Fall 1998

The Structure of Punishment Norms: Applying the Rossi-Berk Model

Joseph E. Jacoby

Francis T. Cullen

THE STRUCTURE OF PUNISHMENT NORMS: APPLYING THE ROSSI-BERK MODEL

JOSEPH E. JACOBY* & FRANCIS T. CULLEN**

I. INTRODUCTION

Over the past two decades research on the nature of public attitudes toward crime and punishment has grown substantially.[1] Much of this research has been descriptive, reporting "what the public thinks" about various crime-related issues. When conceptual frameworks are used to explore the organizing principles of public opinion, they are largely dominated by the ongoing debate between consensus and conflict theory.[2] Researchers typically comment on the implications of their findings for this debate: Do citizens fundamentally agree or disagree

---

* Associate Professor, Department of Sociology, Bowling Green State University
** Distinguished Research Professor, Division of Criminal Justice, University of Cincinnati

[1] See generally Julian V. Roberts, Public Opinion, Crime, and Criminal Justice, in CRIME AND JUSTICE: A REVIEW OF RESEARCH 99 (Michael Tonry ed., 1992); Loretta J. Stalans & Arthur J. Lurigio, Editors' Introduction, Public Opinion About the Creation, Enforcement, and Punishment of Criminal Offenses, 39 AM. BEHAVIORAL SCIENTIST 369 (1996); Mark Warr, Public Perceptions and Reactions to Violent Offending and Victimization, in 4 UNDERSTANDING AND PREVENTING VIOLENCE: CONSEQUENCES AND CONTROL 1 (Albert J. Reiss, Jr. & Jeffrey A. Roth eds., 1995) [hereinafter Warr, Public Perceptions].

[2] As is well known, consensus theory argues that there is widespread agreement in society about what should or should not be illegal. Laws and legal sanctions are thus seen as reflecting the "will of the people." In contrast, conflict theory contends that groups in society, based on competing political and/or economic interests, differ in their views of what should be declared illegal and of what penalties lawbreaking should elicit. Accordingly, laws and legal sanctions are seen as reflecting the ability of competing groups to exercise power and have their interests represented in the criminal law and in the administration of the criminal justice system. See, e.g., Charles W. Thomas et al., Public Opinion on Criminal Law and Legal Sanctions: An Examination of Two Conceptual Models, 67 J. CRIM. L. & CRIMINOLOGY 110 (1976); Mark Warr et al., Contending Theories of Criminal Law: Statutory Penalties Versus Public Preferences, 19 J. RES. CRIME & DELINQ. 25 (1982) [hereinafter Warr et al., Contending Theories].

on the rules that should govern society? Findings of attitudinal agreement or consensus about the seriousness of crime or appropriate punishments for offenders are taken as evidence in favor of consensus theory; cleavages in opinion between social groups—especially along race and class lines—are taken as support for conflict theory. Only rarely, however, do researchers test a full range of hypotheses systematically derived from these competing theories.

Although this body of research has value, the dominance of the consensus/conflict debate may have stifled the development of alternative approaches to examining crime attitudes. It is instructive that the empirical studies attempting to resolve whether consensus or conflict best describes a normative domain typically produce ambiguous results.[3] Researchers seldom find either universal normative consensus or consensus clearly differentiated by interest group membership. Instead they find variation—more intra-individual variation than supports consensus theory and less intra-group variation than supports conflict theory.

It is possible that the "ambiguous" findings of analysis oriented around the consensus/conflict debate are a consequence of the limited vision of both perspectives. The patterns of norms that exist in the real world are not "ambiguous," though they are not explained adequately by either of the dominant perspectives. These perspectives may oversimplify the range of potential normative structures (i.e., normative structures may exist outside the types that are logically derived from either consensus or conflict theory).

Accordingly, we suggest that criminologists studying the structure of crime attitudes should move beyond consensus and conflict theories as guides for their research by employing more comprehensive, sophisticated models. To this end, we use an

---

[3] *See, e.g.,* Stalans & Lurigio, *supra* note 1 (Stalans and Lurigio cite evidence and arguments that support both the conflict and consensus models. In support of the consensus model, they cite numerous studies revealing public consensus around which behaviors are harmful and wrong, as well as widespread public support for the courts and police. In support of the conflict model they site the sharp division along racial lines of the justice of the verdict in the O.J. Simpson murder trial.)

analytical model introduced by Peter Rossi and Richard Berk.[4] The Rossi-Berk model offers a general sociological approach to investigating and mapping normative structures. We apply this model to data gathered through a national survey of public attitudes toward the punishment of street crimes.

The Rossi-Berk model, which is described in detail below, has advantages over both consensus and conflict theory as a guide to exploring normative structures:

1. The Rossi-Berk model is rooted in empirical observation, not ideology. Unlike both consensus and conflict theories, the Rossi-Berk model's validity does not depend on whether consensus or dissensus exists in any normative domain. Scholars embracing either of these competing theories, having a stake in finding or not finding consensus in public attitudes, must treat findings anomalous to their paradigm as somehow not reflecting reality. Consensus theorists dismiss inconsistencies between public opinion and public policy as products of misinformation caused by the entertainment media.[5] Conflict theorists dismiss evidence of widespread public agreement on some issues, asserting that survey respondents who express attitudes divergent from their "class interests" are exhibiting "false consciousness." Though the existence of false consciousness may be impossible to test empirically, it is an effective rhetorical response to evidence that challenges the validity of the conflict model.

2. The purpose of the Rossi-Berk model is to provide a comprehensive tool that may be used to determine whether norms exist and what those norms are in *any* normative domain. Consensus theory, which is rooted in the sociological theory of structural-functionalism, has little to say about normative domains that are not clearly connected to common interests that contribute to the survival of the society; conflict theory has little to say about normative domains that are not clearly related to

---

[4] Peter H. Rossi & Richard A. Berk, *Varieties of Normative Consensus,* 50 AM. SOC. REV. 333, 333-47 (1985) [hereinafter Rossi & Berk, *Varieties*]; Peter H. Rossi & Richard A. Berk, *A Conceptual Framework for Measuring Norms, in* THE SOCIAL FABRIC: DIMENSIONS AND ISSUES 77 (James F. Short, Jr. ed., 1986) [hereinafter Rossi & Berk, *Conceptual Framework*].

[5] Stalans & Lurigio, *supra* note 1, at 370.

groups' political interests. Many normative domains, apparently, do not lend themselves to either consensus or conflict analysis because they are unrelated to either society's survival or groups' political interests.[6]

3. The Rossi-Berk model is designed to deal with the full spectrum of variability of normative judgments—from absolute consensus to absolute dissensus—wherever it exists—within individuals, between individuals in the same group, and between groups of individuals. The model covers the entire range of logically possible normative structures. Neither consensus nor conflict theory predicts the wide variety of patterns of public opinion that actually exist.

4. The Rossi-Berk model is "comfortable" with the continuum of consensus-dissensus that occurs in the real world. Unlike both consensus and conflict theory, the Rossi-Berk model does not require the arbitrary creation of dichotomous categories labeled "consensus" and "dissensus."

5. Neither consensus nor conflict theory suggests any particular methods for testing its validity. The precise language of the Rossi-Berk model provides clear guidance to empirical application of the model through the measurement of variation of normative judgments within each individual, between individuals, and among groups of individuals.

6. Applying the model to a normative domain clarifies how that normative domain is structured relative to other normative domains.

7. Applying the model to the same normative domain in many cultures could clarify whether normative structures are universal or unique to each culture.

8. Applying the model to a large number of normative domains creates the possibility of theorizing about norms at a higher level of abstraction, by revealing whether all normative

---

[6] Conflict and consensus theorists, of course, do not write about phenomena they cannot explain from their perspectives. Examples of normative domains that seem unrelated to society's survival or any group's political interests (and are therefore incapable of being explained by either consensus or conflict analysis) might include the public's preferences for different styles of clothing and varieties of food.

domains are structured similarly or some domains have unique structures.

9.   The Rossi-Berk model exists outside the structural-functionalism/conflict debate, but applying the model to specific normative domains produces empirical findings that can answer questions raised in that debate.

10.   The Rossi-Berk model reconnects the study of crime and deviance to the field of sociology.[7] The study of normative structures uses the concepts and methods of sociology, and is of interest to sociologists studying all kinds of human behavior.

II. STUDYING NORMATIVE STRUCTURES

A. IMPORTANCE

The shape of normative consensus regarding criminal punishment has important implications for punishment policy and the very legitimacy of criminal justice institutions and processes: "[P]ublic opinion research can provide information about people's perceptions of the legitimacy of laws and the institutions that are designed to uphold, protect, and enforce them."[8] At a deeper level, however, it may be equally important to understand how people formulate their preferences about punishment.   What qualities of crimes, offenders, and victims do people consider relevant to punishment?   How do people combine the qualities they consider relevant, leading them to select a particular punishment?   In other words, what norms guide their choice of punishments?

An understanding of normative behavior is central to most social science conceptual schemes.[9] Norms identify deviant be-

---

[7] The field of sociology is concerned with the structure of social relations, generally.   The deterrence, rational choice, control, and biological theories that are currently prominent in criminology are not concerned with these broader issues, so criminological studies applying those theories do not inform the broader field.

[8] Stalans & Lurigio, *supra* note 1, at 371.

[9] *See generally* Judith Blake & Kingsley Davis, *Norms, Values, and Sanctions, in* HANDBOOK OF MODERN SOCIOLOGY 456 (Robert E.L. Faris ed., 1964); V. Lee Hamilton & Steve Rytina, *Social Consensus on Norms of Justice: Should the Punishment Fit the Crime?*, 85 AM. J. SOC. 1117 (1980); Robert F. Meier, *Norms and the Study of Deviance: A Proposed Research Strategy*, 3 DEVIANT BEHAV. 1 (1981); Terance D. Miethe, *Public Consensus on Crime Seriousness: Normative Structure or Methodological Artifact?*, 20 CRIMINOLOGY 515

havior and, relevant to our concerns, prescribe punishments for transgressors. The existence of normative standards is indicated when public opinion is characterized by high degrees of consensus and stability. Knowledge of this normative structure, however, is complicated by the difficulties associated with identifying empirically stable, enduring public preferences. In particular, research on norms has been hindered by two issues: how to measure them, and how to distinguish them from more idiosyncratic preferences.

B. CONSEQUENCES OF DIFFERENT MEASUREMENT STRATEGIES

As noted above, a growing body of research has emerged on public attitudes toward the punishment of crime. Although these studies have provided useful insights about punishment norms, they have tended to be limited by one or more methodological problems. First, most public opinion polls about punishment have called for general responses to very complex questions stated in simple terms.[10] They have not evaluated subtleties in judgments. People have been asked, for example: "In general, do you think the courts in your area deal too harshly, or not harshly enough with criminals?"[11] When asked this question, 85% of respondents to a 1994 national poll responded "not harshly enough," revealing general dissatisfaction with judges' sentencing practices.[12] Respondents were not asked what they believed such practices to be or what practices they preferred. This apparent consensus, therefore, reveals neither respondents' policy preferences nor the norms underlying those preferences.

This criticism applies in particular to conventional polling techniques (e.g., Gallup Polls), whose results often are disseminated widely in the media, strongly influencing policy makers'

(1982); John F. Stolte, *The Formation of Justice Norms*, 52 AM. SOC. REV. 774 (1987); Mark Warr et al., *Norms, Theories of Punishment, and Publicly Preferred Penalties for Crimes*, 24 SOC. Q. 75 (1983) [hereinafter Warr et al., *Norms*].

[10] Michael G. Turner et al., *Public Tolerance for Community-Based Sanctions*, 77 PRISON J. 6, 6-9 (1997).

[11] *Id.* at 7.

[12] BUREAU OF JUSTICE STATISTICS, U.S. DEP'T JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS—1995, at 173 (Kathleen Maguire & Ann L. Pastore eds., 1996) [hereinafter SOURCEBOOK].

understanding of "public opinion."[13]   The broad questions posed in traditional public opinion polls reveal little about normative structure.   Such questions de-contextualize punishment choices from real-life situations where punishment is applied.   They do not simulate actual decision making by people confronted with real punishment decisions (in the courtroom, for example), so they cannot reveal the norms guiding those real decisions.   General questions about punishment tend to elicit very punitive responses characterizing the public's general fear of crime and dissatisfaction with the criminal justice system, rather than carefully thought-out punishment preferences appropriate for specific situations.[14]

Indeed, Thomson compares the results of typical public opinion polls with other types of studies.[15] Thomson claims that public opinion polls that provide little information about specific crimes appeal to fear and outrage, eliciting emotional responses, much like the reactions of a vigilante mob: "Given something like the distorted and insufficient information and visceral incentives of a crowd, they respond something like a crowd. Hang the criminals. Impeach the judges. Build more prisons. Hang the criminals."[16]

Much academic research, as might be anticipated, is more sophisticated and more valuable in furnishing information on punishment norms; but it is not without limitations. Conventional polling techniques have the potential advantage of national coverage, but with the exception of the National Survey

---

[13] *See* JULIAN V. ROBERTS & LORETTA J. STALANS, PUBLIC OPINION, CRIME, AND CRIMINAL JUSTICE 293-94 (1997). *See also* Timothy J. Flanagan, *Public Opinion and Public Policy in Criminal Justice, in* AMERICANS' VIEW OF CRIME AND JUSTICE: A NATIONAL PUBLIC OPINION SURVEY 151, 152-54 (Timothy J. Flanagan & Dennis R. Longmire, eds., 1996).

[14] Brandon K. Applegate et al., *Assessing Public Support for Three Strikes-and-You're Out Laws: Global versus Specific Attitudes,* 42 CRIME & DELINQ. 517, 528-30 (1996) [hereinafter Applegate et al., *Assessing Public Support*]; Anthony N. Doob & Julian V. Roberts, *Social Psychology, Social Attitudes, and Attitudes Toward Sentencing,* 16 CANADIAN J. BEHAV. SCI. 269, 277 (1984).

[15] *See generally* Douglas R. Thomson, Discordant Images of Public Sentiments Toward Criminal Sanctions (1988)(Paper presented at the Annual Meetings of the Law and Society Association in Vail, Colorado, on file with author).

[16] *Id.* at 20.

of Crime Severity,[17] virtually all these studies are based on student, community, or at most, state-wide samples.[18] Further, with few exceptions,[19] many studies do not ask respondents to specify the actual punishments they would prefer for a given criminal event. Instead, they use rating tasks—such as response scales measuring seriousness, general punitiveness, or the fairness of a punishment meted out—that are one step removed from specific judgments on concrete sentencing preferences.[20]

Most important, the standard design used in traditional academic research—asking respondents to judge the seriousness of, or apply punishments to, lengthy lists of criminal offenses—is potentially limited by the core problem found in conventional polling techniques: decontextualized ratings that do not approximate decision making in real-life situations. Thus, respondents are given limited information about a criminal event—in this case, information that varies primarily along only two dimensions, the type of crime and degree of harm caused by the offense.[21] As a result, the data produced by this

---

[17] *See* MARVIN E. WOLFGANG ET AL., U.S. DEP'T OF JUSTICE, THE NATIONAL SURVEY OF CRIME SEVERITY (1985).

[18] *See, e.g.,* Alfred Blumstein & Jacqueline Cohen, *Sentencing of Convicted Offenders: An Analysis of the Public's View,* 14 L. & SOC. REV. 223 (1980); Alexis M. Durham, *Crime Seriousness and Punitive Severity: An Assessment of Social Attitudes,* 5 JUST. Q. 131 (1988); Sandra S. Evans & Joseph E. Scott, *The Seriousness of Crime Cross Culturally,* 22 CRIMINOLOGY 39 (1984); Hamilton & Rytina, *supra* note 9; Peter H. Rossi et al., *The Seriousness of Crimes: Normative Structure and Individual Differences,* 39 AM. SOC. REV. 224 (1974) [hereinafter Rossi, et al., *Seriousness of Crimes*]; Warr et al., *Norms, supra* note 9.

[19] *See, e.g.,* Blumstein & Cohen, *supra* note 18; L. Thomas Winfree, Jr. & Larry E. Williams, *Understanding Public Support for Punitive Criminal Sanctions: Psychological and Sociological Views of the 'Outlier' Phenomenon,* 6 SOC. SPECTRUM 179 (1986).

[20] *See e.g.,* Francis T. Cullen et al., *The Seriousness of Crime Revisited: Are Attitudes Toward White-Collar Crime Changing?,* 20 CRIMINOLOGY 83, 85-87 (1982) [hereinafter Cullen et al., *Seriousness of Crime Revisited*]; Colin Goff & Nancy Nason-Clark, *The Seriousness of Crime in Fredericton, New Brunswick: Perceptions Toward White-Collar Crime,* 31 CANADIAN J. CRIMINOLOGY 19, 22-24 (1989); Rossi et al., *Seriousness of Crimes, supra* note 18, at 227-29; WOLFGANG ET AL., *supra* note 17, at 2.

[21] *See, e.g.,* Blumstein & Cohen, *supra* note 18, at 228; Roland Chilton & Jan De-Amicas, *Overcriminalization and the Measurement of Consensus,* 59 SOC. & SOC. RES. 319, 323 (1975); Cullen et al., *Seriousness of Crime Revisited, supra* note 20, at 88-91; Francis T. Cullen et al., *Dissecting White-Collar Crime: Offense Type and Punitiveness,* 9 INT'L J. APPLIED & COMP. CRIM. JUST. 15, 20-21 (1985); Goff & Nason-Clark, *supra* note 20, at 29; Darnell F. Hawkins, *Perceptions of Punishment for Crime,* 1 DEVIANT BEHAV. 193, 198 (1980); Rossi et al., *Seriousness of Crimes, supra* note 18, at 228-29; Peter G. Sinden, *Per-*

research are limited in their ability to illuminate the way people make real choices about complicated issues, where highly differentiated normative structures with conflicting principles are involved.[22]

Criminologists have recognized this potential limitation and have used vignette methodology to provide respondents with a rating task that approximates more closely the information available in real-life crime events.[23] The first generation of vignette research, however, was faced with the daunting problem that varying too many dimensions in the vignette would produce exponential growth in the number of vignettes respondents would have to rate. Accordingly, these studies tended to vary only a few theoretically salient factors (e.g., culpability and harm).[24]

Factorial survey methodology, however, overcomes this problem by permitting multiple dimensions of a crime event to vary randomly across vignettes to be rated.[25] Although the complexity of real-life crime events can never be duplicated fully, factorial design vignettes operationalize these events more ade-

---

ceptions of Crime in Capitalist America: The Question of Conciousness Manipulation, 13 SOC. FOCUS 75, 79 (1980); Thomas et al., *supra* note 2, at 112-13; Winfree & Williams, *supra* note 19, at 191-92; WOLFGANG ET AL., *supra* note 17, at 247-50.

[22] David Indermaur, *Public Perceptions of Sentencing in Perth, Western Australia*, 20 AUSTRALIAN & NEW ZEALAND J. CRIMINOLOGY 163, 176-80 (1987); Douglas A. Thomson & Anthony J. Ragona, *Popular Moderation versus Governmental Authoritarianism: An Interactionist View of Public Sentiments Toward Criminal Sanctions*, 33 CRIME & DELINQUENCY 337, 339-41 (1987).

[23] *See, e.g.*, Brandon K. Applegate et al., *Determinants of Public Punitiveness Toward Drunk Driving: A Factorial Survey Approach*, 13 JUST. Q. 57, 65 (1997) [hereinafter Applegate et al., *Determinants of Punitiveness*]; Applegate et al., *Assessing Public Support*, *supra* note 14, at 522-24; Turner et al., *supra* note 10, at 12-13.

[24] *See, e.g.*, James Frank et al., *Sanctioning Corporate Crime: How Do Business Executives and the Public Compare?*, 13 AM. J. CRIM. JUST. 139 (1989)(This study varied the culpability of the offender and the harm of the offense.); Valerie P. Hans & M. David Ermann, *Responses to Corporate versus Individual Wrongdoing*, 13 L. & HUM. BEHAV. 151 (1989)(This study varied whether the wrongdoing was done by an individual or a corporation.)

[25] PETER H. ROSSI & STEVEN L. NOCK, MEASURING SOCIAL JUDGMENTS: THE FACTORIAL SURVEY APPROACH 16 (1982) [hereinafter ROSSI & NOCK, MEASURING SOCIAL JUDGMENTS]; Applegate et al., *Determinants of Punitiveness*, *supra* note 23, at 63-64; Joop J. Hox et al., *The Analysis of Factorial Surveys*, 19 SOC. METHODS 493, 493-95 (1991); Turner et al., *supra* note 10, at 12-13.

quately and (arguably) introduce less bias into rating tasks, thus providing a better opportunity for assessing punishment norms.

It is instructive that other criminologists have recognized the value of the factorial design method in the study of punishment judgments. As Rossi, Simpson, and Miller note, offenders may:

> be regarded as complex social objects that vary from one another in many, often contradictory ways—crimes committed, losses or damages inflicted on victims, and social characteristics of both offenders and victims. Hence judgments about appropriate punishments for convicted criminals are a fitting subject for study through the factorial survey approach.[26]

To date, a few studies have employed this method to study punishment preferences, including research on "just punishments" in a Boston SMSA sample,[27] on "punishment repertoires" in American, Japanese, and Russian cities,[28] on sources of punitiveness toward drunk driving,[29] and on racial bias in support for capital punishment.[30] The most extensive use of factorial survey design to study punishment preferences was Rossi and Berk's evaluation of public support for the U.S. Sentencing Commission's sentencing guidelines.[31] This article reports the results of

---

[26] Peter H. Rossi et al., *Beyond Crime Seriousness: Fitting the Punishment to the Crime,* 1 J. QUANTITATIVE CRIMINOLOGY 59, 62 (1985) [hereinafter Rossi et al., *Beyond Crime Seriousness*].

[27] *See e.g.,* JoAnn L. Miller et al., *Perceptions of Justice: Race and Gender Differences in Judgments of Appropriate Prison Sentences,* 20 L. & SOC'Y. REV. 313 (1986) [hereinafter Miller et al., *Perceptions of Justice*]; JoAnn L. Miller et al., *Felony Punishments: A Factorial Survey of Perceived Justice in Criminal Sentencing,* 82 J. CRIM. L. & CRIMINOLOGY 396 (1991) [hereinafter Miller et al., *Felony Punishments*]; Rossi et al., *Beyond Crime Seriousness, supra* note 26.

[28] *See, e.g.,* Joseph Sanders & V. Lee Hamilton, *Is there a "Common Law" of Responsibility?,* 11 L. & HUM. BEHAV. 277 (1987) [hereinafter Sanders & Hamilton, *Common Law*]; Joseph Sanders & V. Lee Hamilton, *Legal Cultures and Punishment Repertoires in Japan, Russia, and the United States,* 26 L. & SOC'Y. REV. 117 (1992) [hereinafter Sanders & Hamilton, *Legal Cultures*].

[29] *See e.g.,* Applegate et al., *Determinants of Punitiveness, supra* note 23, at 57.

[30] *See e.g.,* Brandon K. Applegate et al., *Victim-Offender Race and Support for Capital Punishment: A Factorial Design Approach,* 18 AM. J. CRIM. JUST. 95 (1993).

[31] PETER H. ROSSI & RICHARD A. BERK, JUST PUNISHMENTS: FEDERAL GUIDELINES AND PUBLIC VIEWS COMPARED (1997) [hereinafter ROSSI & BERK, JUST PUNISHMENTS].

the first national, factorial survey study where respondents rated crime vignettes by imposing criminal sentences.

Before describing the specific methods used in the present study, we must first discuss the theoretical framework informing this study of punishment norms.

III. CLASSIFICATION OF NORMATIVE STRUCTURES

A. THE ROSSI-BERK MODEL

Rossi and Berk have elaborated criteria to classify normative structures from survey data.[32] In their scheme, norms are defined as statements of obligatory actions or evaluative rules. As rules governing action, norms specify what should be done in particular situations (e.g., "serious crimes should be punished").[33] As evaluative principles, norms state preference orders (e.g., "assault is more serious than larceny").[34] A normative domain is defined as the set of norms about a homogeneous domain of social action.[35] The normative domain addressed in this paper is punishment for common street crimes. The present research reports the first application of the Rossi-Berk model to the domain of criminal punishment. While the Rossi-Berk model has previously been applied to the study of crime seriousness,[36] it has not been applied to punishment evaluations.

This analysis of normative structures focuses on the way three components of norms vary among individuals and between population groups. The first of these components is *judgments*—in this case, the particular punishments selected for particular crimes. In relation to judgments, the analysis asks,

---

[32] Rossi & Berk, *Varieties, supra* note 4, at 336-44; Rossi & Berk, *Conceptual Framework, supra* note 4, at 84-100.

[33] Rossi & Berk, *Varieties, supra* note 4, at 333; Rossi & Berk, *Conceptual Framework, supra* note 4, at 77-78.

[34] WOLFGANG ET AL., *supra* note 17, demonstrate that consensus exists around the ranking of seriousness of crimes, based on the type of crime and severity of harm.

[35] Rossi & Berk, *Varieties, supra* note 4, at 335; Rossi & Berk, *Conceptual Framework, supra* note 4, at 80.

[36] *See* David Rauma, *The Context of Normative Consensus: An Expansion of the Rossi/Berk Consensus Model, with an Application to Crime Seriousness,* 20 SOC. SCI. RES. 1, 14-16 (1991).

What kind of punishment do people want to apply to offenders who commit particular crimes?

In addressing the conflict-consensus debate about the origin of law, we especially want to know whether, overall, consensus exists on the appropriate type and amount of punishment for each type of crime. From a social policy perspective, we want to know whether there is sufficient agreement on the kind of punishment to impose on criminal offenders, so that social policy could directly reflect "the will of the people" as expressed in such surveys.

We are also interested in the second component of norms, called *thresholds*—in this case we want to know whether people adhere to some internal scale of punishment severity, and whether all people use the same scale. In relation to the conflict-consensus debate, we want to know whether some subgroups of the population (e.g., rich or poor people) use distinctive scales. If subgroups use their own unique scales, and if the application of those scales would clearly benefit the subgroup, we would have strong evidence supporting the conflict perspective on punishment norms. These questions have important implications for social policy, as well. If sentencing laws are to be based on public opinion, there must be widespread agreement on the appropriate severity of punishment.

The final component of the analysis is *error*—in this case, error refers to the dispersion of punishment preferences around the population mean. If the dispersion is small, representing relatively minor disagreements about the kind and amount of punishment, we may claim that substantial consensus exists. If, on the other hand, there is wide dispersion, no such claim for consensus could be made, and the mean of punishment preferences will be an inadequate representation of the will of the people. Identification of normative domains through surveys is made difficult by pervasive measurement error. Though normative order may exist, searching for it with real data is always confounded by measurement error from several sources. Random inconsistencies in judgments of individuals create "noise," obscuring any underlying pattern. People, influenced by changing moods or recent experiences, judge the same situation dif-

ferently every time, and different people understand the rating task differently.

Rossi and Berk specified nine generalized normative structures, depending on how judgments, thresholds, and error vary.[37] The first four models all postulate *absolute* consensus with varying degrees of measurement error (i.e., everyone makes the same choice, with no variation in choices, and no structure to the error).

Model I—*Absolute Consensus and Uniformity*—Every person has a perfect understanding of the norms and subscribes to them to exactly the same degree, without any variation or error.[38] This would describe a situation where every person independently chooses exactly the same punishment (e.g., execution for all crimes and offenders), without any distinction among offenses and offenders.

Model II—*Absolute Consensus and Uniformity with Error Only*—As in Model I, everyone subscribes to the norms to exactly the same degree, but there are random variations in responses caused by different understandings or confusion about the task or variations in mood.[39]

Model III—*Absolute Consensus and Differentiated Judgments with No Error*—Respondents perceive that different situations call for different responses; they all make the same choices without any random variation.[40] If this model described the domain of punishment norms, respondents would all agree that different crimes required different punishments. They would also agree perfectly on the punishment to impose for every type of crime (e.g., all misdemeanor thefts should be punished by one year of probation, all robberies should be punished by five years in prison).

---

[37] Rossi & Berk, *Varieties, supra* note 4, at 336-44; Rossi & Berk, *Conceptual Framework, supra* note 4, at 84-100.

[38] Rossi & Berk, *Varieties, supra* note 4, at 337; Rossi & Berk, *Conceptual Framework, supra* note 4, at 86.

[39] Rossi & Berk, *Varieties, supra* note 4, at 337-38; Rossi & Berk, *Conceptual Framework, supra* note 4, at 86-87.

[40] Rossi & Berk, *Varieties, supra* note 4, at 338; Rossi & Berk, *Conceptual Framework, supra* note 4, at 87-88.

Model IV—*Absolute Consensus, Differentiated Judgments and Error*—As in Model III, respondents assign different responses to different situations, but, though they agree completely in their judgment of every situation, random variation in responses occurs because of mood changes or misunderstandings.[41]

These first four models can be safely ignored because in modern, complex societies absolute consensus on either generalized norms or specific applications is nonexistent.[42] Normative structures in such societies are more complicated; consensus, if it exists at all, is "relative." That is, people agree on the norms, but adhere to those norms with different degrees of intensity. In choosing punishments, for example, people may apply the same general norm that more serious crimes should be punished more severely. They come up with different punishments, however, because they have different "thresholds"—people's scale of punishments vary, with some people preferring consistently harsher punishments than do other people. Brief descriptions of the remaining Rossi-Berk Models, V through IX, all of which include relative consensus, are given below.

In Model V—*Relative Consensus, Differentiated Judgments, Varying Thresholds, and Error*—people do not agree on each judgment, but their disagreement is not random. Each disagrees, by some constant characteristic of that individual, from the average rating of the group.[43] Rossi and Berk refer to this constant as the individual's "threshold."[44] Thresholds represent individual variation in strengths of adherence to norms.

For example, most people would agree that robbery and burglary should be punished by imprisonment. They might also agree that robbery should be more harshly punished than burglary (i.e., their punishment judgments are differentiated by offense type). The periods of imprisonment they choose for

---

[41] Rossi & Berk, *Varieties, supra* note 4, at 338-39; Rossi & Berk, *Conceptual Framework, supra* note 4, at 88-89.

[42] Rossi & Berk, *Varieties, supra* note 4, at 339; Rossi & Berk, *Conceptual Framework, supra* note 4, at 89.

[43] Rossi & Berk, *Varieties, supra* note 4, at 339; Rossi & Berk, *Conceptual Framework, supra* note 4, at 89.

[44] Rossi & Berk, *Varieties, supra* note 4, at 339; Rossi & Berk, *Conceptual Framework, supra* note 4, at 89.

robbery and burglary vary among individuals. Some variation is random (i.e., error in judgments exists), but some variation is systematic: one individual might choose sentences of four and two years, respectively, for robbery and burglary, while a second individual might choose sentences of three and one years. The second individual has a higher punishment "threshold."

Rossi and Berk report several examples of real-world data that conform to the requirements of Model V (i.e., attitudes toward welfare entitlements and crime seriousness ratings), and they speculate that most normative domains in modern societies are best described by Model V.[45]

In Model VI—*Modified Model V, Error Variances Correlated with Individual Differences*—all conditions of Model V apply, but variation in judgments is related to some characteristic of the raters. According to this model, identifiable subgroups of the population differ on the amount of dispersion around the mean rating for the subgroup (as would be the case if, for example, there were greater consensus among women than among men about the appropriate punishment for a particular crime).[46]

In Model VII—*Modified Model V, Thresholds Correlated with Individual Characteristics*—the conditions of Model V apply, and thresholds of individuals are consistently correlated with individual characteristics. Here, subpopulations are distinguishable by the strength of their normative preferences (as would be the case, for example, if men were consistently more punitive than were women).[47] Model VII is of particular interest regarding the conflict-consensus debate. If an identifiable subpopulation were consistently more punitive than the general population, and if its greater punitiveness were consistent with the political interests of that subgroup, such a finding would support the conflict model of punishment norms.

---

[45] Rossi & Berk, *Varieties, supra* note 4, at 340; Rossi & Berk, *Conceptual Framework, supra* note 4, at 91.

[46] Rossi & Berk, *Varieties, supra* note 4, at 340-41; Rossi & Berk, *Conceptual Framework, supra* note 4, at 92-93.

[47] Rossi & Berk, *Varieties, supra* note 4, at 341; Rossi & Berk, *Conceptual Framework, supra* note 4, at 93-94.

In Model VIII—*Segmented Normative Structures: Global Dissensus and Local Consensus*—more than one set of beliefs exists about the norms of a domain. Identifiable subpopulations adhere to each set, but aggregating across the entire population obscures agreement within subgroups.[48] This circumstance would exist if, for example, wealthy people's punishment choices were guided by entirely different principles than were poor people's. We might find such a condition if wealthy people's punishment choices emphasized the financial harm suffered by crime victims, while poor people gave great weight to the employment status of offenders.

Finally, in Model IX—*Structureless Normative Domains*—normlessness exists. Within such domains choices are random.[49] Much consumer behavior (e.g., the volatile enthusiasm for fads in clothing) and public opinion on policy issues that are not salient for people are examples of such domains. With regard to punishment norms, such a condition would exist if people's punishment preferences were completely unrelated to characteristics of crimes, victims, and offenders.

B. RESEARCH STRATEGY IN RELATION TO PRIOR RESEARCH

This paper is designed to determine whether any of the Rossi-Berk models of normative structure adequately describes the normative domain of punishment for common street crimes. Data reflecting public opinions about punishment were first collected. Those data were then evaluated to determine the degree of consensus in public opinion about punishment. This evaluation was ordered by the progression in normative structuring hypothesized in the models: First the degree of overall consensus was determined. Then, where consensus was found to be relative (to qualities of the offense, offender or respondent), the factors which differentiated judgments were examined. Punishment thresholds were examined next, followed by the structure of error.

---

[48] Rossi & Berk, *Varieties, supra* note 4, at 341-43; Rossi & Berk, *Conceptual Framework, supra* note 4, at 94-97.

[49] Rossi & Berk, *Varieties, supra* note 4, at 343-44; Rossi & Berk, *Conceptual Framework, supra* note 4, at 99.

Most prior research suggests that punishment norms are structured according to Rossi and Berk's Model V (*Relative Consensus, Differentiated Judgments, Varying Thresholds, and Error*), but the evidence is inconsistent. Hamilton and Rytina, for example, had 391 respondents in the Boston SMSA match hypothetical crimes with punishments in face-to-face interviews.[50] They found, within individuals, a consistently high correlation between crime seriousness and punishment severity (e.g., relative consensus—punishment severity was related to the type of crime—and differentiated judgments, according to the seriousness of the crime).[51] On the other hand, there was great variation among individuals on the punishments they preferred for each offense (e.g., respondents had varying thresholds).[52]

On the issue of error, Hamilton and Rytina found that higher-income respondents were more likely than others to agree with average punishments (a condition of Rossi and Berk's Model VI, which posits that error variances are correlated with characteristics of individuals).[53] Hamilton and Rytina also found that lower-income and black respondents were "less likely to exhibit the high within-individual correlations between crime seriousness and punishment severity which pervaded the data set."[54] This latter finding is somewhat consistent with Rossi and Berk's Model VIII, under which more than one normative structure exists, though Hamilton and Rytina did not identify any alternative norms that lower-income and black respondents may have used in choosing punishments.

In a vignette study similar to the present one, Miller, Rossi, and Simpson found no differences in the decision rules by which men and women determine punishments.[55] Black respondents were, however, slightly more likely than whites to be

---

[50] Hamilton & Rytina, *supra* note 9, at 1124.

[51] *Id.* at 1130.

[52] *Id.* at 1132 (The authors did not report summary measures of dispersion for their entire sample, but did report significant mean dispersion between subgroups.)

[53] *Id.*

[54] *Id.* at 1140.

[55] Miller et al., *Perceptions of Justice, supra* note 27, at 331.

influenced by a complex combination of offender and offense characteristics.[56]

Sanders and Hamilton searched for segmentation of retributive justice norms among respondents in one U.S. and two Japanese cities.[57] They found little evidence that justice norms were segmented by either sex or educational attainment.[58] There was little difference between men and women, or between highly-educated and less well educated subjects, with respect to either punishment thresholds (i.e., whether a hypothetical offender should be punished) or decision norms (i.e., what punishment the offender should receive).[59]

Rauma's analysis of crime seriousness, not punishment norms, closely parallels the present study.[60] Rauma included crime seriousness rating questions in the 1986 Detroit Area Study.[61] Each of the 578 respondents rated, on a ten-point scale, the seriousness of twenty crime vignettes contained in a self-administered booklet.[62] Rauma explicitly tested the compatibility of his findings with the Rossi-Berk models.

With regard to the decision of what behaviors constitute crimes, Rauma's findings were consistent with Rossi and Berk's Model IV; that is, "widely shared norms about what constitutes a crime that are apparently unaffected by respondent characteristics."[63] With regard to the seriousness of crimes, however, Rauma reported that his findings supported a version of Rossi and Berk's Model VII; seriousness ratings were correlated with several respondent characteristics: race, gender, political affiliation, and education.[64] Respondents who were Whites, Democrats, females, and high school or college graduates gave lower mean seriousness ratings than did Blacks, Republicans, males,

---

[56] *Id.*

[57] *See generally* Sanders & Hamilton, *Common Law, supra* note 28.

[58] *Id.* at 285, 287.

[59] *Id.*

[60] *See* Rauma, *supra* note 36.

[61] *Id.* at 14.

[62] *Id.* at 14-16.

[63] *Id.* at 25.

[64] *Id.*

and respondents with less than twelve years of education.[65] These statistically significant differences between respondent groups were in the range of 0.7-1.0 on the ten-point scale, or 7-10% of the seriousness scale range. Whether differences of this magnitude represent a segmented evaluative structure, as Rauma suggests, is open to debate.

Rossi and Berk's evaluation of the Federal Sentencing Guidelines involved face-to-face interviews with a national probability sample of representatives of 1737 households.[66] Each respondent chose a preferred punishment for each of forty-two hypothetical offense vignettes that described violations of federal laws. The crime types covered by these vignettes included: drug trafficking, fraud, kidnapping, extortion, forgery, money laundering, and robbery, as well as violations of firearms, immigration, civil rights, environmental, and tax laws.

Analysis of the present study, guided by these earlier findings, attempts to resolve whether consensus on punishment norms exists and, if judgments are differentiated, the characteristics of respondents, offenses, offenders, and victims that differentiate them.

IV. METHODS

A. SAMPLE

This study is based on thirty-minute telephone interviews with a national sample of 1920 adults. The interviews were conducted between August and October of 1987. In line with Zimmerman, et al.,[67] we refer to this study as the National Punishment Survey (NPS).

The interview sample was selected from two computerized telephone lists. One list was stratified to be representative of all states, while the other list intensively sampled geographical areas with high concentrations of minority residents. About 1200 respondents came from the first list and 720 from the second.

---

[65] *Id.* at 23.

[66] ROSSI & BERK, JUST PUNISHMENTS, *supra* note 31, at 43.

[67] Sherwood E. Zimmerman et al., *The National Punishment Survey and Public Policy Consequences*, 25 J. RES. CRIME & DELINQ. 120 (1988).

The second list was required to obtain a large enough sample of minority respondents to permit testing the hypothesis suggested by Rossi and Berk's Model VII: that racial minority group members (in particular, Blacks) adhere to a different normative structure than do Whites. The overall response rate was 43%.[68]

The interview sample closely approximated the age, income, and regional distribution of the adult U.S. population; it deviated, however, on sex, race, and educational attainment. Females were overrepresented in the sample. Blacks and other non-whites were overrepresented in the sample, due to intentional oversampling of geographical areas with high concentrations of non-whites. Finally, the sample is, on the average, better educated than the U.S. population, with college educated people overrepresented and people with less than a high school diploma underrepresented. To correct for the sex, race, and education disparities, cases in the sample were weighted on these three characteristics.[69] The distribution of responses reported below should, therefore, closely approximate the attitudes of a representative cross-section of American adults.

We recognize that nonrespondents' attitudes may differ from members of the sample who completed interviews. Nonrespondents may have been more or less punitive, for example. We cannot assess this possibility directly, of course, but it is instructive that the results we report on crime seriousness approximate closely those found by Wolfgang et al.[70] in their national crime seriousness study. Further, the results of our re-

---

[68] The MACATI computer program did not permit saving partial responses, so partial response data were lost. Analysis of call records revealed that 6% of all interview attempts were partially completed—lasting more than three minutes but terminated before completion—deflating the response rate by 6%. Most nonresponses were refusals given in the first minute of interview attempts. Due to limitations of time and money, no attempts were made to convert refusals into completions. Achieving a high response rate also proved difficult for Thomas, Cage, and Foster, who reported a 46.1% response rate to their mailed questionnaire, as well as Blumstein and Cohen, whose mailed questionnaires were returned by only 24% of respondents. Thomas et al., *supra* note 2, at 112; Blumstein & Cohen, *supra* note 18, at 230.

[69] Each case was assigned a weight, the inverse of the sampling proportion for cases in the respondent's sex/race/education group. Every response was multiplied by that respondent's "weight" in analyses of aggregate responses.

[70] *See generally* WOLFGANG ET AL., *supra* note 17.

search are largely consistent with past studies.[71] Accordingly, the data do not appear to signal any clear way in which nonresponse may have affected the results reported here.

B. CRIME VIGNETTES

1. Constructing Vignettes.

To assess normative judgments, interviewers read and asked respondents to rate eight crime vignettes. Each vignette was constructed by a microcomputer program, through the factorial survey approach.[72] Thirteen "dimensions" were selected to be included in the vignettes; these dimensions were related to the type of crime, amount of harm incurred by the victim, offender characteristics, and victim characteristics. Each dimension varied in its number of "levels." For example, the dimension of offender's sex had two levels (male and female), while the dimension of offense had twenty-four levels (that is, twenty-four different crimes).

To construct a given vignette, the computer selected one level from within each of the thirteen dimensions. Each vignette is, thus, a unique, random combination of information. As such, each vignette represents a specific circumstance calling for the application of the norms concerning the proper punishment for crime. Respondents, of course, were asked to rate the vignettes by stating the sentence the offender should receive. Norms concerning proper punishment are revealed in the punishments chosen.

A detailed description of the vignette dimensions and levels is presented in the Appendix. Figure 1 displays a full, sample vignette, giving an example of each level. ·

---

[71] *See, e.g.*, Hamilton & Rytina, *supra* note 9.
[72] *See, e.g.*, ROSSI & NOCK, MEASURING SOCIAL JUDGMENTS, *supra* note 25.

Figure 1

The offender, a twenty-two-year-old male, used a knife to intentionally injure a victim. The victim, a sixty-year-old female, was treated by a doctor and was hospitalized. The offender had never had a steady job. The offender had a mental condition and was drunk when he committed the crime. He had never been convicted before for a violent offense, but had been convicted once before for stealing money or property. He had served one previous sentence of one year in jail.

Finally, we should note that the construction of the vignettes deviated slightly from complete randomization; a few specific combinations of levels were excluded because they would not typically occur in real life. For example, if the offender's age was fourteen, he or she was not "permitted" to have a criminal history involving six prior convictions for violent offenses; in forcible rape offenses, only males were permitted to be offenders and females victims. These deviations from completely random creation of vignettes introduced low intercorrelations among the dimensions.

2. *Choosing Vignette Dimensions*

The dimensions included in this study are, with few exceptions, "legally relevant variables"—characteristics that judges and parole boards may consider when evaluating a case for sentencing or parole. The primary source for these variables was the sentencing guidelines and policies established by the U.S. Sentencing Commission.[73] The Commission listed in detail many additional criteria to be considered as aggravating or mitigating circumstances that could justify harsher or milder sanctions, within or outside the guidelines.

The offender's and victim's sex are not legally relevant, according to the Sentencing Commission. Sex was included, however, based on the belief that respondents would find it eas-

---

[73] *See generally* U.S. SENTENCING COMMISSION, SENTENCING GUIDELINES AND POLICY STATEMENTS (1987).

ier to imagine a "male" or "female" rather than a "person" committing an offense. Including sex of offender may have introduced some random error, as some respondents may have had difficulty imagining some combinations of offender and offense characteristics.

A finite, manageable list of dimensions must be chosen in any study of this type. Limiting the range of factors to legally admissible ones was designed to focus respondents' attention on characteristics that may legally be manipulated in setting punishment policy. Extra-legal dimensions (e.g., race and income) likely influence criminal justice decisions and punishment preferences, but such considerations are beyond the scope of this study. Pretests of the interview schedule showed that expanding the number of dimensions (beyond the thirteen used) to include extra-legal variables would have rendered telephone interviewing unworkable.

The decisions regarding the number of dimensions to include in each vignette and the number of vignettes to pose to each respondent were guided by several overarching considerations: Telephone interviewing was selected because it was the only technique likely to produce a large, national sample of responses in a short time. Interview length was limited to thirty minutes both because longer interviews would be difficult to complete and excessive length would reduce the quality of responses due to respondent fatigue. Vignettes could have been very long, including dozens of dimensions, but one very long vignette would have consumed the entire interview. Pilot testing revealed that eight vignettes could be completed within the thirty minute limit if they contained only dimensions composed of major legally-relevant variables.

A comparison between the NPS and the studies by Rossi and Rauma points up the strengths and weaknesses of both approaches. In their self-administered questionnaire booklets, Rossi, Simpson, and Miller included fifty vignettes constructed from twenty dimensions.[74] Rauma[75] also used self-administered

---

[74] Rossi et al., *Beyond Crime Seriousness, supra* note 26, at 64-66.

[75] Rauma, *supra* note 36, at 14.

questionnaire booklets, in which respondents rated twenty vignettes composed of fourteen dimensions.

Face-to-face interviews permit respondents to rate more vignettes composed of a few more dimensions. They are extremely expensive to conduct with widely dispersed samples, however; so they typically involve geographically restricted samples. Rossi's respondents all resided in the Boston SMSA, while Rauma's respondents all lived in the Detroit metropolitan area. Respondents to the NPS, by contrast, lived all over the U.S. and, in telephone interviews, each rated eight vignettes composed of thirteen dimensions.

To permit testing for intra-individual patterns, it would have been necessary for each respondent to rate at least fifteen vignettes (two more than the number of dimensions).[76] This design limitation precludes analysis of response patterns within individual respondents. Consequently, analysis is limited to aggregate response patterns.

C. SELECTING INDEPENDENT VARIABLES

In the multivariate analyses, the vignette characteristics serve as independent variables. In addition, however, information was collected on respondent characteristics: age, sex, race, education, family income, and region (see Table 6 for the categories within each variable). Furthermore, information was collected on the offense seriousness score given to each vignette— a procedure that warrants further description.

After being read each vignette, the respondents were asked to judge the seriousness of the event. The magnitude estimation approach of Sellin and Wolfgang[77] was used to measure respondents' perceptions of offense seriousness. This procedure involved asking respondents to assign numbers representing the seriousness of offenses relative to a standard offense with a specific score. That is, after listening to a crime vignette, the respondent was asked, "What number would you give this situation [we just described] to show how serious *you* think it is

---

[76] *Id.* at 17; Peter H. Rossi & Andy B. Anderson, *The Factorial Survey Approach: An Introduction, in* ROSSI & NOCK, MEASURING SOCIAL JUDGMENTS, *supra* note 25, at 26.

[77] THORSTEN SELLIN & MARVIN E. WOLFGANG, THE MEASURE OF DELINQUENCY (1964).

compared to the bicycle theft with a score of ten?" This part of the study replicated much of the methodology of the National Survey of Crime Severity (NSCS), in which 52,000 people were surveyed by the Bureau of the Census in 1977 as a one-time supplement to the National Crime Survey.[78] Accordingly, comparison of our results with the NSCS is possible, and is presented later in this paper.[79]

D. SPECIFYING THE DEPENDENT VARIABLE: NORMATIVE PUNISHMENTS

After respondents rated the seriousness of a crime vignette, they were asked a series of questions to determine their punishment preferences. All the commonly available punishments—jail or prison, probation, fine, restitution, and (for homicide offenses only) death—were then offered. Respondents were asked which of these punishment types they would choose for the offender in that crime vignette. If they chose incarceration, they were asked whether the time should be served continuously or periodically and how long the sentence should be. If they chose a fine, they were asked the amount. Respondents could choose as many of these punishment types as they wished for each vignette.[80]

---

[78] WOLFGANG ET AL., *supra* note 17.

[79] Procedures in the National Punishment Survey differed in several important ways from the NSCS:

(1) In the NSCS respondents were interviewed mostly face-to-face; in the NPS interviews were conducted by phone. WOLFGANG ET AL., *supra* note 17, at 39.

(2) The NSCS included crime severity questions as part of a victimization survey, to which many respondents had replied one or more times before; the NPS study of crime seriousness and punishment preferences did not include questions on victimization and involved only one contact with each respondent. *Id.*

(3) In the NSCS only type of offense and amount of loss or harm were given; the NPS included information about offender and victim. *Id.* at 40.

(4) In the NSCS respondents each rated 21 crimes chosen from 204; each NPS respondent gave opinions about eight offenses chosen from 20, most of which were taken from the NSCS. *Id.*

[80] Respondents were also asked a series of questions to elicit the philosophical justifications for their punishment choices for a subsample of vignettes. Analysis of these justifications is not presented in this paper because the strength of factorial methodology is that it permits examination of norms through people's actions (the choices people make). The justifications people offer for their actions may obscure the norms actually guiding their choices.

A logical possibility exists that respondents would differ in their perception of the severity of the various sanctions. Were this the case, it would be impossible to determine whether to attribute differences in punishment preferences to differences in the desired severity of punishment, or to differences in the perception of the severity of punishments. Erickson and Gibbs have explored whether people perceive the severity of punishments differently.[81] Using a magnitude estimation procedure similar to the crime seriousness rating procedure employed in the present study, they had respondents rate the severity of several types and amounts of punishment. Erickson and Gibbs found a high degree of reliability in ratings of punishment severity.[82] They also found that police respondents consistently rated punishments as being more severe than did other citizens.[83] These findings by Erickson and Gibbs provide some reassurance that differences in sentencing preferences among respondents in the present study would be produced largely by differences in the desired severity of punishment.

V. RESULTS

The analysis presented below is designed to determine whether any of the Rossi-Berk models of normative structure adequately describes the normative domain of punishment for common street crimes. The analysis is therefore organized to search for consensus on punishment type and amount in the progressive order hypothesized in the models.

The first set of analyses is guided by Rossi and Berk's Model V, which hypothesizes, in part, relative consensus and differentiated judgments. The analysis therefore covers the degree of differentiation (or variability) in punishment *type* (i.e., imprisonment, probation or fine) and *severity* (i.e., length of prison sentence) based on offense characteristics. Next, the sources of punishment differentiation are examined. These sources include offense type, degree of harm, seriousness (as indicated on

---

[81] Maynard L. Erickson & Jack P. Gibbs, *On the Severity of Legal Penalties*, 70 J. CRIM. L. & CRIMINOLOGY 102, 102-16 (1979).

[82] *Id.* at 108-09.

[83] *Id.* at 116.

the Sellin-Wolfang scale), and dollar loss (for property offenses).

The analysis then shifts its focus from offenses to respondents, to determine whether punishment preferences are structured by characteristics of respondents. Here the most important question addressed is whether there are identifiable subgroups of the population that hold distinctive punishment norms, as suggested by Rossi and Berk's Model VII. Within this analysis the relative importance of offense and respondent characteristics are compared.

The final series of analyses examines the structure of dispersion or error, to determine whether there are subgroups of respondents who share greater consensus on punishment than does the general population, as hypothesized by Rossi and Berk's Model VI, and as Rauma found in evaluations of crime seriousness.[84]

A. PUNISHMENT VARIABILITY

Selection of type of punishment shows strong normative features, with incarceration being chosen overwhelmingly by respondents. Across all twenty-four offense types and all conditions, the most preferred punishment was a jail or prison sentence, chosen for 71% of vignettes. (Variation by offense type is described in the next section). Some respondents combined other types of punishment with imprisonment: Probation was added to imprisonment in 30% of cases, a fine in 24%, and restitution in 35%. It is clear, however, that these alternatives were seldom preferred as substitutes for imprisonment. As Table 1 shows, probation was selected as the most severe penalty in

---

[84] Rauma, *supra* note 36, at 25.

Table 1.
Punishment Preferences Across all Offenses

Punishment Type	Responses That Included This Punishment[a]		Responses Where This Punishment Was the Most Severe Punishment Selected[b]	
	Percent	(n)	Percent	(n)
Death [c]	34.2%	1,872	34.2%	1,872
Jail or Prison	71.4	14,174	71.4	14,174
Probation	29.8	14,174	16.6	14,174
Fine	24.3	14,174	3.8	14,174
Restitution	35.2	14,174	3.7	14,174

[a] After they rated the seriousness of the offenses, respondents were read the four commonly available punishment types in this order: Jail or prison, probation, fine, restitution, and (for homicide offenses only) the death penalty. They were then asked which of these punishments the offender (if arrested and convicted) should receive, and told they could choose as many punishment types as they wished. Where a response included the death penalty, all other punishments were deleted from the analysis of that response. This column does not add to 100 percent because many responses included more than one punishment type for each offense.

[b] Punishments were ranked in the following order, from most to least severe: death penalty, jail or prison, probation, fine, and restitution. Only the most severe punishment of all those chosen for an offense is reported in this column.

[c] In these interviews, respondents could choose the death penalty for only three (homicide) offenses; therefore the percentages presented regarding the death penalty are for responses about these offenses only. Among all 1,872 responses, 41.8% were "No" and 24.0% were "Don't know."

only 17% of all cases, a fine in 4% and restitution in 4%.[85]

The death penalty was an option for only three of the twenty-four offense types—homicides associated with assault, robbery, and forcible rape offenses. Capital punishment was chosen for 34% of the vignettes depicting homicide offenses. This figure is low in comparison with current levels of support for the death penalty as measured by general questions. Some 70% of respondents to a national poll in 1987 (the same year as the NPS) said they "favor[ed] the death penalty for persons convicted of murder."[86] The homicide offenses included in this study—committed in the context of a rape, robbery and assault— may not be typical of all "murders." Rape and robbery homi- cides do constitute felony murders, however, punishable by death in many states. The observed lower level of support for the death penalty for specific offenses is consistent with the proposition that respondents who are given more detail about a crime form less punitive judgments.

---

[85]Caution should be exercised in inferring from these aggregate data. One can only say that for the mix of offenses represented by the 15,360 crime vignettes posed in this study, in 71% of responses imprisonment was the most severe sanction pre- ferred. This aggregate percentage reflects responses to the specific mix of criminal offenses examined in this study. The proportional distribution of types of offenses among vignettes does not resemble the actual distribution of crimes resulting in con- viction in U.S. courts.

Comparison of the distribution of offenses in the NPS with the distribution of fel- ony conviction offenses in U.S. state courts revealed an overrepresentation of the most serious offenses in the NPS. *See* BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, FELONY SENTENCES IN STATE COURTS 1988, at 2 (1990). The authors will fur- nish this comparison upon request. Zimmerman et al. compared the proportional distribution of offense types offered in vignettes in the NPS with the actual distribu- tion of offenses resulting in conviction in New York State. They found the less serious types of offenses—larceny, harassment and DWI—were underrepresented in the NPS, while some very serious offenses—murder/manslaughter and DWI resulting in a death—were overrepresented. Zimmerman et al., *supra* note 67, at 120.

Though the mix of offense types included in the NPS biases the overall set of re- sponses toward severe sanctions (i.e., long prison terms), there is ample justification for this mix. The 24 offense types included represent common street crimes, about which the public is concerned, and which constitute a substantial proportion of of- fenses actually processed by the criminal justice system. The included offenses cover a wide range in seriousness—from larceny of $10 to rape-murder—and a substantial number of behavioral elements crucial for sentencing—assault, threats, unlawful en- try, weapon use, theft, drug use, sexual content—to providing the opportunity to ana- lyze the structure of punishment preferences across this wide range of concerns.

[86] SOURCEBOOK, *supra* note 12, at 185.

B. PUNISHMENT TYPE DIFFERENTIATION

1. Differentiation by Offense Type

There was considerable variability in preferred sanction type according to type of offense (see Table 2). A majority of respondents favored imprisonment for all offenses, with the exception of larceny of property worth $10. Imprisonment is more strongly favored for violent sex offenses than for any other category of offenses; forcible rape offenses elicited imprisonment as the preferred punishment from more than 94% of respondents. Probation was most preferred as an add-on for cocaine use and the $10 burglary.

Repeating the pattern over all offenses, no alternative to imprisonment was preferred as the most severe penalty for any offense (see Table 3). The most popular application of probation as the most severe sanction was for a $10 larceny (35%), $10 burglary (33%), and cocaine use (35%). Even in these cases imprisonment was far more commonly chosen as the most severe sanction. Fines and restitution did not exceed 20% (reaching this peak for the $10 larceny) of most severe punishments for any offense.

2. Differentiation by Degree of Harm

Within offense categories, imprisonment was uniformly more strongly favored for more harmful offenses. For example, 78% favored a prison term for larceny of property worth $10,000, compared to 55% favoring a prison term for larceny of property worth $50. This pattern is consistent across all offense types.

The death penalty, available as an option only for the three homicide offenses, was most preferred (42%) for forcible rapes resulting in death, compared to robberies resulting in death (37%) and fatal assaults (30%).[87]

---

[87] The same caution applies to the interpretation of these results as to the aggregate data: Respondents gave their opinions of appropriate punishments in relation to specific offense descriptions. The distribution of offense characteristics in the vignettes may not resemble the distribution of characteristics of all offenses of a particular type (e.g., all homicides) resulting in conviction in the U.S.

C. SENTENCE LENGTH DIFFERENTATION

1. Differentiation by Offense Type

Incarceration is clearly the preferred punishment for felony crimes, but there is less consensus over appropriate prison sentence lengths (see Table 4). Clearly, respondents differentiated between offense types in assigning sentence lengths. The shortest mean preferred sentence for any offense—burglary of a building netting $10—was twenty-seven months in prison, with a median of twelve months. Drunk driving without an accident received a mean sentence of more than twenty-seven months, with a median of twelve months.

The longest sentences were for violent assaults resulting in death. When sentences of "life" and "death" were included (recoded as forty-year sentences), mean sentences for the three fatal assaults were between thirty and thirty-five years, with a median at the forty-year maximum for all three offenses.

2. Differentiation by Degree of Harm

The five larceny crimes differ only in the dollar value of the amount stolen—$10, $50, $100, $1,000 and $10,000. These five crimes were compared to ascertain the effect of varying pecuniary harm to victims. Cumulative response distributions of sentence length preferences for larceny crimes are shown in Figure 2. The vertical axis represents the percentage of respondents choosing lengths at least as long as the sentence lengths represented on the horizontal axis. Distributions in Figure 2 were truncated at 180 months to permit examination of detailed differences between curves.

Figure 2 shows a set of similarly shaped curves. The curves representing the higher dollar value thefts are flatter and have higher means (i.e., respondents chose longer sentences for thefts with larger dollar losses).

Table 2.

All Punishments Selected, by Offense

Offense Type	Punishment Type Selected[a] (Percent)					
	Death	Jail or Prison	Probation	Fine	Restitution	n
Property Theft & Damage						
Arson-$500,000 Damage	–	81.5%	27.1%	24.3%	39.6%	536
Larceny of $10,000	–	78.4	28.2	22.3	47.4	733
Car Theft-Sale-$5,000	–	72.9	36.1	26.3	59.8	603
Larceny of $1,000	–	67.7	34.4	17.9	43.9	727
Larceny of $100	–	62.3	33.5	22.4	46.1	751
Larceny of $50	–	55.3	38.8	24.0	49.6	768
Larceny of $10	–	45.6	41.9	24.0	48.5	684
Burglary Offenses						
Burglary-Home-$1,000	–	80.7	31.4	23.8	59.6	574
Burglary-Building-$10	–	56.5	46.8	30.7	47.7	530
Robbery Offenses						
Robbery-Gun-Death	37.1	61.7	10.6	6.8	16.8	570
Robbery-Gun-Hospital-$1,000	–	92.1	22.5	22.5	47.6	552
Robbery-Weapon-No Harm-$10	–	74.5	33.4	26.5	34.7	486
Robbery-Threat-No Harm-$10	–	72.2	32.9	31.4	45.2	605

Offense Type	Death	Jail or Prison	Probation	Fine	Restitution	n
Assault Offenses						
Assault-Death	29.7	67.4	11.6	7.6	12.4	557
Assault-Hospital	–	82.3	29.1	19.9	42.4	560
Assault-Doctor	–	78.3	34.2	28.2	43.9	543
Assault-No Injury	–	55.4	39.5	34.3	16.7	484
Forcible Rape Offenses						
Rape-Death	41.7	57.0	5.0	5.3	11.6	633
Rape-Oral Sex-No Other Injury	–	94.7	18.8	19.6	27.0	583
Rape-No Other Injury	–	94.1	21.9	19.2	24.1	553
Drunk Driving Offenses						
Drunk Driving-Death	–	90.6	21.2	29.5	33.6	555
Drunk Driving-No Accident	–	54.1	40.2	57.8	8.0	541
Drug Offenses						
Cocaine-Sold for Resale	–	89.9	27.1	35.4	7.8	565
Cocaine-Used	–	57.9	49.0	33.9	7.3	481
Means	36.4%[b]	71.4%	29.8%	24.3%	35.2%	

[a] Respondents were read the four commonly available punishment types in this order: jail or prison, probation, fine, restitution, and (for homicide offenses only) the death penalty. They were then asked which of these punishments the offender (if arrested and convicted) should receive, and told they could choose as many punishment types as they wished. Where a response included the death penalty, all other punishments were deleted from the analysis of that response. The rows do not add to 100 percent because many responses included more than one punishment type for each offense.

[b] The percentage of respondents who selected the death penalty was averaged over only the three (homicide) offenses for which the death penalty was an optional punishment.

Table 3.
Most Severe Punishment Selected, by Offense

Offense Type	Death	Most Severe Punishment Type Selected[a] (Percent)				
		Jail or Prison	Probation	Fine or Restitution	Totals[b]	n
Property Theft & Damage						
Arson-$500,000 Damage	–	81.5%	11.1%	7.4%	100.0%	536
Larceny of $10,000	–	78.4	12.8	8.8	100.0	733
Car Theft-Sale-$5,000	–	72.9	19.3	7.8	100.0	603
Larceny of $1,000	–	67.7	23.0	9.3	100.0	727
Larceny of $100	–	62.3	23.0	14.7	100.0	751
Larceny of $50	–	55.3	29.0	15.7	100.0	768
Larceny of $10	–	45.6	34.6	19.9	100.1	684
Burglary Offenses						
Burglary-Home-$1,000	–	80.7	14.4	4.9	100.0	574
Burglary-Building-$10	–	56.5	32.6	11.0	100.1	530
Robbery Offenses						
Robbery-Gun-Death	37.1%	61.7	1.2	0.1	100.1	570
Robbery-Gun-Hospital-$1,000	–	92.1	5.6	2.2	99.9	552
Robbery-Weapon-No Harm-$10	–	74.5	19.5	6.1	100.1	486
Robbery-Threat-No Harm-$10	–	72.2	19.2	8.7	100.1	605

Offense Type	Death	Jail or Prison	Probation	Fine or Restitution	Totals[b]	n
Assault Offenses						
Assault-Death	29.7	67.4	2.3	0.6	100.0	557
Assault-Hospital	–	82.3	14.4	3.2	99.9	560
Assault-Doctor	–	78.3	16.3	5.3	99.9	543
Assault-No Injury	–	55.4	28.1	16.6	100.1	484
Forcible Rape Offenses						
Rape-Death	41.7	57.0	0.9	0.4	100.0	633
Rape-Oral Sex-No Other Injury	–	94.7	4.3	1.0	100.0	583
Rape-No Other Injury	–	94.1	4.7	1.1	99.9	553
Drunk Driving Offenses						
Drunk Driving-Death	–	90.6	6.7	2.7	100.0	555
Drunk Driving-No Accident	–	54.1	29.4	16.5	100.0	541
Drug Offenses						
Cocaine-Sold for Resale	–	89.9	7.6	2.5	100.0	565
Cocaine-Used	–	57.9	35.3	6.8	100.0	481
Total						14,174[c]

[a] The entries in this table represent the most severe penalty chosen among all the penalties given by each respondent for each offense type.

[b] Some rows do not total 100% due to rounding.

[c] A total of 15,360 responses were obtained; the remaining 7.7% were recorded as "Don't know" or "No" to all punishment types.

Table 4. Jail or Prison Sentence Length, by Offense

| | Sentence Length[a] (Months) | | | |
Offense Type	Mean	Median	Standard Deviation	n
Property Theft & Damage				
Arson-$500,000 Damage	99.9	60.0	76.7	420
Larceny of $10,000	67.8	36.0	84.5	532
Car Theft-Sale-$5,000	55.5	36.0	76.7	420
Larceny of $1,000	54.8	24.0	89.8	445
Larceny of $100	43.7	12.0	74.5	408
Larceny of $50	37.4	12.0	59.0	379
Larceny of $10	32.9	12.0	64.3	282
Burglary Offenses				
Burglary-Home-$1,000	53.4	24.0	72.4	442
Burglary-Building-$10	27.0	12.0	43.7	270
Robbery Offenses				
Robbery-Gun-Death	365.2[b]	480.0	161.5	548
Robbery-Gun-Hospital-$1,000	123.4	60.0	129.3	482
Robbery-Weapon-No Harm-$10	68.0	36.0	91.0	339
Robbery-Threat-No Harm-$10	46.1	24.0	75.1	406
Assault Offenses				
Assault-Death	349.5[b]	480.0	174.5	536
Assault-Hospital	92.7	60.0	109.7	446
Assault-Doctor	67.3	36.0	100.2	403
Assault-No Injury	42.8	24.0	70.3	239
Forcible Rape Offenses				
Rape-Death	416.4[b]	480.0	132.9	616
Rape-Oral Sex-No Other Injury	202.1	120.0	173.3	529
Rape-No Other Injury	184.9	120.0	155.3	489
Drunk Driving Offenses				
Drunk Driving-Death	141.2	84.0	152.5	486
Drunk Driving-No Accident	27.4	12.0	53.8	258
Drug Offenses				
Cocaine-Sold for Resale	126.3	60.0	142.9	498
Cocaine-Used	66.5	24.0	104.4	262

[a] Only responses where a jail or prison sentence was selected, and the respondent chose a specific sentence length, are included here. All sentence lengths over 40 years and all sentences of "life" were recoded to 40 years, which was considered to be, effectively, a life sentence.

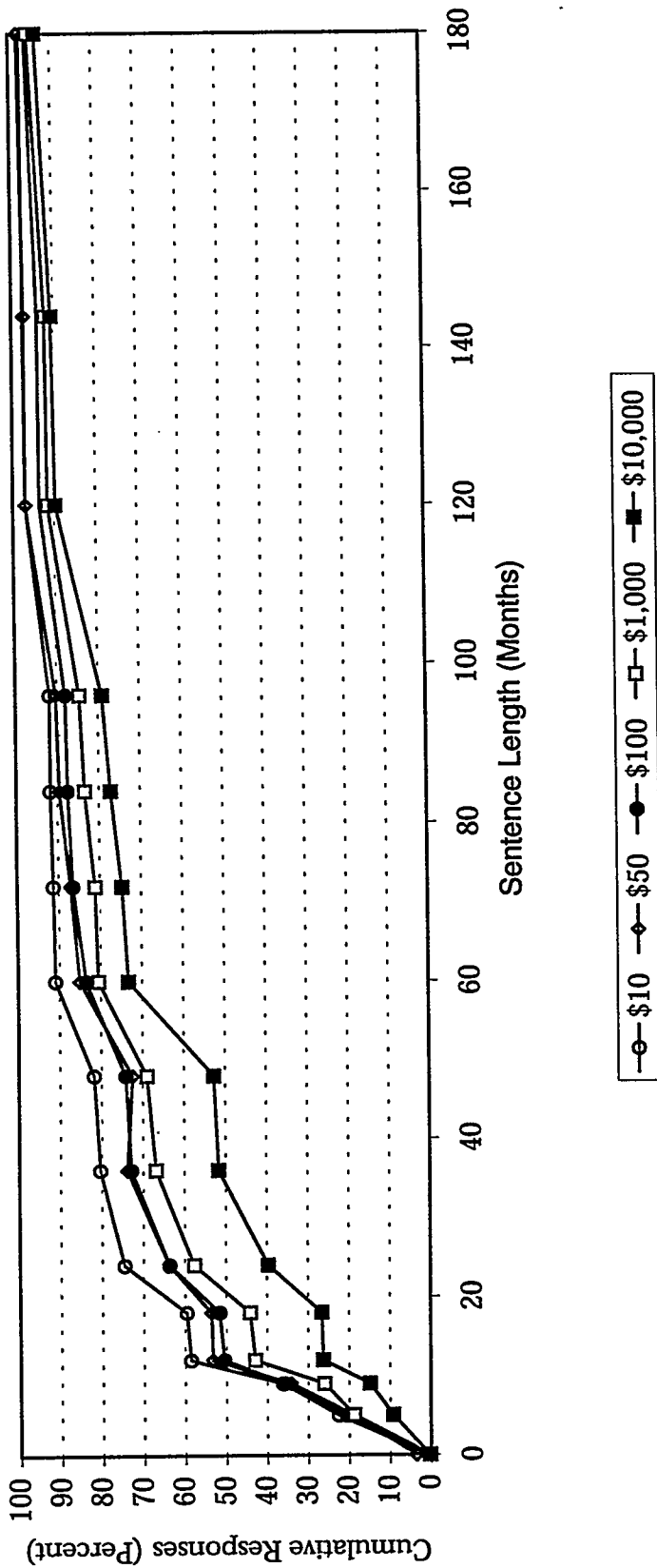[b] Sentences of "death," available only for the homicide offenses, were recoded to 40 years for this analysis.

Figure 2.

Sentence Lengths for Five Larceny Offenses–Cumulative Responses

3. Differentiation by Degree of Seriousness

The offense seriousness magnitude estimation tasks included in the National Survey of Crime Severity produced a classical power function, as shown repeatedly for a wide variety of physical stimuli.[88] In the NSCS the power function of offense seriousness ratings for five larcenies had the form:

$Y = aX^b;$ $Y = 21.88 X^{0.27}$

Where: Y = magnitude of perceived stimulus (seriousness),

a = Y intercept,

X = magnitude of physical stimulus (dollar loss),

b = slope of the function.

A power function also described the relationship between dollar loss and perceived seriousness of these crimes in the NPS. This function has the form: $Y = 21.5 X^{0.19}$

Data from both the NSCS and the NPS described above are plotted on a log/log scale in Figure 3. The two functions have nearly identical Y-intercepts, though the NPS data have a lower slope.[89] The crime seriousness ratings of respondents in the NPS were less sensitive to increases in dollar value of thefts than were the ratings of respondents to the earlier NSCS.

This is the first national survey to combine the Sellin-Wolfgang offense seriousness rating scheme with a measure of preferred punishment. Respondents gave both an offense seriousness score and, if they chose a prison term, a preferred length of confinement. In Table 5 means for these two measures are presented together by offense type. The arithmetic mean is given for sentence length, while the geometric mean is given for the

---

[88] Stanley S. Stevens, *On the Psychological Law*, 64 PSYCHOL. REV. 153, 162 (1957); Stanley S. Stevens & E. Galenter, *Ratio Scales and Category Scales for a Dozen Perceptual Continua*, 54 J. EXPERIMENTAL PSYCH. 377, 409 (1957).

[89] The NSCS and NPS were conducted ten years apart, during which inflation produced a 53% devaluation in the value of U.S. dollars, as measured by change in the Consumer Price Index. The NSCS data were therefore converted to 1987 dollars and the results compared to the unadjusted figures. This correction did not change the slope of the function, but only moved the y-intercept down slightly.
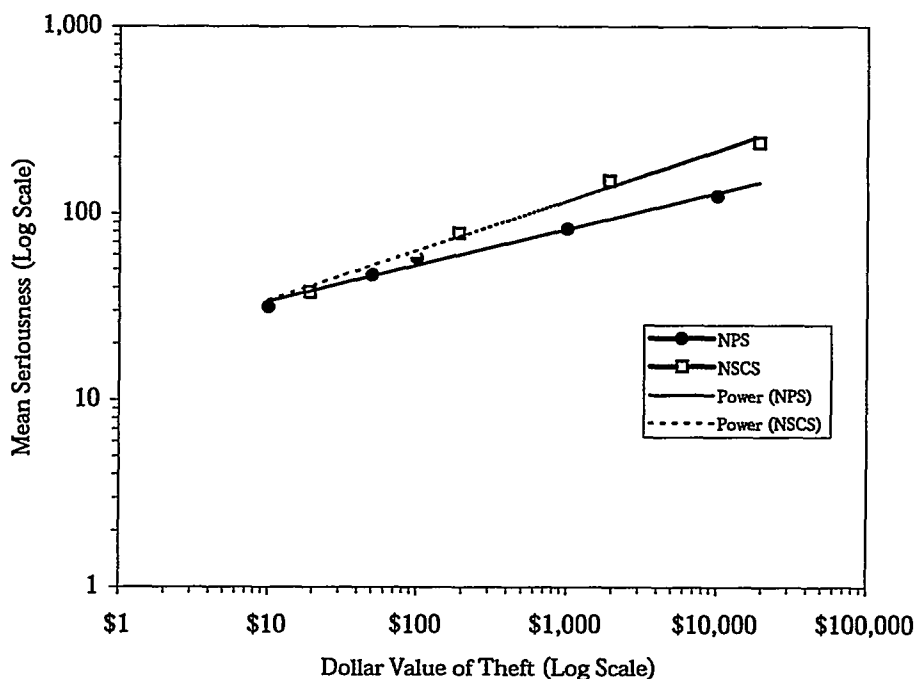
Figure 3.
Seriousness by Dollar Loss : National Punishment Survey
and National Survey of Crime Severity

seriousness score.[90]

Offenses were ranked identically on mean sentence length and seriousness score through the first four offenses. Some variability appears in the ordering below that, though offenses with higher average sentence lengths were generally viewed as more serious.

---

[90] The geometric mean is defined as the positive nth root of the product of the numbers, or the antilog of the mean of the sum of the logs. William G. Hines, *Geometric Mean, in* 3 ENCYCLOPEDIA OF THE STATISTICS SCIENCES 397, 397 (Samuel Kotz & Norman L. Johnson eds., 1983). The geometric mean is the appropriate measure of central tendency for ratio scale scores. It reduces the effect on the mean of outliers in very widely dispersed distributions. In this study seriousness ratings were very widely dispersed—they ranged from 0.3 to 100 billion—because they were presented to respondents as having no upper or lower limits.

Table 5. Sentence Length by Offense Seriousness

Offense Type	Sentence Length[a]			Offense Seriousness		
	Mean (Months)	Rank	(n)	Geometric Mean	Rank	(n)
Rape-Death	416.4[b]	1	616	738.8	1	620
Robbery-Gun-Death	365.2[b]	2	548	629.9	2	600
Assault-Death	349.5[b]	3	536	441.7	3	572
Rape-Oral Sex-No Other Injury	202.1	4	529	414.0	4	602
Rape-No Other Injury	184.9	5	489	390.7	6	585
Drunk Driving-Death	141.2	6	486	400.8	5	594
Cocaine-Sold for Resale	126.3	7	498	217.9	9	575
Robbery-Gun-Hospital-$1,000	123.4	8	482	266.9	7	567
Arson-$500,000 Damage	99.9	9	420	220.7	8	544
Assault-Hospital	92.7	10	446	197.8	10	591
Robbery-Weapon-No Harm-$10	68.0	11	339	178.4	11	550
Larceny of $10,000	67.8	12	532	124.4	14	751
Assault-Doctor	67.3	13	403	140.0	12	593
Cocaine-Used	66.5	14	262	89.1	18	556
Car Theft-Sale-$5,000	55.5	15	420	123.2	15	618
Larceny of $1,000	54.8	16	445	83.0	19	759
Burglary-Home-$1,000	53.4	17	442	133.5	13	620
Robbery-Threat-No Harm-$10	46.1	18	406	91.3	17	645
Larceny of $100	43.7	19	408	57.2	21	807
Assault-No Injury	42.8	20	239	36.6	23	582
Larceny of $50	37.4	21	379	46.7	22	826
Larceny of $10	32.9	22	282	31.5	24	791
Drunk Driving-No Accident	27.4	23	258	95.9	16	579
Burglary-Building-$10	27.0	24	270	60.6	20	546
Totals			10,135			15,073[c]

[a] Only those responses where a jail or prison sentence was selected, and the respondent chose a specific sentence length, are included here. All sentence lengths over 40 years and all sentences of "life" were recoded to 40 years, which was considered to be, effectively, a life sentence.

[b] Sentences of "death," available only for the homicide offenses, were recoded to 40 years for this analysis.

[c] Respondents failed to rate the seriousness of 1.9% of vignettes.

Results of Spearman Rank Order Correlation (with sentence length dependent):

Intercept	-4.66
Slope	0.557
r	0.956
r^2	0.915

The relationship between mean sentence length and offense seriousness is displayed in Figure 4. The correlation between sentence length and seriousness (r = .956) underscores the close correspondence between the two variables. Across twenty-four offense types, 91.5% of the variation in average sentence length is explained by variation in mean offense seriousness.

The close relationship between seriousness and sentence lengths is attenuated at the individual response level. The correlation between paired seriousness ratings and imprisonment lengths across all 9997 individual vignettes for which both ratings were recorded is much lower (r = .336). This difference in magnitude between the two correlations reveals a strong, aggregate, linear relationship between offense seriousness and sentence length; but marked deviations from this pattern exist in individual pairs of ratings.

The difference between individual vignette correlations and correlations between aggregated mean ratings indicates the presence of both considerable error and differing individual thresholds. Threshold differences represent variations from individual to individual in their ratings on both scales. Respondents agreed, on average, on the ordering of the twenty-four crimes in seriousness and deserved sentence length. They did not agree on the appropriate value for seriousness or sentence length.[91]

4. Differentiation by Dollar Loss

This relationship is analogous to the relationship between dollar loss and offense seriousness. Dollar loss can be consid-

---

[91] The error component can be viewed as a function of the rating tasks: Seriousness is measured by a ratio scale; scores have no common unit of measure (e.g., pound or inch), they have no upper limit, and they are not additive. Each respondent applied his/her own set of numbers to the concept "seriousness"; so raw scores varied by many orders of magnitude (as described above), according to whims of respondents. Interpretation of a score is clearest when this score is compared directly with other scores produced similarly by the same respondent. In contrast to offense seriousness, sentence length is an interval scale; individual scores are additive and have a common unit of measure ("a month in jail or prison") which has a commonly understood meaning.

ered an objective stimulus and preferred sentence length interpreted as an indicator of strength of the perceived stimulus.
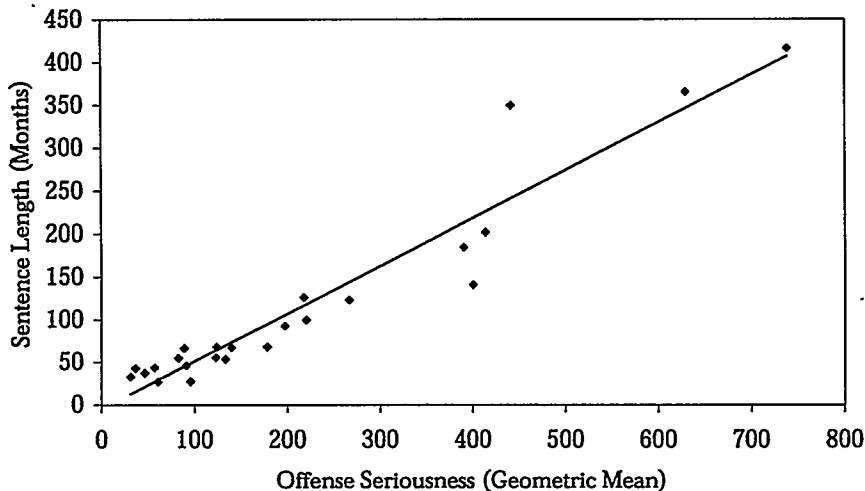

Figure 4
Sentence Length by Offense Seriousness

Figure 5 shows the linear relationship between logs of dollar loss and sentence length. Considering only means of these values (not variation among individuals), r = .9942 (r^2 = .989). Sentence length is a power function of dollar loss. The least squares regression line for this function has the form:

$$Y = 25.69 \, X^{0.107}$$

Where: Y = mean preferred sentence length in months;
X = dollar loss stated in the vignette.

D. CONSENSUS OR DISSENSUS ON SENTENCE LENGTH

Within each category of offenses—theft, burglary, robbery, etc. —there is a consistent pattern of more harmful offenses receiving longer average sentences. Despite this pattern, dispersion of sentence preferences among respondents is high. Medians of sentence lengths are only 27-65% of means for all crimes except capital offenses. Standard deviations are large.
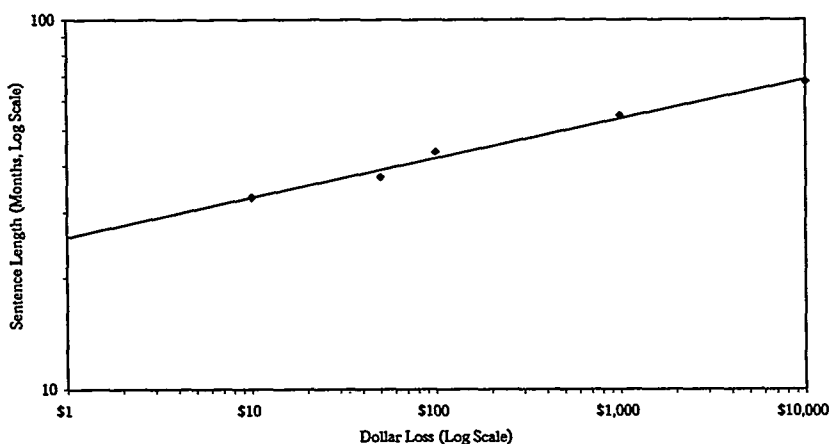
Figure 5
Sentence Length by Dollar Loss for Larceny Offenses

Point estimates (means or medians) of public preference are, therefore, misleading. Point estimates, when used alone, inaccurately suggest consensus.

Returning to Rossi and Berk's analytical model, characteristics of punishment norms found thus far place this domain within Rossi and Berk's Model V (relative consensus, differentiated judgments, varying thresholds, and error): People distinguish among types of offenses in choosing kinds and amounts of punishment. They agree on the kinds of punishment appropriate for different offenses. Punishment severity is consistently related to harm; but people do not agree on the amount of punishment to be applied for each offense (i.e., people have different punishment thresholds). Whether punishment thresholds are patterned by characteristics of respondents (i.e., whether Rossi and Berk's Model VI describes the structure of punishment norms) will be explored below.

E. PUNISHMENT THRESHOLDS BY RESPONDENT CHARACTERISTICS

Because of limited interview time, only a few demographic characteristics were obtained. Differences among respondents along demographic lines were neither large nor systematic, as Table 6 shows.

The second column of Table 6 considers only vignettes depicting capital crimes. Preferences for the death penalty for homicides varied significantly by age, sex, education, and family income, though not in any clear pattern. Males were significantly more likely to choose the death penalty than were females. Differences by race and region were not statistically significant.

The third column of Table 6 considers all crimes, showing the percentage of respondents preferring incarceration as a punishment. The fourth column displays mean incarceration lengths. Preference levels for imprisonment were significantly different by age, education, family income, and region. Two clear patterns emerged: Respondents with the least and most education, and Westerners, favored imprisonment least. Although a statistically significant difference appeared in sentence length by age, this difference is not clearly patterned. Sentence length differences by sex, race, education, family income, and region were not significant.

F. RESPONSE SENSITIVITY

In the analyses presented up to this point, there is abundant evidence that the data are both structured and heterogeneous.

At the aggregate level (i.e., average responses), there is considerable structure, but among individual responses there is considerable variability. How these seemingly uneven characteristics come about can be seen in the analysis presented in this section. Here the vignette becomes the unit of analysis.

With each rated vignette as a unit, multiple regression analyses were conducted on sentence lengths. The regressors were the levels included in the vignettes, each level coded as a dummy variable, and the demographic characteristics of respondents, also represented as dummies. These analyses, structured hierarchically (see Table 7), show respondents' sensitivity

to the various dimensions in choosing sentences, as well as the importance of demographic characteristics.

Each line in Table 7 refers to a separate regression equation. The first equation contains only the offense types as dummy variable regressors. Each successive equation includes all the independent variables from the preceding equations plus an additional set of dummies, as indicated. The purpose of the table is to show how much additional variance is explained by adding successive sets of information to the model.

In the creation of the crime vignettes, assignment of some characteristics of offenses and offenders was contingent upon the prior selection of other characteristics (e.g., information about the age of the victim was given only for crimes involving personal injury to a victim). Hence n's for later equations are reduced under the listwise deletion rule. Because orthogonality exists among vignette characteristics, collinearity among vignette dimensions and levels does not confound judgments about relative sensitivity.[92]

The most noteworthy feature of Table 7 is that respondents were most sensitive to offense type. Offense explains 51% of the variance in sentence length, as shown in Equation 1. Adding other regressors in later equations adds little explanatory power—the highest r^2 in the table is 0.60.

The finding that prior criminal record (whether measured by number of convictions or incarcerations, or length of prior incarcerations) had little effect on preferred sentence length is highly relevant to contemporary changes in sentencing statutes that provide dramatically longer sentences for defendants with previous convictions. In apparent contrast to these findings is the work by Finkel and his colleagues, who studied the responses of college undergraduates to the application of "three

---

[92] A few pairs of dimensions had collinearity imposed in the design by restrictions on combinations. Offender's age, for example, had correlations in the .20 range with number of prior convictions for assaultive offenses and larcenies, and number of prior incarcerations. Correlations between pairs of dimensions whose combinations were not restricted were all below .06.

Table 6.
Punishment by Respondent Characteristics

Characteristic	Selected Death for Homicides		Selected Jail or Prison[a]		Mean Jail or Prison Sentence	
	Percent	(n)	Percent	(n)	Months	(n)
Age						
18-24	25.97%	31	74.98%	234	134.4	234
25-34	31.57	71	74.55	384	134.8	380
35-44	28.36	40	70.99	346	123.8	318
45-54	35.39	40	76.39	288	129.7	285
55-64	26.74	46	76.05	285	152.5	284
65-74	41.27	67	78.61	281	135.5	281
75+	29.75	14	77.48	96	142.0	96
Total/Mean	*32.37*	*310*	*75.20*	*1,917*	*135.2*	*1,881*
Sex						
Male	35.14	147	75.09	907	134.1	879
Female	29.90	163	75.28	1,011	136.2	1,002
Total/Mean	*32.39*	*311*	*75.19*	*1,919*	*135.2*	*1,882*
Race/Ethnicity						
White	33.30	251	75.55	1,617	135.3	1,584
Black	27.15	53	72.43	250	136.9	247
Hispanic	49.10	4	77.46	22	125.0	22
Asian	–	–	65.83	4	102.2	4
Other	30.22	5	80.67	24	126.4	25
Total/Mean	*32.39*	*311*	*75.21*	*1,920*	*135.2*	*1,883*

Characteristic	Selected Death for Homicides		Selected Jail or Prison[a]		Mean Jail or Prison Sentence	
	Percent	(n)	Percent	(n)	Months	(n)
Education						
Elementary	42.39	41	64.85	131	149.6	110
Junior High	25.83	21	79.58	160	134.8	159
Some High School	33.42	34	77.10	238	142.3	234
High School Grad	30.65	108	77.75	728	136.2	727
Some College	32.12	49	74.51	308	138.5	305
College Grad	30.01	35	73.66	231	122.1 ·	227
Post-Graduate	30.97	20	66.36	121	118.1	118
Total/Mean	*32.39*	*311*	*75.21*	*1,920*	*135.2*	*1,883*
Family Income						
<$6,000	21.65	24	77.74	137	146.8	137
$6,000-12,999	42.58	58	76.48	216	134.1	216
$13,000-18,999	37.24	28	75.30	234	126.4	231
$19,000-28,999	34.94	62	71.14	394	136.2	372
$29,000-47,999	29.06	58	76.35	364	133.6	363
$48,000-74,999	30.54	25	75.69	168	137.5	163
$75,000+	21.15	6	75.71	60	132.6	59
Total/ Mean	*33.48*	*263*	*74.93*	*1,575*	*134.8*	*1,543*
Region						
Northeast	35.06	51	74.59	347	135.9	345
Midwest	27.74	82	75.84	545	130.6	544
South	34.60	146	76.77	707	136.8	703
West	30.00	31	71.32	318	139.0	290
Total/Mean	*32.39*	*311*	*75.21*	*1,920*	*135.2*	*1,883*

[a] The entries in this column are the percentage of each respondent category who included jail or prison among the types of punishment they selected.

Table 7.
Results of Hierarchical Regression on Sentence Length

Equation Number	Variables Added	R	R^2	R^2 Added	(n)
1	Offense type	.7116	.5064	–	10,131[a]
2	Respondent Characteristics	.7229	.5225	.0161	8,341[b]
3	Prior convictions	.7293	.5326	.0101	8,241
4	Offender's age	.7329	.5371	.0045	8,241
5	Offense seriousness score	.7395	.5468	.0097	8,097[c]
6	Offender's employment, mental illness, drug/alcohol use	.7406	.5484	.0016	8,097
7	Number of prior incarcerations	.7413	.5496	.0012	8,097
8	Length of prior incarcerations	.7420	.5505	.0009	8,097
9	Victim's age	.7496	.5619	.0114	3,579[d]
10	Offender's and victim's sex	.7732	.5978	.0359	2,357[e]
11	Weapon used	.7665	.5875	-.0103	946[f]

[a] Values for the dependent variable (sentence length) were not recorded for vignettes where the respondent failed to select "jail or prison" or, where the respondent selected jail or prison but failed to choose a specific period of confinement.

[b] Cases were deleted "listwise" in this analysis. Responses where the respondent failed to give all respondent characteristics were excluded.

[c] Cases were deleted for this and subsequent equations if the respondent failed to give a seriousness score.

[d] This equation includes only vignettes containing offenses involving personal victims (i.e., assaults, robberies, and forcible rapes). There were the only offenses where the victim's age was stated in the vignette.

[e] This equation includes only vignettes containing non-sexual assaults and robberies. These were the only offenses for which the victim's sex was given and varied. The victim's sex was given as female for all sexual assaults.

[f] This equation includes only vignettes containing non-sexual assaults, the only offenses for which the type of weapon was identified.

strikes" laws to property offenders.[93]   They found that respondents did favor longer sentences for recidivists, but not sentences as long as many statutes provide: "[Survey subjects] do not limit their punishment to just the last offense when they know there were six priors, but neither do they dramatically, geometrically, or exponentially escalate their punishments because of priors."[94]   After reviewing the research literature on public attitudes toward imposing more severe sentences on recidivists, Roberts concluded:

> Public support is probably restricted to recidivist statutes that target offenders convicted of repetitive violent conduct. Recidivist statutes that result in long-term or indeterminate detention for relatively minor property crimes (and that accordingly violate the desert-based principle of proportionality in sentencing) are in all likelihood contrary to community sentiment.[95]

Roberts, however, noted that the existing research is flawed. When only offense type and criminal history information are provided to experimental subjects, the subjects may interpret the demand characteristics of the experiment as requiring the assignment of great weight to the criminal history information.[96] Roberts concludes: "[a] less obtrusive methodology . . . would necessitate embedding the criminal record within a more complex stimulus array."[97] The present study offers such a "less obtrusive" method, whereby a broad range of information about each crime, offender, and victim is provided the subjects, who must then choose which variables to assign greater weight. The present experimental conditions would seem likely to yield substantially lower weightings for criminal record than were found in the studies reviewed by Roberts, where only the type of crime

---

[93] *See generally* Norman J. Finkel, et al., *Recidivism, Proportionalism, and Individualized Punishment*, 39 AM. BEHAVIORAL SCIENTIST 474 (1996).

[94] Julian V. Roberts, *Public Opinion, Criminal Record, and the Sentencing Process*, 39 AM. BEHAVIORAL SCIENTIST 488, 497 (1996).

[95] *Id.*

[96] *Id.* at 495.

[97] *Id.*

and the criminal record of the offender were provided respondents.

The normative structure of punishment for crimes centers on the nature of the crime committed, with little response sensitivity to offender or victim characteristics, although the characteristics of victims matter slightly more than do offenders'. Respondents' demographic characteristics count for little. Little evidence exists of either threshold effects structured along demographic lines or of alternative normative structures. In short, Rossi and Berk's Model V fits the data best. The structure of sentence length preference norms is dominated by offense type.

G. SOURCES OF ERROR

The remaining variation is "error," so far as this analysis can determine. It is not structured in relation to the respondent or offense characteristics included in this study. We now turn to analyses that seek to understand the extent to which that error is structured.[98] To accomplish this we return to the aggregate level, examining sources of variation in standard deviations around mean sentence lengths. The objective is to specify sources of consensus and dissensus in ratings of offense types.

1. Error by Offense Type

Dispersion around mean sentence lengths varied by offense type. Multiple regression analysis on standard deviations revealed that about 20% of the variation in standard deviations is accounted for by offense type: disagreement among respondents was greater for some crimes than for others.[99]

Within each offense type, much dispersion around the mean existed. This dispersion was not symmetrical, as shown by large differences between higher means and lower medians for all offenses except capital offenses. Most respondents preferred

---

[98] Miethe has severely criticized previous studies of normative structure, specifically in relation to crime seriousness, for failing to consider dispersion. Miethe, *supra* note 9, at 517.

[99] The standard deviations of sentence length for each offense type are displayed in Table 4.

shorter sentences than suggested by the means. Skewness resulted from the very long sentence preferences of a minority of respondents.[100]

2. Error by Degree of Harm

Distributions of preferred sentence lengths for the five larceny offenses, in Figure 2, reveal there was little agreement on appropriate sentence length for any of the larcenies. Each curve rises rapidly from the origin, then levels off slowly through a wide range of sentence lengths, indicating a broad distribution of responses across the entire range. There is much overlap in preferred sentence lengths between offenses. For example, the median sentence for the most serious ($10,000) larceny was thirty-six months, but over one-quarter of responses to the least serious ($10) larceny were higher than thirty-six months.

Standard deviations increased among property theft and damage offenses, though not consistently, with amount of dollar loss. Within other offense categories—burglary, robbery, assault, forcible rape, drunk driving, and drugs—there are consistently positive relationships between amount of harm and sentence dispersion.

3. Error by Respondent Characteristics

Dissensus may be patterned. Some subgroups of people may agree to the same extent (i.e., may share a norm to the same degree), as suggested by Rossi and Berk's Model VI. To test the applicability of Model VI, a multiple regression analysis was conducted on standard deviations of sentence length. An aggregated data set was created, with each case being a unique combination of offense type and offender characteristics. The independent variables were respondent characteristics and offense types, all coded as dummies. The dependent variable was the standard deviation of sentence length. In the first model all

---

[100] This discussion might appear to justify application of the geometric mean to sentence lengths, as was done with seriousness scores; however, the range of sentence lengths was only three orders of magnitude (0 - 480 months), while seriousness scores ranged over 13 orders of magnitude (.3 - 100,000,000,000).

variables were included. The resulting multiple R was .458 (R^2 = .210). In the second model only respondent characteristics were included, producing a multiple R of .079 (R^2 = .006).

The conclusion of this analysis is that most (21%) of the identifiable structure in dispersion of preferred sentence lengths is created by type of offense. Less than 1% of the dispersion was accounted for by all respondent characteristics combined. Though some demographic subgroups have punishment thresholds different from the general population, no subgroups were found that have a higher degree of consensus than the general population has.

VI. CONCLUSIONS

The purpose of this study was to map the underlying normative structure informing citizens' punishment preferences. In undertaking this task, we chose to employ a research strategy that was derived not from theories of law (consensus and conflict), but from the sociological study of normative structures (the Rossi-Berk model). A major advantage of this model is that it makes no *a priori*, ideological assumptions about public opinion. Instead, it provides a systematic scheme for assessing a wide range of possibilities regarding the underlying structure of punishment preferences. Furthermore, because the model is general (i.e., not specific to attitudes about crime or punishment), it allows future researchers to explore whether the normative structure of punishment is unique or similar to other normative domains.

A. THE STRUCTURE OF PUNISHMENT NORMS

The main finding emerging from these the analysis is that views of the public on punishments for crimes are normatively structured. The public, however, disagrees about specific levels of punishments for specific crimes. This lack of consensus obscures the underlying normative structure. Consensus exists on punishing crimes according to relative degrees of harm, but little consensus exists on absolute amounts of punishment.[101] The

---

[101] *See also* Blumstein & Cohen, *supra* note 18, at 259; Miethe, *supra* note 9, at 520; Peter H. Rossi & J. Patrick Henry, *Seriousness: A Measure for All Purposes?, in* HANDBOOK

results of this analysis are consistent with Rossi and Berk's Model V. These data do not support the more stringent conditions of Rossi and Berk's Model VI or VII, because no significant segmentation of punishment norms by respondent characteristics exists.

The following general conclusions about popular views on punishment for crime can be drawn from our analysis:

1. *Prison is the most preferred punishment*

The American public favors imprisonment as punishment for common street crimes.[102] A normative expectation exists, in short, that most offenders convicted of common street crimes will be sanctioned by a prison term.[103]

2. *Punishment preferences are differentiated most by offense type*

Severity of preferred punishment—for both type of punishment and length of imprisonment—is determined mostly by type of offense. In turn, the most influential quality of each offense type is its perceived seriousness. This finding replicates findings by Warr et al.[104] and Blumstein & Cohen.[105] Relative consensus exists on the association between punishment and both objective and subjective measures of harm. This finding is consistent with that of Thomas et al.,[106] who found rank order correlations exceeding .91 for rankings of punishment severity by sex, race, age, income, occupational prestige, and education.

OF CRIMINAL JUSTICE EVALUATION 489, 491 (Malcolm W. Klein & Katherine S. Teilmann eds. 1980).

[102] Zimmerman et al., *supra* note 67, at 130.

[103] *See* Christopher A. Innes, *Recent Public Opinion in the United States Toward Punishment and Corrections*, 73 PRISON J. 220, 227 (1993); Richard C. McCorkle, *Punish and Rehabilitate? Public Attitudes Towards Six Common Crimes* 39 CRIME & DELINQ. 240, 250 (1993); Warr, *Public Perceptions, supra* note 1, at 49-50. t is possible that support for incarceration would have been less pronounced in the NPS if the respondents had been presented with a wider range of intermediate sanctions as sentencing options. ROBERTS & STALANS, *supra* note 13, at 212-13. Further, although citizens often prefer incarceration as a response to crime, their attitudes may be flexible enough to find non-incarceration as an acceptable, if not preferred, penalty. Turner et al., *supra* note 10, at 20-21.

[104] Warr et al., *Norms, supra* note 9, at 88.

[105] Blumstein & Cohen, *supra* note 18, at 236.

[106] Thomas, et al., *supra* note 2, at 110-16.

Prison sentence length preferences vary directly with perceived offense seriousness and (for larceny offenses) with dollar loss, according to the laws of psychophysics. Mean preferred sentence length is a power function of both mean perceived seriousness and dollar loss. Tremblay[107] also reported evidence of a power function between sentence length and dollar loss.

3. Punishment preferences are influenced little by the characteristics of offenders and victims

Only offenders with long criminal records impress the public as requiring substantially more severe punishment. Very young offenders are viewed more leniently than are older offenders.[108] Other factors—such as short records of arrest, conviction, and imprisonment; employment history; drug and alcohol use; and mental health—are accorded little importance in determining appropriate penalties.

4. Respondents have different punishment thresholds

Some respondents consistently prefer higher and some lower levels of punishment. Consequently, a wide range of preferences exists for the punishment of every offense. Unfortunately, the design of the present study, where each respondent rated only eight vignettes, did not permit analysis of the consistency of these thresholds within individuals. However, in a similar study where fifty vignette ratings were collected from each respondent, Rossi et al. found that "respondents' ratings were internally more consistent than were the pooled ratings."[109] Another way to understand the empirical finding of Rossi and his colleagues, is that people are consistently "punitive" or "lenient"—in Rossi and Berk's terms, they really do have "punishment thresholds."

---

[107] Pierre Tremblay, *On Penal Metrics*, 4 J. QUANTITATIVE CRIMINOLOGY 225, 237 (1988)

[108] *See* Francis T. Cullen et al., *Public Support for Correctional Treatment: The Tenacity of Rehabilitative Ideology*, 17 CRIM. JUST. & BEHAV. 6, 12, 15 (1990). *See also* Jody L. Sundt et al., *The Tenacity of the Rehabilitative Ideal Revisited: Have Attitudes Toward Offender Treatment Changed?*, 25 CRIM. JUST. & BEHAV. 426, 437-38 (1998).

[109] Rossi et al., *Beyond Crime Seriousness, supra* note 26, at 88.

5. No normative segmentation exists

Variability of punishment thresholds does not indicate normative segmentation of the American public. No demographically structured segments of the population differ substantially in their punishment preferences.[110] Small but statistically significant differences in punishment preferences appeared by respondents' age, education, income, and geographical region. These findings are consistent with earlier work by Blumstein & Cohen,[111] who found statistically significant differences in punishment severity preferences by sex and race, as well as differences between some religious, income, and occupational groups. Rossi et al.[112] also found comparatively small effects of demographic characteristics of respondents, accounting for under 5% of the total variance in punishment preferences. Punitiveness thus appears to be more a matter of taste than of status. Furthermore, differences among respondents are overshadowed by the much greater differences associated with objective characteristics of offenses.

6. Dissensus about punishment varies by offense type

The extent to which respondents disagree is influenced by the particular offense being rated: there is greater disagreement about the punishment for more serious offenses. This disagreement is not patterned by respondents' demographic characteristics. In other words, there is no population subgroup (identifiable by age, ethnicity, sex, income, or geography) that has more internal agreement on appropriate punishments than does the population as a whole.

---

[110] This generalization does not exclude the possibility that punishment thresholds could be structured by other characteristics that were not evaluated.

[111] *See generally* Blumstein & Cohen, *supra* note 18.

[112] Rossi et al., *Beyond Crime Seriousness*, *supra* note 26, at 81.

7. *American punishment norms are structured according to Rossi and Berk's Model V—relative consensus, differentiated judgments, varying thresholds, and error*

Punishment norms are characterized by relative consensus, differentiated judgments, varying thresholds, and error:

(a) "Relative consensus"— People agree on the rank ordering of punishments for crimes, with crimes perceived as more serious deserving harsher punishment. They generally prefer imprisonment to other types of punishments for all but the most minor offenses. This consensus is "relative" to their (widely varying) internal scales of punishment, however. People do not agree on the specific punishment for specific crimes.

(b) "Differentiated judgments"—People's punishment choices are not random. Preferred punishments vary in a patterned way, influenced primarily by their perception of the seriousness of offenses, and much less by other legally relevant characteristics of offenses and offenders.

(c) "Varying thresholds"—People have a wide range of punishment "thresholds" (i.e., some people are more punitive than others).

(d) "Error"—Wide variations in punishment choices exist, even within each offense type and for offenses perceived as being equally serious.

The domain of punishment norms does not satisfy the more stringent conditions set in Rossi and Berk's Models VI and VII, where error and thresholds are hypothesized to vary with individual characteristics. In short, the normative domain of punishment for criminal offenses is structured the same as other normative domains which have been studied. Despite the intensity of people's emotional response to crime, the logic they use in matching crimes and punishments is structured the same as the logic they apply to their other concerns.

The evidence of widespread consensus on punishment norms provided in this study provides a counterpoint to the widespread disagreement about the appropriate response to a broad spectrum of social problems. When it comes down to the normative principles which should guide punishment for criminal offenders, there is consensus. The preference for incarcera-

tion as punishment for most offenses narrows the range of publicly acceptable choices considerably. The rank order of punishment preferences according to crime seriousness further narrows the range of acceptable punishments.

Rossi and Berk's set of hypothetical normative structures has proved to be a particularly useful guide to searching for the level and location of consensus on punishment. We have answered the sequential list of logically related questions implied by Rossi and Berk's model. In doing so we have systematically documented the location of consensus and identified the norms around which this consensus exists.

B. THEORETICAL AND POLICY IMPLICATIONS

Rossi and Berk provide a set of hypothetical normative structures, a way of searching for the level and location of consensus/dissensus on punishment. We have answered the sequential list of logically related questions implied by the Rossi-Berk model and, in turn, have attempted to document the location of consensus and identified the norms around which this consensus exists. The survey data reveal that people want, more than anything else, for punishment to fit crimes.[113] When given a precisely defined punishment-selection task, people choose a punishment that is proportional to the perceived seriousness of the crime.[114] Characteristics of victims and offenders are much less influential than objective harm or perceived seriousness. Further, there is a normative preference to sanction offenders with imprisonment.

These findings have implications for, but do not constitute a rigorous test of, consensus and conflict theories of criminal justice. These competing theories do not furnish precise guide-

---

[113] *See also* Hamilton & Rytina, *supra* note 9, at 1132.

[114] "Most research on perceptions of crime seriousness treat seriousness as synonymous with harm or damage. Crime seriousness has, however, been found by Warr to consist of two distinct components: "wrongfulness" (i.e., the degree of normative violation) and "harmfulness" (i.e., the amount of damage to the victim). "Seriousness" is generally interpreted as wrongfulness when a crime is perceived as being more wrong than harmful. Seriousness is interpreted as harmfulness when a crime is perceived as being more harmful than wrong. Mark Warr, *What is the Perceived Seriousness of Crimes*, 4 CRIMINOLOGY 795, 821 (1989). The present research did not explore whether respondents applied this distinction.

lines as to how much consensus or dissensus must exist for a given perspective to be confirmed. Conflict theory can also claim that any finding of consensus is really "false consciousness," which, in effect, renders this perspective non-falsifiable with survey data. Earlier studies focusing exclusively on whether the public agreed on the specific type or amount of punishment[115] or on the congruence between public preferences and actual sentencing laws[116] drew ambiguous conclusions on the conflict-consensus argument. The results reported here and in other recent work lend support to consensus theory.[117] The present study looked underneath the surface—the simple expressions of punishment preferences—and found a coherent structure to those preferences. The existence of this structure provides reassurance that the U.S. is not splintered into factions in relation to an issue as fundamental to governing as punishment for crime. Though members of the public disagree on the specifics, they agree on the general principles that should guide government in punishing criminal offenders. Absent in the patterned variation in punishment norms is evidence that any demographically defined population subgroup has a substantially different punishment threshold (i.e., is more or less punitive), adheres to a distinctively different set of punishment norms (i.e., puts unique emphasis on particular characteristics of crimes, offenders, or victims), or shares a higher degree of consensus about appropriate punishment than does the general population.[118]

---

[115] *See, e.g.,* Warr et al., *Norms, supra* note 9.

[116] *See, e.g.,* Warr et al., *Contending Theories, supra* note 2.

[117] *See e.g.,* Warr, *Public Perceptions, supra* note 1, at 49-52.

[118] We note one qualification to these findings. The NPS did not measure respondent's religious affiliation, beliefs, or practices, so punishment practices could vary by religion or religiosity. As Garland notes: "One of the reasons why the influence of religion upon punishment is so evident [historically] to the observer is that religious cultural systems are clearly articulated." DAVID GARLAND, PUNISHMENT AND MODERN SOCIETY: A STUDY IN SOCIAL THEORY 204 (1990). On a general level, Hunter contends that social science research shows "religion's declining significance as an explanatory variable. Whether one is a Protestant, Catholic, or Jew simply does not mean very much when attempting to explain variations in people's attitudes or values." JAMES DAVISON HUNTER, CULTURE WARS: THE STRUGGLE TO DEFINE AMERICA 105 (1991). If denomination *per se* no longer shapes attitudes, however, evidence is increasing that "potentially consequential divisions" are occurring within denominations between

We should caution, however, that it is not clear theoretically why social cleavages in punishment preferences should be expected. In applying conflict theory, it may be mistaken to expect that opinion would be splintered on all crime-related issues; the key consideration is whether group *interests* are markedly affected by the issue being examined. From this vantage point, general consensus might be expected about punishment preferences: Norms favoring proportionality of punishment—as opposed to discretionary, preferential treatment before the law—may, if anything, be in the interests of less powerful groups. Even imprisonment may be viewed as a governmental resource that provides powerless groups protection against crime victimization.[119]

In contrast, public opinion may be more divided by group status when the issues at stake are not whether punishment should be proportionate, but whether state power actually is applied equitably. Thus, research shows clear racial differences in opinions about the extent of racial discrimination in criminal

---

"orthodox" and "progressive" members. *Id.* at 105. The result is the "polarization of a religiously informed public culture into two relatively distinct moral and ideological camps. *Id.* at 106.

Existing criminological research lends some credence to Hunter's view. Several studies have shown that, although religious *denomination* is unrelated to attitudes toward crime, religious *fundamentalism* is positively related to punitiveness. *See, e.g.,* Harold G. Grasmick et al., *Religious Beliefs and Public Support for the Death Penalty for Juveniles and Adults* 16 J. CRIME & JUST. 59, 72-73 (1993); Brandon Applegate et al., Forgiveness and Fundamentalism: Reconsidering the Relationship between Correctional Attitudes and Religion 2 (Mar. 12, 1998) (Paper presented at the annual meeting of the Academy of Criminal Justice Sciences, Albuquerque, NM) (on file with the authors) [hereinafter Applegate et al., Forgiveness and Fundamentalism]. Even so, the relationship of fundamentalism to punitiveness is not apparent across all studies. *See, e.g.,* Applegate et al. *Determinants of Punitiveness, supra* note 23; Marla Sandys & Edmund F. McGarrell, *Beyond the Bible Belt: The Influence (or Lack Thereof) of Religion on Attitudes Toward the Death Penalty* 20 J. CRIME & JUST. 179, 186 (1997). Finally, some contradictory evidence exists, indicating that certain religious beliefs—such as belief in forgiveness—are related to *less* punitiveness. Applegate, Forgiveness and Fundamentalism, *supra,* at 24. Further research is therefore needed to clarify the extent to which various dimensions of religious belief systems influence punishment preferences.

[119] *Compare* John J. DiIulio, Jr., *The Question of Black Crime,* 117 PUB. INTEREST 3, 11-12, 23-24, 31 (1994), *with* JEROME MILLER, SEARCH AND DESTROY: AFRICAN AMERICAN MALES IN THE CRIMINAL JUSTICE SYSTEM 125-28 (1996).

punishments[120] and about police use of deadly force.[121] In any case, testing the conflict-consensus debate through public opinion is complex and will require more precise propositions about which attitudinal domains should be marked by social cleavages.

The findings we report also have implications for theories or justifications of punishment. The pervasive support for the norm of proportionality—the idea that punishment should fit the crime—suggests that citizens believe that retribution or "just deserts" should guide criminal sentencing.[122] Although this is a plausible interpretation, two qualifications may apply: First, as Warr points out, "individuals may invoke seriousness in judging appropriate punishments for reasons having nothing to do with retribution."[123] For example, the public may favor longer prison terms for those committing serious crimes because they perceive such offenders as being more dangerous, requiring longer periods of incapacitation to secure public safety.

Second, criminal sanctions are both multi-faceted and are applied to offenders over long periods. Norms of proportionality or "just deserts" may guide public views toward core features of sanctioning, but not on all of its aspects. This may be one reason why surveys that poll citizens about the goals of punishment or on "what the main emphasis of prisons should be" find support not only for retribution, but also (and at times more strongly) for rehabilitation, incapacitation, and deterrence.[124] Similarly, in a factorial vignette study, Applegate reports that, although offense seriousness is inversely related to support for rehabilitation, this relationship explains only a modest amount

---

[120] Martha L. Henderson et al., *The Impact of Race on Perceptions of Criminal Justice*, 25 J. CRIM. JUST. 447, 453-54 (1997).

[121] Francis T. Cullen et al., *"Stop or I'll Shoot:" Racial Differences in Support of Police Use of Deadly Force*, 39 AM. BEHAVIORAL SCIENTIST 449, 454-58 (1996).

[122] Warr, *Public Perceptions, supra* note 1, at 52.

[123] *Id.*

[124] Brandon K. Applegate et al., *Public Support for Correctional Treatment: The Continuing Appeal of the Rehabilitative Ideal*, 77 PRISON J. 237, 244-46 (1997); Cullen et al., *supra* note 108, at 10-13; Mark Warr & Mark Stafford, *Public Goals of Punishment and Support for the Death Penalty*, 21 J. RES. CRIME & DELINQ. 95, 99-102 (1984).

of the variation in such support.[125] Accordingly, it may be that the underlying normative structures of "rehabilitation preferences" and "punishment preferences" may be different. This is an area for future research through the application of the Rossi-Berk model.

Finally, although it seems only a short step from data on punishment to setting policy, we do not encourage a simplistic "policy by poll" approach. The sentiments of citizens are not irrelevant to setting public policy, but the exact role they should play is complex and beyond the scope of this article. More relevant here, there is a risk that even sophisticated surveys—never mind the one- or two-item polls that receive publicity in the media and are used by politicians—will be misinterpreted.[126] Thus, the willingness of respondents to select very specific sanctions does not mean that public opinion about punishment is rigidly fixed. For example, there is no firm consensus on the punishments to impose on convicted offenders. Within the broad principle that more serious crimes ought to be punished more severely, for most offenses a broad range of punishments receives support. Almost any specific punishment will find some supporters and many opponents. Imprisonment length preferences are widely dispersed; therefore setting prison terms at the means or medians of the distribution of preferences would be inconsistent with the desires of most people.[127]

Furthermore, individuals tend to manifest flexibility in their punishment preferences: they will revise views on crime control when complexities are introduced into the decision-making process. The focus group research of Doble and Doble & Klein reveals, for example, that people will go beyond retribution in

---

[125] Brandon K. Applegate et al., Specifying Public Support for Rehabilitation: A Factorial Survey Approach 197 (1997) (unpublished Ph.D. dissertation, University of Cincinnati) (on file with authors).

[126] Durham, *supra* note 18, at 9-10.

[127]Greater consensus may exist around *ranges* of punishment type and severity than around point estimates. This hypothesis has not been tested. In such a test respondents might be asked, for example, to specify the most and least severe punishments they could *tolerate* for given offenses. It remains for future research to determine if sufficient consensus exists around ranges of punishments to provide precise guidance to social policy.

matching punishments to offenders when they know more about the realities of punishment.[128] Their preference for long prison terms is moderated when they know about the costs and benefits of prison and alternative sanctions. Similarly, the experimental research of Doob & Roberts and Stalans & Diamond also shows that the more people are told about the details of a particular case, the less punitive they are.[129]

Building on this last point, there is a risk of confounding the interrelated but distinct issues of (1) public views on punishing individual offenders and (2) the complex matter of devising correctional policy. Punishment norms derived from surveys such as the NPS speak mainly to the issue of what respondents feel is an appropriate response to individual offenders. Correctional policies, in contrast, reflect not only considerations of what sanctions specific offenders (e.g., robbery versus homicide offenders) should receive, but also an array of organizational, financial, and political factors—such as levels of prison crowding, funding competing governmental needs (e.g., education), and evidence of sanction effectiveness. As a result, attempts to use opinion data to justify specific policy proposals are suspect if they rely on survey data not designed to address the complexities of the policy under consideration.

These considerations are not meant to imply that the punishment norms identified in the NPS and related research are methodological artifacts and irrelevant to a deeper understanding of American correctional policy. It would be unwise to attempt to translate public opinions directly into correctional policy. It would be equally unwise, however, to ignore the constraining influence of public opinion on public policy discourse and decision-making. Norms that favor fitting punishments to crimes, usually with prison sentences, exist as expectations

---

[128] *See generally* JOHN DOBLE, CRIME AND PUNISHMENT: THE PUBLIC'S VIEW (1987); JOHN DOBLE & JOSH KLEIN, PRISON OVERCROWDING AND ALTERNATIVE SENTENCES: THE VIEWS OF THE PEOPLE OF ALABAMA (1989).

[129] *See* Doob & Roberts, *supra* note 14, at 276; Loretta J. Stalans & Shari Seidman Diamond, Biased Recall Effects in Lay Perceptions of Judicial Leniency in Sentencing 24 (1988)(Paper presented at the Annual Meetings of the Law & Society Association in Vail, Colorado)(on file with author); *see also* Applegate et al., *Assessing Public Support*, *supra* note 14, at 528-30.

whose violation can evoke strong reaction; indeed, the cases of Willie Horton, Rodney King, and O.J. Simpson illuminate the salience of these normative expectations. Furthermore, as Scheingold points out, punishment preferences are political capital—ubiquitous, if usually quiescent, sentiments—that law-makers can either brandish against a political opponent or flaunt while claiming they are dealing with crime by passing yet another round of "get tough" legislation.[150] Accordingly, as-sessments of the public's views on criminal punishments, though of limited use for making policy, provide important in-sights into America's reaction to crime.

---

[150] *See* STUART A. SCHEINGOLD, THE POLITICS OF LAW AND ORDER: STREET CRIME AND PUBLIC POLICY 43-45, 54-57 (1984). *See also* KATHERINE BECKETT, MAKING CRIME PAY: LAW AND ORDER IN CONTEMPORARY AMERICAN POLITICS 14-27 (1997).

APPENDIX
OFFENSE VIGNETTE DIMENSIONS AND LEVELS

Vignette Construction Procedures:

In the construction of each vignette, selection was first made from Dimensions C (Larcenies) or D (All other offenses). The selection from C or D was retained regardless of the exclusion of combinations of subsequent dimensions and levels, so that every level in C and D would be contained in equal (1/24) proportions among all vignettes constructed.

For each respondent, every dimension except Dimensions C and D was sampled *with* replacement (i.e. all characteristics, except offense type, could be repeated in vignettes posed to a single respondent).

Within each dimension, except Dimensions B (Offender's sex), G (Offender's Employment History), H (Offender's Mental Condition) and I (Drug Dependence and Alcohol Abuse), every level had an equal probability of being included in every vignette. The proportionate selection of levels within Dimensions B, G, H and I are specified below.

Some dimensions have a level specified as "BLANK." When that level was selected for a dimension, no information about that dimension would be included in the vignette.

Dimension A—Offender's Age
1—The offender, a 14 year old
2—The offender, a 18 year old
3—The offender, a 22 year old
4—The offender, a 28 year old
5—The offender, a 32 year old
6—The offender, a 45 year old
7—The offender, a 65 year old
8—The offender, a

Dimension B—Offender's Sex
1—male   (70%)
2—female   (30%)

Dimension C—Offense Set #1 (Larceny Offenses)
1—stole property worth $10 from outside a building.
2—stole property worth $50 from outside a building.
3—stole property worth $100 from outside a building.
4—stole property worth $1,000 from outside a building.
5—stole property worth $10,000 from outside a building.

Dimension D—Offense Set #2 (All Other Offenses)
1—broke into a building and stole property worth $10.
2—broke into a home and stole $1,000.
3—did not have a weapon. He/she threatened to harm a victim unless the victim gave him money. The victim gave him/her $10 and was not harmed.
4—threatened a victim with a weapon unless the victim gave him/her money. The victim gave him/her $10 and was not harmed.
5—robbed a victim of $1,000 at gunpoint. The victim was wounded and required hospitalization.
6—robbed a victim at gunpoint. The victim struggled and was shot to death.
7—[INSERT DIMENSION N] intentionally injure[d] a victim. As a result, the victim died.
8—[INSERT DIMENSION N] intentionally injure[d] a victim. The victim was treated by a doctor and was hospitalized.
9—[INSERT DIMENSION N] intentionally injure[d] a victim. The victim was treated by a doctor but was not hospitalized.
10—intentionally shoved or pushed a victim. No medical treatment was required.
11—forcibly raped a victim. No other physical injury occurred.
12—forcibly raped a victim. As a result of physical injuries she died.

13—forcibly raped a victim and forced her to perform oral sex on him. No other physical injury occurred.

14—drove his/her car while drunk, but did not cause an accident.

15—drove his/her car while drunk, and caused a traffic accident where a victim was killed.

16—stole a car worth $5,000 and sold it.

17—intentionally set fire to a building, causing half a million dollars worth of damage.

18—sold cocaine to others for resale.

19—used cocaine.

Dimension E—Victim's Age

1—The victim was a 10 year old

2—The victim was a 14 year old

3—The victim was a 20 year old

4—The victim was a 30 year old

5—The victim was a 45 year old

6—The victim was a 60 year old

7—The victim was a 75 year old

8—The victim was a

Dimension F—Victim's Sex

1—male.

2—female.

3—BLANK.

Dimension G—Offender's Employment History

1—The offender was unemployed for a long time, even though he/she had tried hard to get a job.   (10%)

2—The offender has never had a steady job.   (10%)

3—The offender has held a good-paying job for several years.   (10%)

4—The offender makes his living mostly from committing crimes.   (10%)

5—BLANK   (60%)

Dimension H—Offender's Mental Condition

1—The offender had a serious mental illness.   (10%)

2—BLANK   (90%)

Dimension I—Drug Dependence and Alcohol Abuse

1—The offender was under the influence of an illegal drug
when he/she committed the offense.   (10%)

2—The offense was committed to get money to buy drugs.
          (10%)

3—The offender was drunk when he/she committed the
offense.   (10%)

4—BLANK   (70%)

Dimension J—Offender's Prior Convictions for Assault

1—The offender was never convicted before for a violent of-
fense.

2—The offender was convicted once before for a violent of-
fense.

3—The offender was convicted 3 times before for violent of-
fenses.

4—The offender was convicted 6 times before for violent of-
fenses.

5—BLANK

Dimension K—Offender's Prior Convictions for Property Offenses

1—The offender was never convicted before for stealing money
or property.

2—The offender was convicted once before for stealing money
or property.

3—The offender was convicted 3 times before for stealing
money or property.

4—The offender was convicted 6 times before for stealing
money or property.

5—BLANK

Dimension L—Previous Incarcerations

1—The offender had never been sentenced to jail or prison before.

2—The offender had served 1 previous sentence

3—The offender had served 3 previous sentences

4—The offender had served 6 previous sentences

5—BLANK

M—Length of Previous Incarcerations

1—(of or totaling) 6 months in jail.

2—(of or totaling) 1 year in jail.

3—(of or totaling) 3 years in prison.

4—(of or totaling) 5 years in prison.

5—(of or totaling) 10 years in prison.

6—BLANK

Dimension N—Weapon Used in Assaults

1—used a gun to

2—used a knife to

3—used his/her fists to

4—used a lead pipe to

5—BLANK