

Summer 1981

## Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity

Kevin Clancy

John Bartolomeo

David Richardson

Charles Wellford

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

---

### Recommended Citation

Kevin Clancy, John Bartolomeo, David Richardson, Charles Wellford, Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity, 72 J. Crim. L. & Criminology 524 (1981)

This Symposium is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

*The proposed Federal Criminal Code features several innovations aimed at the problem of disparate sentencing. In keeping with the Journal of Criminal Law and Criminology's singular interdisciplinary tradition, we present below two viewpoints rarely encountered together. The first is a criminological evaluation of sentencing disparities based on empirical studies by a research group. In the second, Judge Tjoflat provides a legal analysis of the same issue, drawing on his practical experience on the federal bench. They complement and supplement one another.*

THE EDITORS

## SENTENCE DECISIONMAKING: THE LOGIC OF SENTENCE DECISIONS AND THE EXTENT AND SOURCES OF SENTENCE DISPARITY\*

KEVIN CLANCY,\*\* JOHN BARTOLOMEO,\*\*\* DAVID  
RICHARDSON,† AND CHARLES WELLFORD‡

*[In] the great majority of federal criminal cases . . . a defendant who comes up for sentencing has no way of knowing or reliably predicting whether he will walk out of*

---

\* The views expressed are those of the authors and not necessarily the position of their respective organizations. The research described in this article is part of a large sentencing study funded by the Federal Justice Research Program.

The authors are grateful to the following people who have carefully reviewed and commented on this paper: Arthur White, Robert Baskin, Adam Yarmolinsky, Brian Forst, William Rhodes, and Karen Skrivseth. Special thanks are due to Paul Berger, who provided valuable guidance in developing and applying several innovative statistical programs, Michael Rouse, who provided research assistance, and all the operational staffs of Yankelovich, Skelly and White, Inc., who contributed to the successful completion of this survey.

\*\* Senior Vice President at Yankelovich, Skelly and White, Inc. and Professor of Marketing at Boston University Graduate School of Management. Dr. Clancy has directed numerous sociological and marketing research studies, and has written extensively on the subject of response bias in surveys. Ph.D., New York University, 1971.

\*\*\* Vice President at Yankelovich, Skelly and White, Inc. Dr. Bartolomeo is an internal consultant to the firm, providing study design and analysis assistance on special projects. Ph.D., University of Pennsylvania, 1977.

† Vice President of the Government Studies Division at Yankelovich, Skelly and White, Inc. He was project manager and principal researcher on the Federal Sentencing Guidelines Research Project, on which this paper is based. M.A., University of Minnesota School of Public Affairs.

‡ Administrator, Federal Justice Research Program, United States Dep't of Justice. Ph.D., University of Pennsylvania, 1969.

*the courtroom on probation, or be locked up for a term of years that may consume the rest of his life, or something in between.*

Judge Marvin E. Frankel<sup>1</sup>

## I. INTRODUCTION

To make sentencing in federal courts more evenhanded and rational, the proposed revision of the Federal Criminal Code would establish a sentencing commission.<sup>2</sup> Within limits established by the legislature, the sentencing commission would establish sentencing guidelines for different types of offenders, offenses, and purposes of sentencing. In addition, the commission would establish a research program to generate information on the effectiveness of sentences imposed and to use that research to recommend modifications of statutes relating to sentencing that produce more effective, humane, and rational sentencing. The research described in this paper is directed towards assessing the extent to which disparity is likely in current federal courts' sentencing practices, and to describe a logic that best accounts for current sentencing practices. This research, and the larger project from which it is drawn, was developed to establish an empirical basis for the initial work of a sentencing commission, and to estimate the extent of the problems in sentencing to which the reforms are addressed.

Most observers of the criminal justice system agree that there are unfair disparities in the sentences meted out in the courts. Indeed, a majority of federal judges believe that the problem exists at least to some extent.<sup>3</sup> However, agreement ends with this, for there is considerable debate about the extent of the problem and the degree to which disparity is attributable to the judge's sentencing style, philosophy on the role of criminal sanction, overall tendency to be lenient or harsh, and other factors.

There is little empirical evidence on the extent and sources of sentence disparity in the federal courts, and virtually no empirical investigation of the phenomenon on a national scale. Judge Marvin Frankel has argued that federal sentence disparity is a serious problem, but has offered only primarily anecdotal documentation.<sup>4</sup> Partridge and Eldridge<sup>5</sup> demonstrate significant variation in the responses of judges in

---

<sup>1</sup> M. FRANKEL, *CRIMINAL SENTENCES: LAW WITHOUT ORDER* 6 (1973).

<sup>2</sup> *See* S. 1722, 96th Cong., 1st Sess. tit. III, §§ 994-95. *See also id.* tit. I, §§ 2001-09, 2101-06, 2201-04, 2301-06, 3725; tit. III, §§ 991-93, 996-97 (1980).

<sup>3</sup> Yankelovich, Skelly & White, Inc., *Judicial Reactions to Sentencing Guidelines* 3-4 (1980).

<sup>4</sup> M. FRANKEL, *supra* note 1, *passim*.

<sup>5</sup> A. PARTRIDGE & W. ELDRIDGE, *THE SECOND CIRCUIT SENTENCING STUDY* (Federal Judicial Center 1974).

one circuit to hypothetical cases. A more recent study by Diamond and Zeisel provides empirical evidence for the existence of serious disparity, but their findings are limited to two federal jurisdictions.<sup>6</sup> Research conducted for the New York State Executive Advisory Committee on Sentencing using simulation techniques, demonstrated disparity among New York state judges.<sup>7</sup> Finally, Rhodes and Conly,<sup>8</sup> using a large sample of cases drawn from several jurisdictions, provide empirical analyses of the factors concerning the defendant and the case that affect sentencing decisions. However, this analysis of actual sentencing decisions permits limited inferences regarding the effect of an individual judge on the sentencing decision.

This paper reports the results of the first national survey of federal judges on the extent and causes of sentence disparity. The study had three principal objectives: determination of the extent of sentence disparity, examination of the effect of specific case characteristics on sentence decisions, and determination of the effect of various attributes of the judge and the judicial environment on sentence decisions.

#### A. DETERMINING THE EXTENT OF SENTENCE DISPARITY

To determine sentence disparity, the study looked at the relative amounts of consensus and dissensus among federal judges regarding the possible sentences for sixteen hypothetical cases. There are two types of sentence disparity. The first type is interjudge disparity which occurs when there is dissensus among judges over the appropriate sentence for cases with effectively identical offense/offender characteristics. The second type is intrajudge disparity. It occurs when there is instability over time of a given judge's sentences for cases with identical or extremely similar offense/offender profiles. Table 1 illustrates the two types of disparity with hypothetical data. The table reveals dissensus among the three judges concerning the length of prison terms and instability over time for Judges X and Z, but not Judge Y. This paper will deal with only the first type of disparity—dissensus among judges, or interjudge disparity. Consequently, any disparity revealed by the analysis probably understates its actual extent. Throughout this paper the phrases

---

<sup>6</sup> S. DIAMOND & H. ZEISEL, *SENTENCING COUNCILS: A STUDY OF SENTENCE DISPARITY AND ITS REDUCTION* (forthcoming). This study was conducted in the Northern District of Illinois and the Eastern District of New York.

<sup>7</sup> NEW YORK STATE EXECUTIVE ADVISORY COMM. ON SENTENCING, CRIME AND PUNISHMENT IN NEW YORK: AN INQUIRY INTO SENTENCING AND THE CRIMINAL JUSTICE SYSTEM (1979) [hereinafter cited as *CRIME AND PUNISHMENT*].

<sup>8</sup> W. RHODES & C. CONLY, *ANALYSIS OF FEDERAL SENTENCING* (Inst. for Law and Soc. Research) (forthcoming).

“dissensus among judges” and “sentence disparity” will be used interchangeably, except where this could prove confusing.

**TABLE 1**

HYPOTHETICAL DATA TO ILLUSTRATE THE TWO  
DIMENSIONS OF SENTENCE DISPARITY

(Prison sentences for identical cases (in years) by judge and time)

TIME	JUDGE			MEAN SENTENCE
	<u>X</u>	<u>Y</u>	<u>Z</u>	
T <sub>1</sub>	3	5	4	4
T <sub>2</sub>	4	5	6	5
T <sub>3</sub>	2	5	2	3
MEAN SENTENCE	3	5	4	

Consistent with this conceptualization, we have operationalized dissensus as the degree to which judges disagree about the sentence that should be served by offenders found guilty in sixteen hypothetical cases. For each of these cases, judges received seven pieces of information about the offender and the offense and were asked to make a sentencing decision. As the sixteen cases were identical, or “held constant,” for all judges in the survey, disagreement about sentences is a measure of interjudge disparity, or dissensus.<sup>9</sup> Thus, for these sixteen cases, the first objective of the analysis was to partition the variance in sentence decisions into two parts: (1) that on which judges agree and which is directly attributable to characteristics of the case; and (2) that on which there is dissensus not attributable to the case, but rather to the judge or elements of the judge’s environment. In statistical terms, this partition separates that proportion of decision variance accounted for by the characteristics of the case from the unexplained variance that is considered attributable to the judge.

---

<sup>9</sup> This is not the only possible interpretation of this disagreement. Since the cases presented to judges contained seven pieces of information—less than would occur in the real world—arguably, at least some of the disagreement about sentences would diminish were additional information presented to the judge. However, the former interpretation is supported by Wilkins and Gottfredson, and by Hogarth, who provided evidence that sentencing and parole decisionmakers do not base their decisions on more than about five factors, or sets of factors. *See* D. GOTTFREDSON & L. WILKINS, *THE UTILIZATION OF EXPERIENCE IN PAROLE DECISION-MAKING: A SUMMARY REPORT* (Nat’l Inst. of Law Enforcement and Crim. Just. 1974); J. HOGARTH, *SENTENCING AS A HUMAN PROCESS* (1971). In any event, they are based on hypothetical cases, so the figures reported in this article should be interpreted only as indicators of the degree of dissensus that occurs among judges rather than precise measures of actual disparity in real sentences.

The second and third objectives involve more detailed analyses of each portion of this partitioned variance in sentence decisions—the role of specific case characteristics and the role of particular attributes of the judge and the judicial environment.

B. DETERMINING THE EFFECT OF SPECIFIC CASE CHARACTERISTICS ON SENTENCE DECISIONS

The study examined the influence of the following seven case characteristics on sentence decisions: (1) the offense itself (two were considered, bank robbery and fraud), (2) the offender's age, (3) the offender's criminal record, (4) the offender's role as principal or accomplice to the offense, (5) method of case disposition (plea or trial), (6) the dollar amount involved, and (7) the role of membership in an ongoing criminal organization (for fraud offenses) or the use of a weapon (for bank robbery). The study used these seven factors because other empirical investigations of sentence decisions conducted at the state and local level have indicated their importance.<sup>10</sup>

An additional element of this objective was to determine the ways in which certain offense or offender profiles elicit judgments about the appropriate goals of a sentence which, in turn, influence the sentence itself.<sup>11</sup> For example, what attributes of the case "cue" the judge that the principal goal of the sanction should be general deterrence rather than, say, incapacitation? Moreover, how do these case-specific goals influence the nature and extent of the sentence imposed? Case-specific goal articulation is a potential source of both dissensus and consensus in sentence decisions. Goal articulation aggravates the problem of sentence disparity to the degree that identical cases yield disparate goals which in turn, produce disparate sentences. If judges agree on sentence goals in specific cases, and agreement on goals does indeed lead to similar sentences, then goal articulation contributes to the maintenance of uniformity in sentencing.

C. DETERMINING THE EFFECT OF VARIOUS ATTRIBUTES OF THE JUDGE AND THE JUDICIAL ENVIRONMENT ON SENTENCE DECISIONS

Is dissensus about the appropriate sentence related to characteristics of the judge and the judicial environment in which he works? Iden-

---

<sup>10</sup> L. WILKINS, J. KRESS, D. GOTTFREDSON, J. CALPIN & A. GELMAN, *SENTENCING GUIDELINES: STRUCTURING JUDICIAL DISCRETION* (Nat'l Inst. of Law Enforcement and Crim. Just. 1978) [hereinafter cited as *SENTENCING GUIDELINES*].

<sup>11</sup> For a discussion of judges' and other legal professionals' views of the criminal justice system's overall goals, see Yankelovich, Skelly & White, Inc., *Sentencing Goals and their Application in the Federal Courts* (August 1980).

tifying the source of interjudge disparity in sentence decisions could furnish valuable insights into methods of curtailing future disparity. The study analyzed five sets of factors having no direct bearing on the case itself to understand what accounts for disparate sentences given for identical cases: (1) differential perceptions of the severity of different sentences, (2) the judge's propensity to give relatively harsh sentences, (3) the judge's overall goal orientation regarding criminal sanction, for example, his predisposition regarding the functions of sentences, regardless of the specifics of the case, (4) background characteristics such as demographics and career history, and (5) the region of the country in which the judge's district is located.

The discussion below is keyed to these three objectives, gradually unfolding the logic of sentencing decisions that judges made in the survey interview. However, before presenting these findings, it is necessary to describe the methodology employed.

## II. METHODOLOGY

There are three key elements of the methodology: the sample and field work, the questionnaire, and the analytic techniques.<sup>12</sup>

### THE SAMPLE AND FIELD WORK

Two-hundred-sixty-four active federal district judges, interviewed between September and November, 1979, comprised the sample. Professional interviewers personally conducted interviews in the judges' chambers or other offices located in the court building. The average interview lasted ninety minutes.

### THE QUESTIONNAIRE

The results reported in this paper reflect responses to a significant portion of the survey interview. All of the questionnaire items used in these analyses were structured, except that judges were free to give whatever sentences they chose to hypothetical case scenarios. There were four sets of questions.

The first set required judges to make sentencing decisions based on attributes of the offender and the offense. This involved three tasks:

---

<sup>12</sup> Technical descriptions of major aspects of the methods employed are available upon request from the authors.

TABLE 2

## EXAMPLES OF CASE SCENARIOS

## CARD # 1

Offense:	The offense committed was bank robbery
Age:	The offender is between 18 and 24 years old
Record:	The offender has no past criminal record
Role:	The offender was an accomplice; he was following the lead of others
Plea/Trial:	The offender entered a plea of guilty
Extent of Injury to Victim:	A weapon was used and at least one of the victims suffered serious personal injury
Dollar Amount:	The total amount involved was less than \$1,000

## CARD # 2

Offense:	The offense committed was fraud
Age:	The offender was 35 to 65 years old
Record:	The offender has a dozen previous arrests, most for serious offenses, and has been convicted four times
Role:	The offender was an accomplice; he was following the lead of others
Plea/Trial:	The offender entered a plea of guilty
Organization:	The offender is not a member of any criminal organization
Dollar Amount:	The total amount involved was less than \$1,000

*Task 1.* Interviewers presented the judges with sixteen offense/offender case "scenarios" and asked them to choose the appropriate sentence. Two examples of these scenarios appear as Table 2. Judges were free to administer sentences that comprised any combination of prison time (in months), supervised time (in months), and fine (in dollars). Moreover, they were instructed to indicate the actual sentence that should be served by the offender. In order to maximize the poten-



tial for drawing statistical inferences about the effect of each factor on sentence decisions, the scenarios reflect an incomplete orthogonal experimental design.<sup>13</sup> Note that the study used two opposite extreme values of each factor in the scenarios. Across all scenarios, these extreme levels of each factor were evenly balanced to insure orthogonality. Table 3 provides a full display of the factors, and the levels within factors, for the sixteen offense/offender scenarios.

**TABLE 3**  
**EXPERIMENTAL DESIGN FOR 16 OFFENSE/OFFENDER**  
**SCENARIOS**

SCENARIO #	CRIME	AGE	RECORD	ROLE	GUILTY BY	DOLLAR AMOUNT	CRIMINAL ORGANIZATION	WEAPON
1	Robbery	Young	None	Accomplice	Plea	Low	N.A.	Yes
2	Fraud	Old	Long	Accomplice	Plea	Low	No	N.A.
3	Robbery	Young	Long	Accomplice	Trial	Low	N.A.	No
4	Fraud	Old	None	Accomplice	Trial	Low	Yes	N.A.
5	Robbery	Old	None	Principal	Trial	Low	N.A.	Yes
6	Fraud	Young	Long	Principal	Trial	Low	No	N.A.
7	Robbery	Old	Long	Principal	Plea	Low	N.A.	No
8	Fraud	Young	None	Principal	Plea	Low	Yes	N.A.
9	Fraud	Old	Long	Principal	Trial	High	Yes	N.A.
10	Robbery	Young	None	Principal	Trial	High	N.A.	No
11	Fraud	Old	None	Principal	Plea	High	No	N.A.
12	Robbery	Young	Long	Principal	Plea	High	N.A.	Yes
13	Fraud	Young	Long	Accomplice	Plea	High	Yes	N.A.
14	Robbery	Old	None	Accomplice	Plea	High	N.A.	No
15	Fraud	Young	None	Accomplice	Trial	High	No	N.A.
16	Robbery	Old	Long	Accomplice	Trial	High	N.A.	Yes

SPECIFICATIONS OF EXTREME LEVELS USED IN THE SCENARIOS

<i>Age:</i>	Young = 18-24 years old Old = 35-64 years old
<i>Record:</i>	None = No prior arrests Long = 12 arrests and 4 convictions for serious offenses
<i>Dollar Amount:</i>	Low = Less than \$1,000 High = More than \$100,000
<i>Criminal Organization/Weapon:</i>	N.A. = Does not apply to this scenario

This entire sequence of questions constitutes an experiment that was incorporated into the questionnaire. The sixteen case scenarios are logically equivalent to experimental treatments or stimuli, and the sentence imposed is the response to the stimuli. In keeping with this experimental parlance, any variance in the sentence decisions imposed that is not attributable to the stimuli (*i.e.*, the specific characteristics of the case and the offender) is assumed to be attributable to the subject (*i.e.*, the judge). Thus, the amount of variance unaccounted for by the factors in the scenarios is interpretable as the degree of dissensus among judges about sentence decisions.

*Task 2.* The second task was a simplified version of the first. Judges gave monadic (*i.e.*, individual, one-at-a-time) ratings of the influence that extreme and intermediate levels of each factor should have on the severity of a sentence. The study used a standard seven-point scale,

<sup>13</sup> See W. COCHRAN & G. COX, EXPERIMENTAL DESIGNS 439-82 (1957).

anchored at either end by "should make the sentence much tougher", versus "should make the sentence much lighter." Gathering these monadic ratings is necessary to facilitate interpolations of the effects of intermediate levels of each factor. (Recall that only extreme levels were incorporated into the scenarios used in Task 1.)

*Task 3.* Finally, judges were asked to indicate the goal they hoped to achieve with their sentence for four of the sixteen offense/offender scenarios. The four scenarios were randomly selected in each interview. The goal alternatives presented to the judge were general deterrence, special deterrence, rehabilitation, incapacitation, and retribution/deservedness.

The second set of questions, asked separately from the first, look at the perceived severity of sentences. This was done because judges do not give sentences in a single or metric "currency." Rather, sentences are composites of at least three components: prison time, supervised time, and fine. Possibly judges have differential perceptions of the severity of each of these components of a sentence; or perhaps they disagree about the severity equivalences across these components. The study contained questions designed to map the three components of a sentence into a single composite index of perceived sentence severity. The questioning procedure parallels the one described above for sentence decisionmaking. It entailed two tasks:

*Task 1.* Judges indicated their perception of the severity of each of nine composite sentences composed of prison time, supervised time, and fine.<sup>14</sup> The severity scale had seven points, ranging from "extremely severe" to "extremely lenient." The nine sentence composites were constructed according to the same experimental design principles that guided the construction of the sixteen offense/offender scenarios. Table 4 displays the nine sentence composites and reveals the experimental design that supported their construction.

*Task 2.* The second task involved monadic ratings of the severity/leniency of extreme and intermediate levels of prison time, supervised time and fine.

The third set of questions elicited attitudinal data on current sentencing policies, including the perceived importance of various goals of criminal sanction. The final set collected background information on the judge, such as age, years of experience, prior career history, region of the country, and so forth.

---

<sup>14</sup> A number of judges found it difficult to evaluate severity without having a specific crime or offender whom they were judging. In analyzing these data, each judge's responses were individually constructed into a ranking of relative severities of alternative sentence combinations.

**TABLE 4**  
 EXPERIMENTAL DESIGN FOR THE NINE SENTENCE  
 COMPOSITES

SENTENCE COMPOSITE #	PRISON TIME (YEARS)	SUPERVISED	
		TIME (YEARS)	FINE (\$)
1	0	0	0
2	2	2	5,000
3	25	5	30,000
4	0	2	30,000
5	2	5	0
6	25	0	5,000
7	0	5	5,000
8	2	0	30,000
9	25	2	0

#### ANALYTIC TECHNIQUES

The study employed several analytic techniques. Repeated measures analysis of variance indicated the role of case characteristics and the judge in sentencing decisions. Interpolation and extrapolation techniques estimated the effects of intermediate levels of offender/offense characteristics that were not included in the scenarios, but instead were asked monadically.<sup>15</sup> Analysis of variance determined the role of case-specific goals in sentence disparities. Micromodeling analyzed the perceived severity of sentences for each judge. Finally, the effect of judges' background characteristics on sentence decisions was determined through analysis of variance followed by step-wise regression analysis.

<sup>15</sup> Responses to these monadic questions were also used to test the reliability of responses to the scenario questions. Because the simulation procedures used to elicit sentences to the hypothetical scenarios are somewhat artificial, we were concerned that judges' responses might be unstable. Therefore, a correlation analysis was run, relating the effects of each factor (e.g., offense) in the scenario analysis to other data provided by the judge later in the interview on the seriousness of each factor. For instance, the judge's individual ratings of the effect of a weapon on their sentence was correlated with the effects observed in scenarios that included use of a weapon by the offender. These correlation analyses showed a very strong relationship—in some cases over .9—between monadic factor score ratings and scenario ratings of the same factors, suggesting that responses to the scenarios provide reliable estimates of sentences.

**TABLE 5**  
PARTITIONING OF INCARCERATION DECISION VARIANCE  
(N = 208)

<u>SOURCE OF VARIANCE:</u>	<u>UNIQUE VARIANCE IN DECISION TO IMPRISON</u>
Offense/Offender Characteristics (main effects)	37%
Judge/Environment (main effects)	10%
Interaction Effects (between offense/offender characteristics <i>and</i> between judge and offense/offender characteristics)	53%

**TABLE 6**  
SUMMARY: PARTITIONING OF SENTENCE DECISION  
VARIANCE (N = 208)

<u>SOURCE OF VARIANCE:</u>	<u>UNIQUE VARIANCE FOR:</u>		
	<u>PRISON TIME</u>	<u>SUPERVISED TIME</u>	<u>FINE</u>
Offense/Offender Characteristics (main effects)	45%	1%	4%
Judge Environment (main effects)	21%	55%	38%
Interaction Effects (between offense/offender characteristics <i>and</i> between judge and offense/offender characteristics)	34%	44%	58%

### III. FINDINGS

#### A. HOW MUCH DISPARITY?

The survey results suggest that sentence disparity does exist. The amount of such disparity varies depending on which aspect of the sentencing decision is examined.<sup>16</sup> If one concentrates strictly on the decision to incarcerate, the disparity is modest (see Table 5). However, when one examines the full range of sentencing options simultaneously,

<sup>16</sup> Commentators have posited several theories about how the sentencing decision is actually made. One is that the process has two steps, the first of which is to determine whether or not the offender should be incarcerated, and the second, to determine the length of incarceration. See SENTENCING GUIDELINES, *supra* note 10, for discussion of this view. An alternative view is that judges weigh all their sanctioning alternatives simultaneously. The discussion in this article concentrates on the latter view.

TABLE 7

DETAILED ANALYSIS OF VARIANCE RESULTS: EFFECTS OF  
OFFENSE/OFFENDER CHARACTERISTICS,  
JUDGES/ENVIRONMENT AND  
INTERACTIONS ON PRISON TIME

<u>SOURCE OF VARIANCE</u>	<u>UNIQUE VARIANCE</u>	<u>SUM OF SQUARES</u>	<u>MEAN SQUARES</u>	<u>F-TEST RATIO</u>	<u>SIGNIFICANCE</u>
Offense/Offender Characteristics (main effects)	45%	49,735	3316	271	< .001
Judge/Environment (main ef- fects)	21%	23,916	116	9.7	< .001
Interaction Effects (between of- fense/offender characteristics and between judge and of- fense/offender characteristics)	34%	37,985	12		

more dissensus appears. More variance in sentences is explained by differences among individual judges than by any other single factor. See Tables 6, 7, 8, and 9. Differences among judges play a particularly strong role in accounting for the amounts of supervised time and fines, completely overwhelming the variance explained by all of the offender/offense characteristics measured in the survey results. The judge is less a factor in explaining the length of a convicted offender's prison term, but is nonetheless significant. The judge accounts for 21% of the variance in the prison term imposed, compared to 45% that is attributable to characteristics of the case.

This partitioning of sentence decision variance understates the amount of disparity exhibited in responses to the sixteen cases tested in the survey because additional disparity is harbored in the interaction effects that are also displayed in Tables 6, 7, 8 and 9. To understand this point, it is necessary to digress and provide a brief explanation of interaction effects. Interactions are conditional effects. The relationship between an independent variable and the dependent variable depends on the value of another independent variable—being greater for certain values of the second independent variable than for others.

Table 10 provides an example of an interaction between two case characteristics.<sup>17</sup> Offenders who are accomplices get shorter prison terms than those who are principals (6.2 years vs. 7.8 years). Similarly, offenders who plead guilty are sentenced to less time in prison than of-

<sup>17</sup> This example can be elaborated to include interactions between three, four, or more independent variables.

TABLE 8

DETAILED ANALYSIS OF VARIANCE RESULTS: EFFECTS OF  
OFFENSE/OFFENDER CHARACTERISTICS,  
JUDGES/ENVIRONMENT AND  
INTERACTIONS ON SUPERVISED TIME

<u>SOURCE OF VARIANCE</u>	<u>UNIQUE VARIANCE</u>	<u>SUM OF SQUARES</u>	<u>MEAN SQUARES</u>	<u>F-TEST RATIO</u>	<u>SIGNIFICANCE</u>
Offense/Offender Characteristics (main effects)	1%	350.7	23.4	6.7	< .001
Judge/Environment (main ef- fects)	55%	13,798	65.4	18.7	< .001
Interaction Effects (between of- fense/offender characteristics and between judge and of- fense/offender characteristics)	44%	11,045	3.5		

TABLE 9

DETAILED ANALYSIS OF VARIANCE RESULTS: EFFECTS OF  
OFFENSE/OFFENDER CHARACTERISTICS,  
JUDGES/ENVIRONMENT AND  
INTERACTIONS ON FINE

<u>SOURCE OF VARIANCE</u>	<u>UNIQUE VARIANCE</u>	<u>SUM OF SQUARES</u>	<u>MEAN SQUARES</u>	<u>F-TEST RATIO</u>	<u>SIGNIFICANCE</u>
Offense/Offender characteristics (main effects)	4%	7,455	497	11.7	< .001
Judge/Environment (main ef- fects)	38%	71,112	421	9.95	< .001
Interaction Effects (between of- fense/offender characteristics and between judge and of- fense/offender characteristics)	58%	107,335	42.3		

fenders found guilty through a trial (6.8 vs. 7.2 years). However, even though offenders who plead guilty get lighter sentences on the average (6.8 vs. 7.2), the impact on pleading is quite large for accomplices (5.6 vs. 6.7), but actually reversed for principals (8.0 vs. 7.6). Clearly there is some effect other than the cumulative impact of the two independent factors that only interaction can explain.

The interaction effects that are indicated in Tables 6, 7, 8 and 9 are each composites of two basic types of interactions. The first type includes interactions between the various offense/offender characteristics and, thus, are like the example given above. The second type of interaction is between the judges and specific offense/offender characteristics

TABLE 10

INTERACTION BETWEEN OFFENDER'S ROLE IN THE CRIME  
AND THE METHOD OF DISPOSITION ON MEAN  
PRISON TIME (IN YEARS) (N = 208)

ROLE IN THE CRIME	METHOD OF DISPOSITION		MEAN
	PLEA OF GUILTY	FOUND GUILTY BY TRIAL	
Accomplice . . . . .	5.6	6.7	6.2
Principal . . . . .	8.0	7.6	7.8
Mean . . . . .	6.8	7.2	

and represents patterned differences between judges in the influence of offense/offender characteristics on their sentence decisions. This second type of interaction is an additional form of sentence dissensus.

This point returns us to the earlier claim that our initial partitioning of decision variance underestimated the disparity measured in the survey, for the majority of these interactions are of the second type, between the judge and specific offense/offender characteristics. Table 11 furnishes a breakdown of these interactions. The most important interactions affecting prison term and supervised time involve the offense itself and the offender's record. Indeed, the rank order of the magnitude of all six interactions is identical for both prison and supervised time. When the two crimes tested were bank robbery and fraud, the sizeable interaction variance attributable to the offense itself (11.7% for prison term and 8.5% for supervised time) suggests that judges have rather different perceptions of the seriousness of common crime and white-collar crime. However, the amount of fine imposed is highly contingent on the judge's concern about the dollar damage inflicted by the crime. Thirty-six percent of interaction variance is attributable to dollar damage, compared to 4.8% for the offense and 1% for the offender's record. These findings regarding interactions generally strengthen the inference that sentence dissensus is a major factor in explaining sentence decisionmaking. They suggest that disparity is not simply random divergence of opinion about the sentence that should be imposed, but is also a consequence of patterned differences of opinion about the influence that specific case attributes should have on the sentence.

Then, too, the lesser role played by interactions between case attributes themselves (the first type of interaction discussed above) suggests that the logic underpinning sentence decisions is relatively straightforward and "additive." This is not to say that the decision logic is simple. The data do suggest, however, that decisions are generally not characterized by multiple contingencies or conditional logic whereby any given datum about the offender or the offense takes on altered

**TABLE 11**  
 TWO-WAY INTERACTIONS BETWEEN SPECIFIC  
 OFFENSE/OFFENDER CHARACTERISTICS AND THE  
 JUDGE (N = 208)

<u>JUDGE INTERACTION WITH:</u>	DECISION VARIANCE EXPLAINED FOR: SUPERVISED		
	<u>PRISON TIME</u>	<u>TIME</u>	<u>FINE</u>
Offense.....	11.7%	8.5%	4.8%
Offender's Record.....	4.8%	8.4%	1.0%
Dollar Damage.....	2.3%	2.6%	36.0%
Offender's Age.....	1.5%	2.0%	1.9%
Offender's Role in the Crime .....	1.5%	2.0%	2.3%
Method of Disposition ..	1.0%	1.4%	0.3%
Total variance attributable to two-way interactions between offense/offender characteristics and judge:	22.8%	24.9%	46.3%

meaning given other data about the case. The logic of sentence decisions is explored further in the following section, which treats the role of each case characteristic and the role of the goal the judge hopes to achieve with the sentence imposed.

**B. THE LOGIC OF SENTENCE DECISIONS: THE ROLE OF SPECIFIC CASE ATTRIBUTES AND CASE-SPECIFIC GOALS**

*Case Attributes*

Not surprisingly, the attributes of the offender and the offense have a significant bearing on the sentence imposed. However, these effects are anything but uniform across the three components of sentences. All of the case characteristics studied have a significant impact on the prison term, but only the past record of the offender influences the amount of supervised time, while the crime and the age of the offender affect the fine imposed (see Table 12). These patterns support the inference (also not surprising) that prison term is the most important element of a sentence and as such, is most sensitive to the various nuances of the case.

The most important determinant of the prison term variable is the crime itself, suggesting that fundamental convictions held among the judiciary about the relative gravity of common crime as opposed to



white-collar crime, and that common crimes are better punished with longer incarceration terms. This is corroborated by the tendency of judges to impose relatively heavy fines for the fraud cases. Also quite important is a set of factors related to the amount of harm caused by the crime and indicators of the harm that the offender might potentially have on society. In rank order these include the offender's past criminal record, the use of a weapon and the injury inflicted, the offender's membership in a criminal organization, the dollar damage of the crime, and the prominence of the offender's role in the crime. The method of case disposition has a modest, though statistically significant, impact on the prison sentence. The relationship between prison sentence and the offender's age is somewhat more complicated. Judges show some leniency for juveniles (under 18 years of age) and younger offenders (18 to 24 years), are somewhat harsher for middle-age offenders (25 to 60 years), and are lenient for older offenders (over 60 years of age).<sup>18</sup> There is, of course, an artificial quality to these data. When actually sentencing, judges may construct composite sentences of prison time, supervision, and fine. To examine the effect of different factors on each of these sentence components separately does not account for trade-offs among them. Even prison time, clearly the most important component of the criminal sanction, does not constitute the complete sentence. A full understanding of the sentencing decision process requires that the three components be considered simultaneously for any given case. A subsequent section of this report therefore combines these three sentence components into a single measure to evaluate sentence disparities and the influence of specific case characteristics. However, before these findings

<sup>18</sup> As a test of the reliability of these findings, these data were compared to the Rhodes/Conly analysis of actual sentence decisions. See note 8 *supra*. Direct comparisons are difficult because of the formulation of the analytic equations and differences in both independent and dependent variable definition. However, four key variables were ranked for their influence on prison terms with the following results.

Rank Order on Effect of:	Yankelovich Skelly and White Simulation Impact on Prison Term	INSLAW - Actual Data Impact on Prison Term	
		Fraud	Bank Robbery
Prior record .....	1	1	1
Age .....	2	4	2
Role in crime .....	3	2	3
Plea .....	4	3	4

The only difference occurs in age in actual fraud cases, and is probably a reflection of differences in the age distribution of actual offenders and the hypothetical offenders described in the survey.

**TABLE 12**  
**EFFECT OF SPECIFIC OFFENSE/OFFENDER**  
**CHARACTERISTICS ON PRISON TIME, SUPERVISED**  
**TIME, AND FINE**

AVERAGE SENTENCES<sup>a</sup> = 84 months of prison time/20 months of supervised time/\$1536 of fine.

N = 208

CASE CHARACTERISTICS	INCREASE (+) OR DECREASE (-) IN AVERAGE SENTENCE DUE TO CASE CHARACTERISTICS		
	PRISON TIME (MONTHS)	SUPERVISED TIME (MONTHS)	FINE (\$)
<u>Type of Crime</u>			
Bank Robbery .....	+33	n.s. <sup>c</sup>	-617
Fraud .....	-33	n.s.	+617
<u>Age</u>			
Under 18 <sup>b</sup> .....	-12.0	n.s.	-536
18 - 24 .....	- 6.8	n.s.	-306
25 - 35 <sup>b</sup> .....	+ 4.3	n.s.	+193
35 - 60 .....	+ 6.8	n.s.	+306
Over 60 <sup>b</sup> .....	- 3.5	n.s.	+155
<u>Record</u>			
No past record .....	-21	+1	n.s.
No record; but suspected <sup>b</sup> .....	0	0	n.s.
1 arrest; no convictions <sup>b</sup> .....	- 6	0	n.s.
Multiple arrests; no convictions .	+ 1	0	n.s.
5 arrests; 1 conviction <sup>b</sup> .....	+ 9	0	n.s.
12 arrests; 4 convictions .....	+21	-1	n.s.
<u>Role</u>			
Accomplice .....	- 9	n.s.	n.s.
Principal .....	+ 9	n.s.	n.s.
<u>Method of Disposition</u>			
Plead guilty .....	- 3	n.s.	n.s.
Found guilty .....	+ 3	n.s.	n.s.
<u>Dollar Damage</u>			
Under \$1,000 .....	-11.4	n.s.	n.s.
over \$100,000 .....	+11.4	n.s.	n.s.
<u>Membership in Criminal Organization</u>			
Not a member .....	-14.6	n.s.	n.s.
Member .....	+14.6	n.s.	n.s.
<u>Use a Weapon</u>			
No weapon .....	-16.4	n.s.	n.s.
Weapon and injury .....	+16.4	n.s.	n.s.

<sup>a</sup> This represents the average across all sixteen scenarios of the sentences judges assign. In statistical terms, it is the *grand mean*. It provides a base for comparison of individual factor effects on sentences, and has no meaning by itself. Because of the artificial nature of the cases, numerical effects of specific characteristics should be used only as indicators and should not be compared to real-world sentences.

<sup>b</sup> Estimate based on interpolation/extrapolation.

<sup>c</sup> Effect of this variable is not statistically significant.

are presented it is valuable to discuss the role of case-specific goals on prison time.

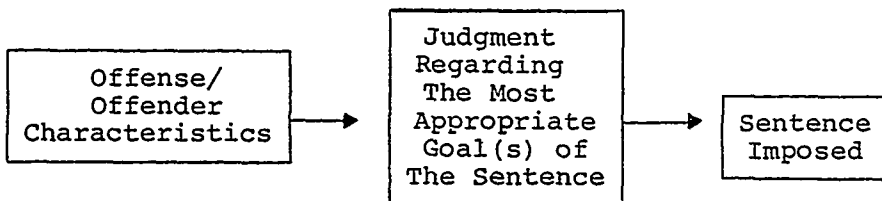
### *Case-Specific Goals*

The foregoing section indicates the role of several offense and offender characteristics in judges' sentencing decisions. However, it leaves unexamined the cognitive process underpinning those decisions. An implicit first step in sentencing is the articulation of a specific goal or intended function of the sentence imposed.<sup>19</sup>

In this analysis we examined how, in each specific case, certain configurations of case attributes are associated with a particular goal or goals. Since this process of goal articulation is fitted to the attributes of each case, it may bear little relationship to the judge's overall philosophy regarding the functions of criminal sanctions.<sup>20</sup> In other words, these case-specific goals intervene between the offense/offender characteristics and the sentences, and result in sentences different from the ones that would have occurred had other goals been sought. This proposition of the role of case-specific goal articulation is captured in Figure 1.

**FIGURE 1**

#### THE ROLE OF CASE-SPECIFIC GOALS



The crucial point to be made about case-specific goals is that they are a potential source of sentence disparity. If identical cases produce divergent judgments about the appropriate goal and if the choice of a

<sup>19</sup> Most theoretical and empirical work on sentencing discusses sentencing goals to some extent. Some, such as W. GARFIN, *PARTIAL JUSTICE: A STUDY OF BIAS IN SENTENCING* (1975), simply describe the major goals of sentencing on which there is a presumed consensus among the judiciary. Others focus on more specific issues, for example, the goals of imprisonment, *see* N. MORRIS, *THE FUTURE OF IMPRISONMENT* (1974); REPORT OF THE 20TH CENTURY FUND TASK FORCE ON CRIMINAL SENTENCING, *FAIR AND CERTAIN PUNISHMENT* (1975); or on the achievement of specific goals, *see* F. ZIMRING & G. HAWKINS, *DETERRENCE: THE LEGAL THREAT IN CRIME AND CONTROL* (1973). A few researchers have considered how goals should be incorporated into sentencing guidelines. *See, e.g.*, *CRIME AND PUNISHMENT*, *supra* note 7.

<sup>20</sup> Another report in this series discusses the impact on sentences of both case-specific and overall goal philosophies. *See* D. Richardson, *Sentencing Goals and their Application in the Federal Courts*, Yankelevich, Skelly & Wright, Inc. (Aug. 1980).

particular goal is, in turn, related to the sentence that is imposed, disparity can result.

The survey provides data to test these hypotheses. After judges sentenced the hypothetical offenders in the sixteen case scenarios, interviewers asked them to indicate the principal goal they hoped to achieve with their sentence.<sup>21</sup> Table 13 shows the distribution of responses, which appears to substantiate the proposition that identical cases yield

**TABLE 13**  
THE RELATIVE DISTRIBUTION OF CASE-SPECIFIC GOALS  
FOR SIXTEEN OFFENSE/OFFENDER SCENARIOS

OFFENSE/OFFENDER CASE SCENARIOS	NUMBER OF OBSERVA- TIONS <sup>1</sup>	PRINCIPAL CASE-SPECIFIC GOALS				
		GENERAL DETER- RENCE (%)	SPECIAL DETER- RENCE (%)	REHABILI- TATION (%)	INCAPACI- TATION (%)	RETRIBU- TION (%)
1 - Bank robbery <sup>2</sup> . . . .	101	39	18	25	8	10
2 - Fraud . . . . .	78	26	34	5	26	9
3 - Bank robbery . . . . .	96	36	14	4	33	13
4 - Fraud . . . . .	81	30	26	38	2	4
5 - Bank robbery . . . . .	64	52	20	5	12	11
6 - Fraud . . . . .	56	20	24	20	20	16
7 - Bank robbery . . . . .	60	30	10	0	47	13
8 - Fraud . . . . .	53	15	21	55	0	9
9 - Fraud . . . . .	36	25	19	0	37	19
10 - Bank robbery . . . . .	40	30	23	25	10	12
11 - Fraud . . . . .	43	51	16	12	14	7
12 - Bank robbery . . . . .	54	22	30	7	30	11
13 - Fraud . . . . .	48	34	23	10	23	10
14 - Bank robbery . . . . .	41	51	29	12	0	8
15 - Fraud . . . . .	25	24	40	12	0	24
16 - Bank robbery . . . . .	47	43	15	4	34	4

<sup>1</sup> Any one judge answered this question for just the first four of the sixteen scenarios. Though scenario cards were to be shuffled, there is some overrepresentation of the first cards and underrepresentation of the last.

<sup>2</sup> Refer to Table 3 for the specific combination of factors associated with each scenario.

divergent goals among judges. In any given case, a goal rarely is cited by less than 10% of the judiciary, but in only three cases does a majority seek any single goal: *general deterrence* in cases 5, 11, and 14. Deterrence is the goal most likely to drive sentence decisions. The average frequency of citation for general deterrence is 33%, followed by special deterrence (23%). Incapacitation (18%), rehabilitation (15%), and retribution (11%) are less frequently pursued. Finally, the nature of the crime seems to have some bearing on the intended purpose of the sen-

<sup>21</sup> Each judge indicated the goal he hoped to achieve for only four of the 16 case scenarios. Interviewers rotated the case scenarios about which they asked this question.

tence. Bank robbery is more likely than fraud to prompt the judge to seek the goals of general deterrence, incapacitation, and retribution. Conversely, special deterrence and rehabilitation are more often sought in the fraud cases (see Table 14).<sup>22</sup> While these are the goals expressed as most important in specific cases, most judges also consider at least one other goal somewhat applicable.<sup>23</sup>

TABLE 14

AVERAGE PERCENTAGES OF CASE-SPECIFIC GOALS BY TYPE  
OF CRIME

(Number of Observations)	TOTAL	BANK	FRAUD
	CASE	ROBBERY	SCENARIOS
	SCENARIOS	SCENARIOS	SCENARIOS
	(923)	(503)	(420)
CASE-SPECIFIC GOALS:	%	%	%
General deterrence.....	33	38	28
Special deterrence .....	22	18	26
Incapacitation .....	19	22	15
Rehabilitation .....	15	11	21
Retribution .....	11	11	11

Does the intended purpose of the sentence influence the sentence itself—the amount of prison time, supervised time, or fine? The goal sought by the judge strongly influences the amount of prison time imposed (see Table 15). The influence of case-specific goals on the amount of supervised time and fine is less pronounced. Such influence is largely limited to the fraud cases (see Tables 16 and 17). In half of the cases studied, the goal of incapacitation yields significantly longer prison terms, while the goal of rehabilitation yields lighter prison sentences. The desire to rehabilitate is also associated with more extensive periods of supervision and lighter fines. The goals of retribution and deterrence are somewhat more likely to prompt contracted periods of supervision and heavier fines.

<sup>22</sup> These results contrast somewhat with those presented in Mann, Wheeler & Sarat, *Sentencing the White Collar Offender*, 17 AM. CRIM. L. REV. 479 (1980).

<sup>23</sup> A further qualifier on these findings is that judges stated these goals *post hoc*. This analysis does not demonstrate conclusively that the articulation of a goal is a causal force in the formulation of a sentence.

TABLE 15

## THE EFFECT OF CASE-SPECIFIC GOALS ON PRISON TIME

OFFENSE/OFFENDER CASE SCENARIOS	NUMBER OF OBSER- VATIONS	SIGNIFI- CANCE LEVEL OF F- TEST	IF SIGNIFICANT:	
			GOAL ELICITING HEAVIEST SENTENCE	GOAL ELICITING LIGHTEST SENTENCE
1 - Bank robbery <sup>a</sup>	101	***	Incapacitation	Rehabilitation
2 - Fraud	78	**	Special Deterrence	Rehabilitation
3 - Bank robbery	96	**	Incapacitation	Rehabilitation
4 - Fraud	81	**	Incapacitation	Rehabilitation
5 - Bank robbery	69	*		
6 - Fraud	56	*		
7 - Bank robbery	60	*		
8 - Fraud	53	*		
9 - Fraud	36	*		
10 - Bank robbery	40	*		
11 - Fraud	43	***	Incapacitation	Special Deterrence
12 - Bank robbery	54	***	Incapacitation	Rehabilitation
13 - Fraud	48	**	Incapacitation	Rehabilitation
14 - Bank robbery	41	**	Incapacitation	Rehabilitation
15 - Fraud	25	*		
16 - Bank robbery	47	***	Incapacitation	Rehabilitation

<sup>a</sup> Refer to Table 3 for specific combinations of factors associated with each scenario.

\* = Not significant \*\* = Significant at .20 level \*\*\* = Significant at .05 level

Divergent perceptions of the appropriate case-specific goals apparently aggravate the problem of disparity. In their response to the survey, judges did indeed pursue quite different goals for identical cases and these disparate goals, in turn, resulted in disparate sentence decisions.

#### C. THE PSYCHOLOGICAL/COGNITIVE, CONTEXTUAL, AND ATTITUDINAL CORRELATES OF SENTENCE DECISIONS

To this point, the discussion has focused on the importance of case-related matters in the sentence decisionmaking process. The segment of decision variance that remains unexplained by the attributes of the case is labeled dissensus, or more simply, disparity. The analysis also indicated that sentence disparity is not a chance occurrence, whereby each judge, in a totally random and idiosyncratic fashion, reacts differently to a case. Instead, the survey uncovered patterns of disparity, namely: (a) interactions between certain offense/offender characteristics and the

TABLE 16

## THE EFFECT OF CASE-SPECIFIC GOALS ON SUPERVISED TIME

OFFENSE/OFFENDER CASE SCENARIOS	NUMBER OF OBSERVATIONS	SIGNIFICANCE LEVEL OF F-TEST	IF SIGNIFICANT:	
			GOAL ELICITING HEAVIEST SENTENCE	GOAL ELICITING LIGHTEST SENTENCE
1 - Bank robbery <sup>a</sup>	101	***	Rehabilitation	Retribution
2 - Fraud	78	***	Rehabilitation	General Deterrence
3 - Bank robbery	96	*		
4 - Fraud	81	***	Incapacitation	General Deterrence
5 - Bank robbery	64	*		
6 - Fraud	56	*		
7 - Bank robbery	60	*		
8 - Fraud	53	***	Rehabilitation	Retribution
9 - Fraud	36	*		
10 - Bank robbery	40	*		
11 - Fraud	43	*		
12 - Bank robbery	54	***	Rehabilitation	Special Deterrence
13 - Fraud	48	*		
14 - Bank robbery	41	***	Rehabilitation	General Deterrence
15 - Fraud	25	*		
16 - Bank robbery	47	*		

<sup>a</sup> Refer to Table 3 for specific combinations of factors associated with each scenario.

\* = Not significant \*\* = Significant at .20 level \*\*\* = Significant at .05 level

judge, and (b) divergences of opinion among judges about the goal of the sentence for each specific case.

This section poses additional questions about the sources of disparity by focusing on certain psychological/cognitive attributes of the judge—primarily, patterned divergences in perceptions of the severity of sentences and in predispositions to be harsh or lenient in sentencing—and on judges' general attitudes toward sentencing and the judicial environment in which sentencing decisions are made. The key point about these factors is that their relationship to the case itself is indirect at best. They constitute the cognitive and social filter through which data about the case must flow before a sentence decision is reached.

**TABLE 17**  
**THE EFFECT OF CASE-SPECIFIC GOALS ON FINE**

OFFENSE/OFFENDER CASE SCENARIOS	NUMBER OF OBSER- VATIONS	SIGNIFI- CANCE LEVEL OF F- TEST	IF SIGNIFICANT:	
			GOAL ELICITING HEAVIEST SENTENCE	GOAL ELICITING LIGHTEST SENTENCE
1 - Bank robbery <sup>a</sup>	101	__b	__b	__b
2 - Fraud	78	***	Retribution	Rehabilitation
3 - Bank robbery	96	__b		
4 - Fraud	81	***	Retribution	Rehabilitation
5 - Bank robbery	64	*		
6 - Fraud	56	*		
7 - Bank robbery	60	*		
8 - Fraud	53	***	Special Deterrence	Rehabilitation
9 - Fraud	36	*		
10 - Bank robbery	40	*		
11 - Fraud	43	*		
12 - Bank robbery	54	*		
13 - Fraud	48	*		
14 - Bank robbery	41	*		
15 - Fraud	25	*		
16 - Bank robbery	47	*		

<sup>a</sup> Refer to Table 3 for specific combinations of factors associated with each scenario.

<sup>b</sup> Fines not given for these cases.

\* = Not significant \*\* = Significant at .20 level \*\*\* = Significant at .05 level

*Psychological/Cognitive Factors: The Role of Judges' Perceptions of Sentence  
Severity and Predispositions to be Lenient or Harsh*

A revealing pattern cuts across Tables 15, 16 and 17—a kind of reciprocal relationship between the goals sought by judges and the type of sentence imposed. Thus, rehabilitative sentences tend to yield diminished prison time and increased supervised time, while retribution and deterrence are associated with heavier fines *and* contracted periods of supervision. This relationship is not particularly surprising, for it conforms to intuitive expectations. However, it is quite noteworthy as an indicator of an important feature of sentence decisionmaking, namely the cognitive process whereby judges simultaneously weigh both the significance of case attributes *and* the relative severity of the sentence options available. These latter considerations add to the complexity of sentence decisions and constitute an additional potential source of disparity.

No single agreed-upon metric exists for the severity of the sentences



TABLE 18

DETAILED ANALYSIS OF VARIANCE RESULTS: EFFECTS OF  
OFFENSE/OFFENDER CHARACTERISTICS,  
JUDGES/ENVIRONMENT AND  
INTERACTIONS ON COMPOSITE MEASURE OF SENTENCE SEVERITY

SOURCE OF VARIANCE	UNIQUE VARIANCE	SUM OF SQUARES	MEAN SQUARES	F- TEST	SIGNIFICANCE
Offense/Offender Characteristics (main effects)	20%	1251	83	161	< .001
Judge/Environment (main effects)	65%	4087	33	64	< .001
Interaction Effects (between of- fense/offender characteristics <i>and</i> between judge and offense/offender characteristics)	15%				

that judges impose. Judges appear to perceive the severity of any given component of a sentence differently. They may also have markedly different perceptions of the severity equivalences across the various components of a sentence. For example, how severe is six months in prison as contrasted with two years of supervision? Given these differential perceptions, two judges quite possibly could reach identical conclusions about the seriousness of each of the hypothetical cases presented in the survey, but impose radically different sentences<sup>24</sup> because they adhere to radically different notions about the sentences' severity. Conversely, it is possible that identical sentences given by two judges are a deceptive indication of consensus, because differences between their perceptions of the seriousness of the case may be balanced by an opposite view about the severity of the sentence.

To address this issue of sentencing tradeoffs, we conducted an analysis of perceived severity beginning with the creation of a composite measure of sentence severity for each judge using a micromodeling procedure. Micromodeling consists of an examination of each judge's perception of the severity of a particular sentence using as a data base their ratings of various sentence combinations. This composite sentence was then used as a dependent variable in an analysis of variance predicting the sentences a judge imposed. The results appear in Table 18. Twenty percent of the decision variance is attributable to the offense/offender characteristics (main effects), and 15% is due to interactions among the offense/offender characteristics *and* between judges and offense/offender characteristics. The balance of the decision variance—a substantial 65%—is attributable to the judge given the assumptions of our design.

This confirms and enlarges on previous findings. Judges do indeed

<sup>24</sup> Radically different sentences to observers are not necessarily different at all to judges.

give different sentences for the same cases. The amount of dissensus among judges is quite significant. In fact, the effect of the individual judge on the composite sentence imposed is greater than it is on each component separately. This suggests that one element of sentencing guidelines might be to agree upon the units of exchange between alternative components of sentences.

Having established that alternative tradeoff patterns among sentencing alternatives account for some interjudge sentence disparity, we are still left with the question: To what degree is sentence disparity attributable to differences in the perceived severity of sentences rather than in the perceived seriousness of the case? A related question deals with patterns of actual leniency/harshness of sentences: To what degree is sentence disparity attributable to basic predispositions, or tendencies, among judges to give sentences that are more or less harsh? The answers to these queries are laden with policy implications. If discrepant perceptions of sentence severity and predispositions to sentence harshly or leniently contribute measurably to disparity, the problem of sentence disparity is appropriately addressed by seeking consensus on the desired underlying metric of sentence severity. This task could prove infinitely more difficult than the development of guidelines because perceptions of sentence severity are value-charged and psychologically and socially determined.

To explore these questions, we conducted another stage of the analysis in which: (a) judges were categorized based on their perceptions of the relative severity of sentences; (b) they were further categorized based on the relative actual harshness or leniency of their sentences; and (c) the effects of these two dimensions on sentences were estimated. The first step of this analysis was to calculate the average sentences (prison time, supervision, and fine) across all judges for each of the sixteen scenarios. These average sentences were then fed through the judge-specific equations that related the three components of sentences in order to estimate *each* judge's perception of the relative severity of the *average* sentence. This was repeated for all sixteen of the hypothetical cases. The sixteen severity estimates were then averaged for each judge, yielding a reliable estimate of each judge's perception of the relative severity of different types of sentences.

This exercise enables us to classify judges according to their *perceptions* of the severity of sentences. It enables us to answer questions such as: Does Judge X believe that three years in prison is more or less severe than Judge Y does? Or, more complexly, does Judge X believe that three years of prison, plus one year of supervision, plus \$1,000 fine is more or less severe than Judge Y does? However, it does not enable us to classify judges according to their *actual leniency or harshness* in

**TABLE 19**  
**NINE-WAY CLASSIFICATION MATRIX FOR OVERALL PERCEIVED SENTENCE SEVERITY AND LENIENCY OR HARSHNESS OF SENTENCES GIVEN IN SCENARIOS**

		OVERALL PERCEIVED SENTENCE SEVERITY				
		High	Medium	Low		
OVERALL LENIENCY/ HARSHNESS IN RESPONSE TO SCENARIOS	HARSH	*	15.097	13.624	12.770	13.830
		**	-0.050	-0.320	0.371	0.828
		***	0.527	0.223	0.365	
		( )	(13)	(14)	(14)	
	MEDIUM	*	14.284	13.070	11.938	13.095
		**	-0.129	-0.140	0.268	0.093
		***	0.201	0.149	0.262	
		( )	(14)	(14)	(14)	
	LENIENT	*	13.577	12.655	10.011	12.081
**		0.179	0.460	-0.639	-0.921	
***		0.213	0.146	1.358		
( )		(14)	(14)	(13)		
Mean		14.319	13.116	11.471		
Effect		1.317	0.114	-1.431		

\* = Mean \*\* = Effect \*\*\* = Standard deviation ( ) Number in group.

sentences. This is an important distinction. Two judges with different perceptions of sentence severity may nevertheless make quite similar sentence decisions.

We accounted for the harshness/leniency factor by classifying judges into a two-dimensional, nine-cell matrix. The first dimension of the matrix took each judge's perceived sentence severity score and assigned each judge to one of three levels of perceived severity. For each of these levels, judges were further classified into one of three levels of relative leniency or harshness. Thus, judges who regard sentences as being rather severe were subdivided into those whose sentences for the sixteen cases were relatively harsh, those whose sentences were relatively light, and those whose sentences fell in the middle of the spectrum. The nine-way classification is displayed in Table 19.

While the procedure used to develop this classification system is

TABLE 20

REPEATED MEASURE ANALYSIS OF VARIANCE RESULTS:  
EFFECTS OF OFFENSE/OFFENDER CHARACTERISTICS,  
OVERALL PERCEPTIONS OF SENTENCE SEVERITY AND  
OVERALL ACTUAL LENIENCY OR HARSHNESS  
ON SENTENCE DECISIONS

<u>SOURCE OF VARIANCE</u>	<u>UNIQUE VARIANCE</u>	<u>F-TEST</u>	<u>SIGNIFICANCE</u>
Offense/Offender Characteristics	14%	161	< .001
Judge:			
- Sentence severity perceptions	39%	296	< .001
- Actual leniency/harshness	16%	120	< .001
- Interaction: severity perceptions and leniency/harshness	3%	13	< .001
Residual Variance and Other Interactions	28%		

complicated, this should not obscure the straightforwardness of its underlying logic and its power as an analytic tool. The question that prompted this entire sequence of analyses can now be posed in more refined form. To what degree is sentence dissensus the result of (a) divergent perceptions of the severity of the sentences imposed, (b) patterned predispositions of judges to be relatively harsh or lenient, and (c) an interaction between perceptions of severity and predispositions toward harshness or leniency? Perceived severity, relative harshness or leniency, and case characteristics were entered as *independent* variables into a repeated measures analysis of variance to determine their effects on sentencing decisions.<sup>25</sup> The results appear in Table 20. Thirty-nine percent of the decision variance is linked to divergent perceptions of the severity of sentences, and another 16% is attributable to judges' general predispositions to be relatively harsh or lenient when sentencing. The cumulative importance of these perceptions and practices of the judge overwhelm the importance of case attributes in the sentence decisions made for the sixteen hypothetical cases.

<sup>25</sup> The case attributes were entered as sixteen levels, for the six hypothetical cases, in the repeated measures design. This isolated the variance due to case attributes from the variance due to perceived severity and relative harshness.

Seemingly, then, the disparities we observed in sentences are not random events. Rather, a substantial portion of this dissensus is attributable to divergent perceptions of the severity of the sentences that are being meted out and predispositions on the part of judges to be more or less harsh when sentencing.

*The Role of Factors that Have No Direct Bearing on the Case*

Up to this point, the discussion has focused on factors that are either directly related to a case—the offender's background, the nature of the offense, and the goal that the judge hopes to achieve with the sentence—or that are directly related to the nature of the sanction—notably the relative severity and harshness of the judges. We now turn to four sets of factors that have little or no direct bearing on the case. The first set includes the overall goal orientation of the judge (as opposed to case-specific goals discussed previously) and the judge's perception of how satisfactorily the federal criminal justice system achieves these goals. We analyzed the following orientations: general deterrence, special deterrence, rehabilitation, incapacitation, retribution/deservedness, and restitution. The second set includes the judge's evaluation of the general quality of the federal sentencing process and his or her perception of the degree to which sentence disparity constitutes a problem for the federal criminal justice system. The third set encompasses background characteristics of the judge: political ideology (self-identification as liberal or conservative); career variables, such as the number of years served as a federal judge and years in the legal profession; the type of community in which the judge was raised (urban/suburban/rural); and the judge's race. The fourth factor is the region of the country in which the judge's jurisdiction is located—with judges separated into four major sections of the country. These variables were entered as independent variables into a step-wise multiple regression in which the dependent variable was mean prison time given across all sixteen cases.<sup>26</sup> The results appear in Table 21.<sup>27</sup>

Approximately 40% of the variance in prison time sentences is attributable to these factors—a surprising and important finding in view

---

<sup>26</sup> Thus this analysis disregarded the role of offense/offender factors. The analysis reported here is for a subsample of 116 judges who received all of these questions. There were some missing observations for a number of respondents, but analysis of judges for whom there were not missing observations yielded parallel results.

<sup>27</sup> To this point, all analyses have involved analysis of variance or cross-statistics. The primary reasons for the use of ANOVA has been (a) the orthogonality in the design of the sentence decisionmaking questions; and (b) the fact that the case-specific goals were nominal variables. Conversely, the shift to regression analysis in this section stems from two considerations: (a) the multicollinearity among the independent variables; and (b) the fact that the independent variables are nominal and scalar.

**TABLE 21**  
**THE EFFECT OF NONCASE FACTORS ON PRISON TIME**

<u>JUDGE PERCEPTIONS OR CHARACTERISTICS</u>	<u>INCRE- MENTAL VARIANCE</u>	<u>B- COEFFICIENT</u>	<u>BETA COEFFICIENT</u>	<u>T- TEST</u>	<u>SIGNIF- CANCE</u>
The system's achievement of special deterrence goal	7.3%	-29.671	0.319	-3.72	.01
Importance of incapacitation goal	5.6%	18.079	0.245	3.26	.01
The system's achievement of rehabilitation goal	4.9%	17.802	0.231	3.05	.01
Region of the country	4.8%	10.924	0.236	3.02	.01
Political ideology	3.5%	11.644	0.196	2.59	.05
The system's achievement of retribution goal	2.8%	-13.344	-0.180	-2.29	.05
Importance of general deterrence goal	2.4%	13.621	0.177	2.12	.05
Other variables combined	8.2%				
Multiple R <sup>2</sup> = .395					

of their remoteness from the case. Even more interesting, however, is the particular variables that come into play. All of the significant variables except for region of the country represent value orientations or attitudes rather than background characteristics of the judge.<sup>28</sup> Of these value orientations, political ideology appears to be the one that is most removed from the issue of criminal sanctions, with those judges who regard themselves as liberal giving less prison time than their conservative counterparts. The bulk of the variance, however, relates to the judge's perceptions regarding the overall goals of sentences—both the importance attached to the goals and evaluations of how well the goals are being accomplished. Thus, the more important the goals of incapacitation and general deterrence, particularly the former, the greater the prison time given by the judge. Judges' views about which goals the system actually is accomplishing are associated with different sentences. Judges who believe that the system is helping to rehabilitate criminals give heavier sentences on average. On the other hand, those who believe the system is achieving special deterrence and retribution give lighter sentences on average. Southern judges give substantially heavier sentences than others.

These findings challenge the view that guidelines will impair the

---

<sup>28</sup> The failure of these background variables to enter the regression equation is primarily a consequence of their modest relationship with the dependent variable rather than multicollinearity with other independent variables.

individualized justice of the current sentencing process. According to this view, the judge is a complex and meticulous information processor. Every nuance of the case is considered and weighed against extensive data on the offender's past and subtle hypotheses about his or her probable future criminal behavior. How can guidelines replicate this complexity and produce sentence decisions as wise and just as those produced by judges? These analyses do not discredit the view that sentencing decisions are complex, nor the inference that those who design guidelines should strive toward commensurate complexities. Quite the contrary, findings reported earlier in this paper, as well as the results from analyses of actual sentence decisions, reinforce these claims.

Yet, the conclusion that emerges from these analyses is that the characteristics of the case and offender tell only part of the story about sentence decisions. The judge's general value orientation and, to a lesser degree, his or her regional environment, also imprint decisions about the amount of time a defendant should spend in prison.

To a significant degree, prison sentences given for the sixteen hypothetical cases represent specific applications, or projections, of a judge's political philosophy and his or her core values about the desired functions of criminal sanctions. Then, too, the tendency of prison sentences to expand or contract depending on how well the system is believed to achieve certain goals suggests that particular sentences also serve as remedies for the perceived failures in the system at large. In a sense, then, judges appear already to be using self-made guidelines in reaching sentence decisions. The shortcoming of these self-made guidelines is the lack of consistency of composition across judges.

#### IV. SUMMARY

The survey results suggest a considerable need for guidelines that would structure judicial sentencing discretion. Disparity is a widespread phenomenon. At least for the hypothetical cases studied in the survey, substantial dissensus exists among judges about the sentences that convicted offenders should serve. When one considers that only one type of disparity was measured in the survey (*i.e.*, dissensus among judges and *not* instability over time), the level of real world disparity is likely to actually exceed that which is reported here.

Moreover, the survey sheds light on the nature and sources of disparity. Part of the problem of disparity reflects unexplainable, perhaps random, differences of judgment. However, a significant amount of disparity is not at all random; judges do not extemporize sentences each time they reach a sentence decision. Rather, disparities are anchored in patterned differences regarding such matters as: (a) overall value orien-

tations about the functions of criminal sanction; (b) judgments about the appropriate goal of case-specific sentences; (c) perceptions about the severity and the sentences themselves; (d) predispositions to sentence in a relatively harsh or lenient fashion; and (e) perceptions of the seriousness of particular attributes of a case.

These patterns are related not just to the case of the offender who is being sentenced, but bear a significant relationship to the characteristics of the judge who is doing the sentencing. Thus, guidelines will yield two general consequences for the criminal justice system. The first manifest consequence is to diminish disparity by specifying, more or less precisely, the importance that each case and offender attribute should assume in arriving at a sentence. However, by enhancing the order and structure of these decisions as they relate to case/offender characteristics, guidelines will have an equal and opposite effect on the hidden order that characterizes current decisionmaking. This is the second, latent function of guidelines—to help make explicit norms and practices that currently guide sentence decisions. Guidelines need not be viewed narrowly as a means of imposing order on chaos; they can be viewed more broadly as a mechanism for replacing one type of order, which is created by the judge, with another type, which is created by the sentencing commission.