commission to produce better ones. The second thing commissioners should do is read and develop its own understanding of its history. If the new commission opens its files, studies its past, holds open meetings, and explains its own decisions, then it may empower judges and scholars to do their part. Together, they might yet rescue the broken federal sentencing system.

Notes
7  USSG Intro. pt. 4(b) (departures) (1998). In the 1987 introduction, the Commission included arson in its list of illustrative crimes where physical injury mattered; in 1998, arson had been deleted. Had the underlying data changed?
8  See MILTON HUEMANN, PEA BARGAINING (1978) (exploring the function of courtroom workgroups in the context of plea bargains).
10  See, e.g., United States v. Sanchez-Rodriguez, 161 F.3d 556 (9th Cir. 1998); United States v. Leshy, 169 F.3d 433 (7th Cir. 1999) (relying heavily on Koon and the heartland concept to hold that “most, but not all, of the conduct regulated by those statutory provisions explicitly referenced would be considered ‘typical’ and within the ‘heartland’ of cases covered by that guideline. By implication, cases embodying conduct regulated by statutory provisions that are not referenced by a particular guideline logically cannot be found to fall within the ‘heartland’ of that particular guideline.”); United States v. Woods, 159 F.3d 1132 (8th Cir. 1998) (holding that “because the underlying offense was bankruptcy fraud, and not drug trafficking or some other offense typical of organized crime, the facts of her money laundering did not fall into the ‘heartland’ of cases involving that offense.”).
12  Marc Miller, Purposes at Sentencing, 66 S. CAL. L. REV. 413 (1992). Congress mentioned the purposes of sentencing eighteen times in the SRA.