PUTTING DESERT IN ITS PLACE

Christopher Slobogin & Lauren Brinkley-Rubinstein*

Based on an impressive array of studies, Paul Robinson and his coauthors have developed a new theory of criminal justice, which they call "empirical desert." The theory asserts that, because people are more likely to be compliant with a legal regime that is perceived to be morally credible, a criminal justice system that tracks empirically derived lay views about how much punishment is deserved is the most efficient way of achieving utilitarian goals, or at least is as efficient at crime prevention as a system that focuses solely on deterrence and incapacitation. This Article describes seven original studies that test the most important hypotheses underlying empirical desert theory. The authors' conclusions, which throw doubt on much of empirical desert theory, include the following: (1) while consensus on the ordinal ranking of traditional crimes is relatively strong, agreement about appropriate punishments—which arguably is the type of agreement empirical desert requires in order to work—is weak; (2) the relationship between people's willingness to abide by the law and the law's congruence with their beliefs about appropriate punishment is complex and not necessarily positive; further, any noncompliance that results from the law's failure to reflect lay views about desert is probably no greater than the noncompliance triggered by a failure to follow lay views about the role utilitarian goals should play in fashioning criminal dispositions; (3) while the relative crime control benefits of a desert-based system and a prevention-based system are hard to evaluate (and are not directly examined here), people are willing to depart from desert in cases that do not involve the most serious crimes if they believe that preventive goals can be achieved in some other way. The Article ends by discussing the implications of these findings for criminal justice policy, especially with respect to determinate and indeterminate sentencing.

---

* Christopher Slobogin is Milton Underwood Professor of Law at Vanderbilt University Law School; Lauren Brinkley-Rubinstein is in the post-doctoral program at Peabody College, Vanderbilt University. For their comments on drafts of this Article, the authors would like to thank the editors of the *Stanford Law Review*, Shima Baradaran, Ed Cheng, Deborah Denno, Mary Fan, David Fontana, David Gray, Rob Mikos, Nancy King, and Dan Simon. Participants in workshops at the American Psychology-Law Society Annual Meeting, Miami, Florida, March 6, 2011; at the Law & Society Annual Meeting, San Francisco, June 4, 2011; and at Brigham Young, Maryland, and Vanderbilt University Law Schools also provided helpful comments and insight. The authors would also like to thank Vanderbilt University Law School for its generous support of this research. The raw data for the studies reported in this Article are available from the authors.

INTRODUCTION

The age-old debate among criminal law theoreticians over whether desert or prevention ought to drive criminal justice has taken on a new cast during the past few years. The old debate featured deontology against utilitarianism: put simply, should offenders be punished according to their moral blameworthiness or should concerns about protecting society be the focus of punishment? Numerous thinkers have staunchly staked out positions at opposite ends of the spectrum, while others have tried to reach some type of compromise between the two.[1] But the consistent theme in the debate has been that a theory that bases punishment on the offender's degree of culpability is frequently in significant tension with the view that punishment should focus on how subsequent offending can be prevented.[2]

In the new debate both sides are more optimistic about resolving this tension, perhaps because both sides are willing to abandon the deontological view

1. For a description of the debate, see Michael Tonry, *Introduction* to WHY PUNISH? HOW MUCH?: A READER ON PUNISHMENT 3, 3 (Michael Tonry ed., 2011).

2. *See id.* at 7 ("The fundamental contrast between the ideas that punishment is morally justified because people have behaved wrongly and that punishment is morally justified only when it has good consequences has long existed and most likely always will.").

of desert. In the new debate the value of a culpability-based punishment system is no longer assessed through the prism of moral philosophy but rather in terms of its utility at achieving the goals of the criminal justice system. Thus, where the two sides differ is not over methodology but over whether crime control and respect for the law is best achieved through a system focused on desert or through a prevention-based regime that is sensitive to desert only when ignoring it would have criminogenic impact.

Most of the literature in this new debate has favored the first stance. The principal proponent of that view has been Paul Robinson, who has coined the term "empirical desert" to capture the idea that a criminal justice system that tracks empirically derived societal views of desert may best facilitate the law's ability to assure compliance with legal prohibitions.[3] Relying on a considerable amount of research, much of which he has helped conduct, Robinson contends that "a criminal justice system that distributes liability and punishment in concordance with the citizens' shared intuitions of justice . . . may provide greater utility than a distribution following the more traditional instrumentalist approach of optimizing deterrence or incapacitation."[4]

The competing view is not as well developed, and in fact has yet to be explicitly articulated. This Article aims to provide that articulation. The theory is that while liability rules should still depend primarily on desert, punishment rules that focus on the utilitarian goals of specific deterrence, rehabilitation, and incapacitation ("individual prevention" goals[5]) are not only superior at accomplishing crime prevention but can also usually assuage society's urge for retribution well enough to avoid vigilantism, norm breakdown, and other negative effects. This Article presents original research that supports this approach to punishment, which could be called "preventive justice."

Both empirical desert and preventive justice are leery of deontological theories. Both also try to reconcile retributive instincts with more explicitly preventive goals. The distinction between the two lies in how they evaluate the utility of desert. This Article's investigation suggests that empirical desert exaggerates desert's utility as a crime prevention mechanism while unduly minimizing the efficacy of preventive justice. When added to the other possible deleterious effects of reliance on desert as the linchpin of punishment policy—effects ranging from unusually harsh sentences to expensive imprisonment schemes—the unclear preventive payoff associated with empirical desert may mean that preventive justice is a superior approach from a utilitarian

3. *See* PAUL H. ROBINSON, DISTRIBUTIVE PRINCIPLES OF CRIMINAL LAW: WHO SHOULD BE PUNISHED HOW MUCH? 139-40 (2008).

4. *Id.* at 175-76.

5. These utilitarian goals, which focus on reducing recidivism in individual offenders, are meant to be distinguished from general prevention goals, the best example of which is general deterrence. *See generally* RICHARD J. BONNIE ET AL., CRIMINAL LAW 25-37 (3d ed. 2010) (categorizing specific deterrence, incapacitation, and rehabilitation as "individual prevention" goals, and distinguishing those goals from general deterrence).

perspective. If so, a sentencing regime that favors broad and flexible dispositional ranges should be favored over one that is centered on desert as defined by surveys of the public.

Part I of this Article describes empirical desert theory in more detail and summarizes it in terms of three testable hypotheses. The first hypothesis is that societal consensus about punishment exists, at least for "core crimes" such as homicide and robbery, and that this consensus is based on desert considerations (the consensus hypothesis). The second hypothesis is that a failure to adhere to this consensus will induce dissatisfaction with the law, which in turn will increase noncompliance with it (the compliance hypothesis). The third hypothesis is that the noncompliance produced by a punishment regime that departs substantially from empirically derived desert will exceed the noncompliance such a regime would prevent, even if the regime is designed to further utilitarian goals such as general deterrence, incapacitation, and rehabilitation (the crime control hypothesis).

Part II reports seven studies that test these hypotheses from various angles. The first two studies explore the consensus hypothesis. The results of the first study suggest that to the extent consensus about the ordinal ranking of crimes exists, it is based as much on utilitarian considerations as on desert.[6] And the findings of the second study indicate that consensus about cardinal, as opposed to ordinal, punishment is hard to come by even with respect to the most common crimes, a result that may doom empirical desert's objective of constructing a punishment system that reflects societal views.[7] The next three studies call into question the compliance hypothesis, because they find little evidence that divergence from societal views about punishment—whether desert- or utilitarian-based—significantly increases willingness to break the law or diminishes respect for it, especially once the passage of time makes the divergence less salient.[8] Further, to the extent that these studies do find such a relationship, they suggest—contrary to the third (crime control) hypothesis—that noncompliance is just as likely when the criminal law departs from utilitarian goals. The final two studies begin to investigate a counterhypothesis to the crime control hypothesis: that punishment explicitly focused on utilitarian goals can satisfy retributive urges.[9] They suggest that information about plausible treatment programs directly affects assessments of desert and creates a tendency to prefer indeterminate over determinate sentences.[10]

Part III, which discusses the implications of this research, concentrates on these latter findings. Punishment focused on prevention is unlikely to cause major dissatisfaction with the justice system, or more noncompliance with the law,

6. *See infra* Part II.A.1.
7. *See infra* Part II.A.2.
8. *See infra* Part II.B.
9. *See infra* Part II.C.
10. *See infra* Tables 5-7.

than it preempts through incapacitation, rehabilitation, and deterrence,[11] unless it represents a radical departure from desert (which is most likely in connection with very serious crimes). Thus, from a utilitarian perspective—the perspective taken by empirical desert—desert should probably play a secondary role. The optimal sentencing system from a crime control standpoint is likely to be one focused explicitly on prevention goals, so long as doing so is not so antithetical to desert that it creates a real possibility of social disruption.

I. EMPIRICAL DESERT THEORY

Joe has just been convicted of armed robbery. What should happen to him now? A retributivist—one who thinks desert should drive the punishment decision—would want Joe's disposition to be proportionate to his culpability, as measured by the relative gravity of the offense, the blameworthiness of the offender, and the harm to the victim or victims.[12] Most desert theorists would probably conclude that armed robbery falls between murder and burglary on the gravity and harm spectra. So if a murderer deserves a twelve-year sentence and a burglar a two-year term, Joe might receive seven years in prison or an equivalently punitive disposition, perhaps boosted upward by a couple of years if he was the mastermind of the heist, or downward by a few years if he appeared to be a bit player in the crime or was cajoled into it.

A utilitarian—one who is interested primarily in crime prevention—might arrive at a quite different sentence, depending upon how costs and benefits are calculated. A desert-based sentence for someone like Joe might be unnecessarily long or too short to achieve general deterrence.[13] Furthermore, if Joe is a career criminal he may require enhanced confinement, whereas if he was simply a tagalong first-time offender or repents his actions in a way that suggests he will not reoffend, the utilitarian might significantly reduce or conceivably even eschew the confinement that desert theory would require for armed robbery (although some form of community supervision would probably be warranted).[14]

---

11. As noted earlier, *see supra* note 5, this Article's references to "preventive" purposes of punishment are meant to focus on the "individual" prevention goals of incapacitation, rehabilitation, and specific deterrence, rather than general deterrence. The assumption throughout this Article is that, while some sort of punishment is crucial to deterring crime, no particular punishment is necessary to ensure that goal. As Michael Tonry summarized the research of three National Academy of Sciences panels, "[i]maginable increases in severity of punishments do not yield significant (if any) marginal deterrent effects." Michael Tonry, *Purposes and Functions of Sentencing*, 34 CRIME & JUST. 1, 28 (2006).

12. This is a standard account of the criteria retributivists use in calculating desert. *See, e.g.*, MODEL PENAL CODE: SENTENCING § 1.02(2)(a)(i) (Tentative Draft No. 1, 2007).

13. On the difficulties of assigning sentences based on general deterrence theory, see BONNIE ET AL., *supra* note 5, at 21-23.

14. As Herbert Packer put it, "the incapacitative theory is at its strongest for those who, in retributive terms, are the least deserving of punishment." HERBERT L. PACKER, THE LIMITS OF THE CRIMINAL SANCTION 50-51 (1968).

If Joe's involvement in the robbery were linked to treatable problems—such as unemployment, mental disability, or substance abuse—a utilitarian would also consider rehabilitation programs that would last as long as necessary to remedy the problem, a period that might be shorter or longer than a desert-based disposition and which, again, would not necessarily involve confinement.[15]

This simple example exposes the tension between desert-based retributive theory and prevention-oriented utilitarian goals. The best known attempt to construct a system that diminishes the conflict is Norval Morris's limiting retributivism, which sets sentencing ranges based on relative desert but allows utilitarian objectives to determine the precise duration of a sentence within the range.[16] Asymmetrical desert, which defines the maximum sentence according to desert but permits downward departures based on preventive goals, is another effort at reconciliation.[17] Under these hybrid approaches, Joe might receive a sentence somewhere between five and nine years (depending on his individual culpability) or might simply be given a maximum sentence of nine years, with the precise time of release dictated by the risk of reoffending that he poses.

These hybrids can satisfy many mixed theorists. But they frustrate those policymakers, scholars, and others who occupy the ends of the spectrum: those near the desert pole who contend that retribution permits at most a narrow band of punishment for a given level of culpability,[18] and those near the utilitarian pole who believe that desert-based confinement is often counterproductive for the offender now willing to be law abiding, as well as insufficiently protective of society when dangerous offenders are released after they have finished the sentence they "deserve."[19] As Andrew von Hirsch has stated, "[a] patched-together compromise . . . between ideas of prevention and those of equity would be likely to satisfy neither those preoccupied with crime prevention, nor those concerned seriously with questions of justice."[20]

---

15. *See* WAYNE R. LAFAVE, CRIMINAL LAW 33 (5th ed. 2010) ("The rehabilitation theory would let the criminal go when (and perhaps *only* when) he had been reformed." (footnote omitted)).

16. *See* NORVAL MORRIS, THE FUTURE OF IMPRISONMENT 59 (1974); *see also* Richard S. Frase, *Sentencing Principles in Theory and Practice*, 22 CRIME & JUST. 363, 363-78 (1997) (describing in detail Morris's theory of limiting retributivism).

17. *See* K.G. Armstrong, *The Retributivist Hits Back*, *in* THE PHILOSOPHY OF PUNISHMENT: A COLLECTION OF PAPERS 138, 155 (H.B. Acton ed., 1969) ("For a variety of reasons (amongst them the hope of reforming the criminal) the appropriate authority may choose to punish a man less than it is entitled to, but it is never just to punish a man more than he deserves.").

18. *See, e.g.*, Andrew von Hirsch, *Equality, "Anisonomy," and Justice: A Review of* Madness and the Criminal Law, 82 MICH. L. REV. 1093, 1105-06 (1984) (suggesting a range of two to three years for a first offense of armed robbery).

19. *See, e.g.*, Michael Vitiello, *Reconsidering Rehabilitation*, 65 TUL. L. REV. 1011, 1051, 1053 (1991) ("[W]hen transformation occurs it is difficult to find a continued justification for imposing suffering on that offender.").

20. Andrew von Hirsch, *Recent Trends in American Criminal Sentencing Theory*, 42 MD. L. REV. 6, 16 (1983).

Enter empirical desert theory. The theory first surfaced in a book Paul Robinson published with John Darley in 1995, entitled *Justice, Liability, and Blame.*[21] Using artfully crafted crime scenarios that would become the staple of Robinson's empirical work, the research reported in the book revealed significant discrepancies between the liability gradations lay people prefer and those endorsed in the common law and modern statutory templates like the Model Penal Code. For instance, contrary to both the common law and the Model Penal Code, Robinson and Darley's subjects on average wanted to impose liability for omissions even in the absence of a duty and also wanted to impose liability for reckless complicity, but rejected full principal liability for purposeful complicity.[22]

Robinson and Darley argued that both desert theorists and utilitarians should pay attention to these types of findings. For the first group, Robinson and Darley suggested, surveys of community sentiment could be "one source of determining what counts as the just desert" and "should be of considerable utility" in resolving disputes between policymakers as to the moral intuitions of their constituencies.[23] Meanwhile, they noted, utilitarians might find the work useful because a substantial gap between community sentiment and the criminal law could reduce the law's "moral credibility," and thus people's willingness to comply with it, not only in everyday life but also when they become directly involved in the criminal process as jurors, witnesses, and offenders.[24]

Robinson repeated these two points in later work. He has elaborated on the first point by distinguishing between "deontological desert" and "empirical desert."[25] While only the former version of desert "can give us the truth of what is deserved, insulated from the vicissitudes of human irrationality and emotions,"[26] deontology is also subject to the biases and idiosyncratic intuitions of philosophers.[27] Thus, "it may be that empirical desert offers the best practical approximation of deontological desert."[28]

---

21. PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW (1995).

22. *Id.* at 42-50 (omissions study); *id.* at 96-105 (complicity study); *id.* at 14-28 (attempt study).

23. *Id.* at 6, 214.

24. *See id.* at 6.

25. Paul H. Robinson, *Competing Conceptions of Modern Desert: Vengeful, Deontological, and Empirical*, 67 CAMBRIDGE L.J. 145, 148-50 (2008).

26. *Id.* at 167.

27. *See id.* at 164 (noting that "moral philosophers simply disagree about just how [moral blameworthiness] translates into specific punishment in a given case").

28. Paul H. Robinson, *The Ongoing Revolution in Punishment Theory: Doing Justice as Controlling Crime*, 42 ARIZ. ST. L.J. 1089, 1110 (2011) [hereinafter Robinson, *Doing Justice as Controlling Crime*]; *see also* Paul H. Robinson, *The Role of Moral Philosophers in the Competition Between Deontological and Empirical Desert*, 48 WM. & MARY L. REV. 1831, 1840 (2007) [hereinafter Robinson, *The Role of Moral Philosophers*] (suggesting that empirically sound research methodologies in social psychology can provide moral philosophers with "more reliable" information about intuitions of justice).

Robinson has also expanded substantially on his thesis that "the crime-control power of the criminal law depends in some significant part upon how well it tracks the community's shared intuitions of justice."[29] According to Robinson, tracking these shared intuitions earns the system moral credibility, which in turn has several positive effects. A morally credible system, he argues, is better able to minimize vigilantism by people unhappy with its results, encourage the citizen cooperation that the criminal process needs to function efficiently, and "harness the power of stigmatization" as a means of influencing people's conduct.[30] Such a system also has "a greater chance of gaining compliance in borderline cases where the actual condemnability of the conduct may be unclear" and influences "the larger public conversation by which societal norms are shaped."[31] Robinson concludes in another piece that "strong arguments suggest greater utility in a distribution based on shared intuitions of justice than in a distribution based upon optimizing deterrence, rehabilitation, or incapacitation."[32]

Several commentators have challenged these arguments. Many of these challenges are theoretical in nature.[33] Of more relevance to this Article are the challenges to the view that empirical desert has strong instrumental effects apart from the deterrence and incapacitation that inevitably flow from punishment. Douglas Berman suggests that reversal of duly enacted laws based on an elite's empirical study of lay views, as Robinson apparently proposes, is more likely to undermine the criminal justice system's credibility than bolster it.[34] Others wonder whether society's views about something as complex as punishment can be reliably ascertained, and how the law should respond to

29. Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1107.

30. *Id.* at 1106.

31. *Id.*

32. Robinson, *The Role of Moral Philosophers*, *supra* note 28, at 1836.

33. For instance, Mary Sigler disputes the relevance of empirical desert to deontological desert, since philosophers are not trying to "measure or describe something as it is," and argues further that empirical desert's positive nature provides no means of reconciling the disagreements that, in a democracy, are inevitable. Mary Sigler, *The Methodology of Desert*, 42 ARIZ. ST. L.J. 1173, 1181, 1185-87 (2011); *see also* Zachary R. Calo, *Empirical Desert and the Moral Economy of Punishment*, 42 ARIZ. ST. L.J. 1123, 1131 (2011) ("[E]mpirical desert . . . has no resources for explaining why an offender deserves punishment other than that doing so achieves [utilitarian] benefits."). Others, pointing to our racist history and similarly undesirable motivations to punish, have noted that societal views may diverge radically from deontologically sound prescriptions. *See, e.g.*, Youngjae Lee, *Desert, Deontology, and Vengeance*, 42 ARIZ. ST. L.J. 1141, 1149 (2011) ("Punitive passions . . . can also be excessive and driven by other less desirable, yet no less common, sentiments such as cruelty, sadism, inhumanity, and racial hatred and prejudice."). Robinson recognizes this potential problem and states that empirical desert should not always trump deontological desert (although he does not provide a methodology for deciding when trumping should occur). *See* ROBINSON & DARLEY, *supra* note 21, at 80, 205.

34. Douglas A. Berman, *A Truly (and Peculiarly) American "Revolution in Punishment Theory,"* 42 ARIZ. ST. L.J. 1113, 1118 (2011).

inevitable changes in those views.[35] Several scholars have called for more proof bolstering the assertions that there are widely shared intuitions about punishment and that failing to heed these intuitions will increase social disruption and noncompliance with the law.[36]

In this Article we join the latter group in calling for such proof, and for the first time provide evidence that directly challenges empirical desert theory.[37] Given the empirical nature of Robinson's argument, our approach is empirical as well, and relies as much as possible on the same types of materials that Robinson and his coauthors have used. In an earlier article, one of us proposed ten hypotheses that test the assumptions underlying empirical desert theory.[38] Here we reduce these to three: the consensus hypothesis, the compliance hypothesis, and the crime control hypothesis.[39]

---

35. *See* Calo, *supra* note 33, at 1133-34; *see also* Deborah W. Denno, *The Perils of Public Opinion*, 28 HOFSTRA L. REV. 741, 744 (2000) ("[P]ublic opinion research may fail to measure accurately the public's fundamental values or 'moral intuitions.' . . . [I]t may be far more successful in reflecting individuals' erroneous knowledge and distorted attitudes, which can range enormously depending on demographics and personal experiences.").

36. *Cf.* Donald Braman et al., *Some Realism About Punishment Naturalism*, 77 U. CHI. L. REV. 1531, 1551-56 (2010) (discussing "empirical evidence contradicting the central claim that evaluations of serious wrongfulness do not vary across social conditions or individuals"); Adam J. Kolber, *How to Improve Empirical Desert*, 75 BROOK. L. REV. 433, 434, 441-48 (2009) (claiming that Robinson has "cherry-pick[ed] . . . intuitions that are elicited at a particular level of abstraction in ways that hide some of our more passionate intuitions that, for better or worse, motivate many people"); Alice Ristroph, *Third Wave Legal Moralism*, 42 ARIZ. ST. L.J. 1151, 1165, 1167 (2011) (stating that Robinson's scenarios do not simulate real crimes and questioning whether a failure to follow desert causes significant "social disutility").

37. Braman et al., *supra* note 36, provides some empirical evidence that challenges the consensus hypothesis at the margins, but in the main the attack is conceptual.

38. *See* Christopher Slobogin, *Some Hypotheses About Empirical Desert*, 41 ARIZ. ST. L.J. 1189 (2011).

39. The ten hypotheses were: (1) that rank ordering of core crimes will vary according to factors other than desert; (2) that significant disagreement exists about appropriate punishments for core crimes; (3) that for crimes outside the core, there is significant variation in both the rank orderings and the appropriate punishment; (4) that "[m]ost departures from empirical desert will not be noticed by the public"; (5) that those departures that are noticed will not occasion negative reaction because they are perceived to achieve other legitimate goals, or (6) will occasion negative reaction only because they depart from goals other than desert; (7) that criminal laws that depart from empirical desert will not cause noncompliance, or (8) will cause noncompliance only in connection with the specific laws at issue; (9) that any noncompliance that does result from departures from empirical desert will be less than the noncompliance resulting from failure to directly address preventive goals; and (10) that preventive dispositions can often satisfy desert. *Id.* Hypotheses 1-3 are related to the consensus hypothesis described in the text, hypotheses 5-7 are related to the compliance hypothesis, and hypothesis 9 is the converse of the control hypothesis. Hypotheses 4, 8, and 10 are not addressed by the research in this Article, although they are addressed by other research. *See infra* note 109 (hypothesis 4); *infra* notes 83-84 (hypothesis 8); *infra* note 138 (hypothesis 10).

The *consensus* hypothesis associated with empirical desert theory is that there is a strong societal consensus about deserved punishment, at least with respect to core crimes such as homicide, assault, and theft. Significant support for this hypothesis comes from a 2007 article in which Robinson and Robert Kurzban canvassed literature showing a surprising degree of consistency, both in this country and elsewhere, in how subjects ranked the relative severity of these types of crimes.[40] They also reported their own study, which, they asserted, found an "astounding level of agreement" on the ordinal ranking of a wide array of different homicide, assault, theft, and fraud offenses.[41] In the 2007 article and elsewhere, Robinson further asserts that this type of agreement is based on an assessment of desert rather than on factors having to do with dangerousness or deterrence.[42]

The *compliance* hypothesis is that if the criminal law fails to adhere to this consensus, noncompliance with the criminal law will increase. As noted above, Robinson argues that, in the wake of such failure, the law's moral credibility will suffer, and as a result will be diminished in its ability to promote and shape norms, forestall vigilantism, and enhance cohesion.[43] Initially, Robinson's support for this hypothesis came largely from the work of Tom Tyler, whose book *Why People Obey the Law* reported research finding a strong correlation between respect for authority and compliance with the law.[44] More recently, Robinson has relied on studies that directly explore whether unjust results increase willingness to flout the law[45] and on his own research, conducted with two coauthors, purporting to show the "disutility of injustice."[46]

The *crime control* hypothesis is that the noncompliance resulting from a failure to follow desert will be at least as great as the noncompliance that results from a failure to implement preventive goals directly through a system

40. *See* Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829, 1848-65 (2007).

41. *Id.* at 1867-72.

42. *See id.* at 1832; *see also* Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1105 ("When . . . asked to assign punishment, [laypersons] don't look to the factors that determine dangerousness or deterrence, but rather to the offender's moral blameworthiness.").

43. Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1107 ("[T]he crime-control power of the criminal law depends in some significant part upon how well it tracks the community's shared intuitions of justice.").

44. *See* ROBINSON & DARLEY, *supra* note 21, at 7 (citing TOM R. TYLER, WHY PEOPLE OBEY THE LAW (1990)); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 471, 474-75 (1997) (citing TYLER, *supra*).

45. Paul H. Robinson & John M. Darley, *Intuitions of Justice: Implications for Criminal Law and Justice Policy*, 81 S. CAL. L. REV. 1, 26 (2007) (citing Janice Nadler, *Flouting the Law*, 83 TEX. L. REV. 1399, 1417 (2005); Erich J. Greene, Effects of Disagreements Between Legal Codes and Lay Intuitions on Respect for the Law, at iv-v (June 2003) (unpublished Ph.D. dissertation, Princeton University) (on file with authors)).

46. Paul H. Robinson, Geoffrey P. Goodwin & Michael D. Reisig, *The Disutility of Injustice*, 85 N.Y.U. L. REV. 1940 (2010).

that ignores empirical desert or treats it as a secondary consideration.[47] Robinson has yet to offer any direct support for this proposition; to date, he and other researchers have focused on the noncompliance effects of what he calls "injustice,"[48] not on a comparison of those effects to the effects on crime of preventive dispositions that do not follow desert-based formulas. Ideal research testing this third hypothesis would need to find in the real world or construct in the laboratory desert and utilitarian regimes in their pure form, and then measure their efficacy at "crime control"—a term that, for Robinson, entails prevention not only of criminal law violations but also minimization of both civil disobedience and the deterioration of norms.[49]

The project reported in this Article is designed to probe each of these three hypotheses more deeply. While the research underlying empirical desert is generally sophisticated, it fails to test fairly the impact of utilitarian considerations on lay decisions about punishment. When those considerations are injected into the research agenda, a different sense about all three hypotheses begins to emerge. To the extent consensus about punishment exists, its contours look quite different when risk-related factors are thrown into the mix. Similarly, to the extent noncompliance can result from failing to adhere to punishment consensus, it is just as likely to occur when the law ignores preventive objectives as when it ignores desert. In any event, our research also leads to the conclusion that both the consensus and the compliance hypotheses are only weakly sustained. Finally, our research suggests that the best way to reconcile the tension between desert and preventive considerations may well be to focus primarily on the latter rather than the former. From a utilitarian perspective, a system aimed at prevention, only loosely constrained by desert, may well be optimal.

II. THE RESEARCH

To test the consensus, compliance, and crime control hypotheses, we conducted seven studies. Like previous studies on empirical desert, all are based on survey data, not field research. With two exceptions, the sample populations were solicited via the Internet and answered our queries online. All of the samples were reasonably demographically diverse.[50] In large part, the results of these studies fail to support the three hypotheses.

---

47. *See, e.g.*, *id.* at 1948 ("The shift to a desert distribution—specifically, empirical desert—will not seriously undermine the criminal justice system's crime-control effectiveness, and indeed may enhance it . . . .").

48. *See, e.g.*, *id.* at 1947 ("[A] distribution of liability and punishment that tracks lay intuitions of justice would significantly reduce the injustice now present.").

49. *See id.* at 1948.

50. As revealed below in Appendix D, except for Study 5 (which used law students) and Study 7C (which did not include any African Americans), our participants represented a wide range of races, ages, political affiliations, income categories, and education. Note, however, that in one of our samples (used in Studies 1, 2, 3, and 6), the gender split (63% male, 37% female) was far from even. For a discussion of the difficulties inherent in

A. *Hypothesis I: Consensus*

Robinson and Kurzban's results reporting consensus on the ordinal ranking of core crimes relied on twenty-four simple fact patterns depicting the offender's act and mental state.[51] Here are two examples, from opposite ends of the severity spectrum:

> The owner has posted rules at his all-you-can-eat buffet that expressly prohibit taking food away; patrons can only take what they eat at the buffet. The owner has set the price of the buffet accordingly. John purchases dinner at the buffet, but when he leaves he takes with him two whole pies to give to a friend.
>
> . . . .
>
> John kidnaps an 8 year-old girl for ransom, rapes her, then records the child's screams as he burns her with a cigarette lighter, sending the recording to her parents to induce them to pay his ransom demand. Even though they pay as directed, John strangles the child to death to avoid leaving a witness.[52]

When asked to indicate the relative amount of punishment John deserved in twenty-four such scenarios, virtually all of Robinson and Kurzban's subjects ranked them in the same order.[53] From these results and their survey of other research,[54] Robinson and Kurzban concluded that "human intuitions of justice about core wrongdoing . . . are deep, predictable, and widely shared."[55] Robinson and Kurzban also imply that these intuitions about punishment are based on blameworthiness and nothing else.[56] In another work, Robinson makes the latter point even more explicitly: "Laypersons see punishment as something that is properly imposed according to desert, that is, blameworthiness. When . . . asked to assign punishment, they don't look to the factors that determine dangerousness or deterrence, but rather to the offender's moral blameworthiness."[57]

All of this may be true, but Robinson's empirical research does not clearly establish it. Note in particular two aspects of Robinson and Kurzban's

---

assessing "ordinary" citizens' views about punishment, see Denno, *supra* note 35, at 747-58, 762-64 (arguing, in the course of discussing a Robinson and Darley study, that various demographic variables, including race, gender, education, and political views can "substantially impact . . . results").

51. Robinson & Kurzban, *supra* note 40, app. A at 1894-98 (describing scenarios).

52. *Id.* app. A at 1894, 1898.

53. *Id.* at 1869-71.

54. *Id.* at 1848-65.

55. *Id.* at 1892. Robinson and Kurzban also suggest that, given the depth of this agreement, it is "unrealistic to expect the population to all 'rise above' its desire to punish wrongdoers, or to expect the government to 'reeducate' people away from their interest in punishing wrongdoers." *Id.*

56. *Id.* (referring to "intuitions of justice" solely in terms of blameworthiness); *see also* John M. Darley, Kevin M. Carlsmith & Paul H. Robinson, *Incapacitation and Just Deserts as Motives for Punishment*, 24 LAW & HUM. BEHAV. 659, 659 (2000) (reporting a study finding that desert predominates over incapacitative concerns).

57. Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1105.

methodology. First, as illustrated by the two examples given above (and ten other scenarios reproduced in Appendix A), the information given to their survey participants consisted solely of facts about John's conduct and his mental state. In other words, in deciding how scenario-punishments should be ranked, the subjects were not provided any information that might be relevant to utilitarian concerns—information about John's relative risk or treatability, for instance—but instead were given only those facts most relevant to retributive judgments. Second, Robinson and Kurzban framed the punishment question in terms of what John "deserve[d],"[58] which again strongly suggested to the subjects that they were to consider only desert, not utilitarian factors, in coming to their conclusions about punishment.[59]

We wondered whether the rankings that Robinson and Kurzban obtained would persist if people were given additional facts that were not as closely tied to blameworthiness. Thus, in the first study that we report below, our hypothesis, contrary to Robinson's assertion, is that people asked to assign punishments think about many other factors besides desert when given the chance. And if that is true, then either the consensus about ordinal rankings might not be as robust as Robinson and Kurzban posit, or any consensus that does exist might rest as much or more on utilitarian concerns. (Whether it is true, as Robinson and Kurzban assert, that punishment assigned on utilitarian grounds has nothing to do with "justice" is a topic we defer until Part III.)[60]

Furthermore, consensus about ordinal ranking does not mean there is consensus about the precise punishment meted out in particular cases. Robinson does not contest this commonsense assertion and, indeed, some of his earlier research supports it.[61] Nonetheless, we felt it was important to investigate further, in the second study reported below, the nature of this dissensus over the appropriate absolute punishment because of the interaction of Robinson's consensus hypothesis with his compliance hypothesis. The latter hypothesis states that dissatisfaction with criminal punishments can lead to greater noncompliance with the law. Presumably people can be dissatisfied not only with the ordinal ranking of crimes within a given system, but also with the absolute

58. Robinson & Kurzban, *supra* note 40, app. A at 1898 (describing instructions to subjects).

59. Similar comments can be made about the study that Robinson conducted with John Darley and Kevin Carlsmith, which purported to find that desert, not dangerousness, is the primary determinant of people's judgments about punishment. *See* Darley et al., *supra* note 56, at 659, 663 (describing the instructions given to survey participants and the conclusions from their study); *see also* Slobogin, *supra* note 38, at 1190 & n.6 (criticizing the study).

60. *See infra* text accompanying notes 154-161.

61. *See* ROBINSON & DARLEY, *supra* note 21, app. A at 226 (reporting that participants' punishment choices resulted in standard deviations of over 3.50 in 20% of the presented scenarios); *see also* Alfred Blumstein & Jacqueline Cohen, *Sentencing of Convicted Offenders: An Analysis of the Public's View*, 14 LAW & SOC'Y REV. 223, 223 (1980) (reporting a study that obtained agreement about "*relative* severity of sentences to be imposed . . . but disagreement over the *absolute* magnitude of these sentences").

punishment that is imposed after conviction. If there is widespread disagreement as to the latter, the empirical desert project may be in trouble from the start because there is no societally accepted baseline to which the criminal justice system can adhere even if it wants to track lay views on punishment.

1. *Study 1: Is there consensus about the ordinal ranking of crimes?*

Methodology. Our first study made use of twelve scenarios drawn from the twenty-four Robinson and Kurzban developed for their study. We administered these twelve scenarios (described in Appendix A) to 530 demographically diverse subjects whom we contacted through an online service.[62] Our control group, consisting of just under half the subjects (264), received the twelve scenarios unchanged. Our experimental group, consisting of the other 266 subjects, received the same twelve scenarios but with six of them modified to reflect utilitarian considerations. Thus, in a scenario involving a T-shirt theft, the experimental subjects were told, "After John is caught, he becomes remorseful. He pays the owner for the T-shirt and apologizes." In a scenario involving theft of a drill, they were told, "John has previously served time for one other, similar theft," and in the scenario involving an assault at a football game they read that "John goes home, realizes he acted rashly, and signs up for an anger management therapy program that he successfully completes." Appendix A describes the manipulations in full.

All of these manipulations were designed to introduce factors having to do with relative dangerousness or treatability—utilitarian considerations—into the experimental group's calculus. It might be argued that at least some of these manipulations (for instance, the expression of remorse in the T-shirt theft scenario or the prior crime in the drill theft scenario) are factors relevant to desert analysis as well.[63] However, Robinson himself considers apology, remorse, restitution, bad character, and postoffense changes in behavior to be "extra-

62. This Study and Studies 2, 3, and 6 made use of the eLab, run by Vanderbilt University's Owen Graduate School of Management. As indicated on eLab's website, "eLab researchers have leveraged use of a diverse international consumer panel to conduct cutting-edge research through robust web-based experiments and surveys." ELAB AT VANDERBILT UNIV., http://elab.vanderbilt.edu (last visited Dec. 19, 2012). The demographics for the samples are found in Appendix D, below. The control and experimental demographics are not reported independently because they are essentially the same here and in the other studies reported in this Article.

63. On the relationship of remorse to retribution, see Michael Wenzel et al., *Retributive and Restorative Justice*, 32 LAW & HUM. BEHAV. 375, 378 (2008) (distinguishing retribution as backward-looking and remorse as restorative). On the relationship of prior criminal acts to retribution, compare ANDREW VON HIRSCH, PAST OR FUTURE CRIMES: DESERVEDNESS AND DANGEROUSNESS IN THE SENTENCING OF CRIMINALS 88-91 (1985) (arguing that repeat offenders deserve somewhat greater punishment), with GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW 460-66 (1978) (questioning whether a prior record should increase an offender's culpability to any degree).

legal punishment factors" that are at best tangential to desert.[64] Even if a manipulation is as consistent with desert as it is with prevention, any deviation from a purely desert-based punishment that it produced would, at the least, leave unclear whether desert is the dominant basis for ascribing punishment that Robinson says it is. Further parsing of the effect attributable to particular manipulations is discussed below.[65]

To make the survey participants' jobs easier, we gave them the twelve scenarios in pairs, with each pair in the experimental group including one manipulated scenario (for the pairings, see Appendix A). We then asked each subject to indicate which offender within the pair should receive the most punishment. We expected that the manipulations given to the members of the experimental group would affect the extent to which their ordinal ranking of the pairs corresponded with the rankings obtained in Robinson and Kurzban's study.

Results. Our expectations were partially met. As indicated in the third column of Table 1, our control group—the group that received scenarios identical to those used by Robinson and Kurzban—overwhelmingly ranked the pairs in the same order as Robinson and Kurzban's subjects, thereby confirming the results of their study.[66] But as the last column of Table 1 illustrates, the majority of the subjects in the experimental group reversed the ranking in three of the six pairs (involving thefts and assaults). Moreover, in one of the three remaining pairs (involving petty theft), the percentage of experimental subjects who ranked the scenarios differently from both Robinson and Kurzban's group and our control group increased from 6% to 41%. In the final two pairs, neither the ranking nor the percentage willing to change ranking changed appreciably, although in both, the small change that did occur was in the direction we

64. Paul H. Robinson et al., *Extralegal Punishment Factors: A Study of Forgiveness, Hardship, Good Deeds, Apology, Remorse, and Other Such Discretionary Factors in Assessing Criminal Punishment*, 65 VAND. L. REV. 737, 739 (2012) (listing as "extralegal punishment factors" apology, remorse, history of good or bad deeds, public acknowledgement of guilt and various other factors that "go beyond the factors that the criminal law formally recognizes" because they are not "desert-based factors" having to do with the "seriousness of the harm or the evil of the offense and an offender's culpability and mental capacity"); *see also* Dean J. Spader, *Megatrends in Criminal Justice Theory*, 13 AM. J. CRIM. L. 157, 168 (1985) ("Factors dealing with the present, such as the defendant's social or economic status, repentance and desire to make restitution, or factors dealing with the future, such as the defendant's predicted dangerousness, rehabilitation, and recidivism, are not truly compatible with pure desert sentencing.").

65. *See infra* text following note 68.

66. The left-hand column of Table 1 ranks the crimes in the same order as Robinson and Kurzban's subjects did. However, contrary to Robinson and Kurzban's findings, *see* Robinson & Kurzban, *supra* note 40, at 1868-69 & tbl.1, in our study the average sentence for the microwave theft slightly exceeded the average sentences for both the assault at the store and the assault at the football game. The microwave theft scenario was also one of the two unmanipulated scenarios (the other involving assault at a store) that produced a different average sentence (slightly lower) in the experimental group compared to the control group.

predicted. Of note, these pairs involved the most serious crimes, with all four scenarios describing homicides.

TABLE 1

Average Sentences and Ordinal Ranking of Twelve Paired Crime Scenarios

(*italics* indicate reversed rankings between control and experimental groups)

	Control Group Ranking		Experimental Group Ranking	
Pairs	Avg. Sent. (rounded)	% Ranking Crime >[†]	Avg. Sent. (rounded)	% Ranking Crime >[†]
Pair 1:				
Theft of pie	1 day	6	1 day	41
Theft of T-shirt* (apology)	2 months	94	2 weeks	59
Pair 2:				
Theft of drill (prior offense)*	*1 year*	*25*	*1 year*	*81*
Theft of microwave	*1 year*	*75*	*6 months*	*19*
Pair 3:				
Slap at store	*6 months*	*25*	*2 months*	*63*
Head-butt (treatment)*	*6 months*	*75*	*2 months*	*47*
Pair 4:				
Assault at station (vow to repeat)*	*3 years*	*20*	*4 years*	*52*
Robbery and beating	*4 years*	*80*	*4 years*	*48*
Pair 5:				
Death by pitbull* (hide dogs)	10 years	22	10 years	33
Death of baby in car	12 years	78	12 years	67
Pair 6:				
Burning woman* (psychopath)	50 years	9	45 years	11
Rape and killing of child	Life	91	Life	89

* Manipulated scenario in experimental condition. Only the experimental group actually received the manipulated scenario.

[†] "% Ranking Crime >" means the percentage of survey participants who considered the indicated crime to be more blameworthy than the other crime in the pair.

These results show, consistent with other research in this vein,[67] that when people make decisions about relative punishment they are willing to consider

67. *See, e.g.*, Jeremy A. Blumenthal, *Perceptions of Crime: A Multidimensional Analysis with Implications for Law and Psychology*, 38 MCGEORGE L. REV. 629, 649 (2007)

more than desert. This is especially so when the crimes are minor or mid-level felonies. The finding that people are willing to consider relative risk, relative treatability and other "extralegal" factors undermines Robinson's thesis that desert, defined in terms of blameworthiness at the time of the crime, is the only or even the principal determinant of society's views on ordinal ranking of these sorts of crimes.[68]

Although we did not include enough scenarios to allow us to investigate the type of utilitarian factors that have the most impact on crime rankings, of interest is that the manipulated scenarios that occasioned the most dramatic difference between our results and Robinson and Kurzban's involved, in order of impact, a prior crime (56% change), a willingness to undergo treatment (38%), an apology and restitution (35%), and a vow to recidivate (32%). Thus, according to our results, both facts that increase risk and facts that reduce it have significant influence on punishment decisions. However, in the two pairs involving homicides, facts that appear to increase risk had very little impact, suggesting that desert is much more influential in this type of case.

While we found that most of the rankings changed significantly when utilitarian factors were added to the scenarios, we also found a fair amount of consensus within both the control and experimental groups, albeit not quite as much as Robinson and Kurzban did. Robinson and Kurzban obtained a 0.88 Kendall coefficient, indicating an extremely high level of agreement about the ordinal ranking of the twenty-four scenarios. For the twelve scenarios that we used in this study, we obtained a 0.836 coefficient for the control group and a 0.816 coefficient for the experimental group,[69] likewise showing that when people are given the same facts—whether relevant to desert, risk, or both—they tend to arrive at similar conclusions about relative punishments.

(reporting a study finding that "people may agree about how serious a crime may be, but not why it is so serious"); Dena M. Gromet & John M. Darley, *Restoration and Retribution: How Including Retributive Components Affects the Acceptability of Restorative Justice Procedures*, 19 SOC. JUST. RES. 395, 411, 423 (2006) (finding that subjects assigned reduced sentences to offenders who successfully completed restorative procedures but also finding that subjects insisted on some punitive prison for serious offenses). One review of public opinion surveys circa 2000 that looked at data about public attitudes toward both relative and absolute punishment concluded that "the central tendency in public opinion is to be punitive and progressive. . . . In short, do justice, protect society, and reform offenders." Francis T. Cullen et al., *Public Opinion About Punishment and Corrections*, 27 CRIME & JUST. 1, 60 (2000) (capitalization altered).

68. Robinson's own research verifies this point. *See* Robinson et al., *supra* note 64, at 825 (summarizing studies finding that "there is broad support for the use of some [extralegal punishment factors] in determining punishment"). It is not clear how he reconciles these results with his statement that people base their punishment decisions on desert.

69. Like Robinson and Kurzban, we calculated the Kendall coefficient, which ranges from 0, indicating no agreement, to 1, indicating complete agreement or concordance. We did not run statistical significance tests on the differences between our results and theirs because we did not have access to their raw data.

2. *Study 2: Is there consensus about the appropriate punishment for crimes?*

Methodology. While our subjects did tend to agree about ordinal ranking, consensus disappeared in both the control and experimental groups when the question shifted to the specific punishments for the scenarios. In Study 2 we directed the same subjects involved in Study 1 to assign punishment in each of the twelve scenarios they read, using a scale that allowed them to select specified punishments ranging from no liability to the death penalty. In all, the subjects were able to choose from thirty different punishment options on this scale.[70]

Results. On the assumption that, with this many options, disagreement would be fairly substantial, we decided that the coefficient analysis relied on in Study 1 was not the best way to measure the degree of consensus. Instead we used the simplest method of discerning variance within a sample population, the standard deviation. Generally speaking, the larger the standard deviation, the wider the dispersion—in this case, the wider the dispersion of punishment choices.[71] As the two "standard deviations" columns in Table 2 indicate, the lowest standard deviation for any of the twenty-four scenarios (the twelve read by the control group and the twelve read by the experimental group) was 2.903, and only two other standard deviations were below 3. Four of the standard deviations were above 7, six above 4, and the remaining eleven ranged from 3.048 to 3.949. Note that the eight scenarios involving homicides (four each for the control and experimental groups) were all associated with very high standard deviations (ranging from 3.7 to 7.2); thus, although Study 1 found consistency in the ordinal ranking of these crimes, disagreement within the control and experimental groups about specific punishments for these same crimes was remarkably high.

Two other measures of punishment agreement are reported in Table 2. The "range within two standard deviations" columns depict, for the control and experimental groups respectively, the range of punishments within two standard deviations. Because this range reflects the choices of only about 95% of the participants,[72] it is less affected by any "irrational" outliers. Note that even for the lesser crimes, this range is fairly significant (for instance, for the theft of the pie, from no liability to four years) and for the more serious crimes it is often

---

70. The punishment options were: not liable, liable but no punishment, one day, two weeks, two months, six months, one year through fifteen years in one-year increments, twenty through fifty years in five-year increments, life sentence, and death penalty.

71. 1 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY, app. B at 252 (2005) ("The [standard deviation] indicates how far a typical element deviates from the average.").

72. *See id.* at 204 n.4 ("[W]hen the distribution follows the normal curve, about 68% of the data will be within one standard deviation of the mean, and about 95% will be within two standard deviations of the mean."). Our distributions were normal.

huge (for instance, for the death of the baby in the car, from two weeks to the death penalty). Finally, the "% at Mode" columns indicate that, except for the two most serious and two least serious crimes, the percentage of subjects who chose the modal sentence fell between 9% and 25%, suggesting only a minimal central tendency in the results (although, as noted, the large number of choices reduces the likelihood that a high percentage of participants would pick the same sentence). The bottom line, not surprisingly, is that people vary widely in their assessments about specific punishments, whether they are focused on desert (as with the control group) or are given additional factors relating to dangerousness and treatability (as with the experimental group).

TABLE 2
Standard Deviations, Range Within Two Standard Deviations, and Modes on Punishments Assigned to Twenty-Four Scenarios

	Control Group			Experimental Group		
Scenarios	Stand. Dev.	Range w/in 2 Stand. Dev.*	% at Mode	Stand. Dev.	Range w/in 2 Stand. Dev.*	% at Mode
Theft of pie	3.258	0 to 4 years	50	3.048	0 to 3 years	51.7
Theft of T-shirt†	3.459	0 to 6 years	23.1	3.933	0 to 6 years	40.6
Theft of drill†	3.829	0 to 8 years	19.2	2.903	0 to 7 years	21.9
Theft of microwave	3.474	0 to 8 years	19.9	3.770	0 to 8 years	25.2
Slap at store	3.470	0 to 7 years	25	2.993	0 to 6 years	23.2
Head-butt at game†	3.109	0 to 6 years	18.8	2.921	0 to 5 years	18.3
Assault at station†	4.284	0 to 12 years	19.8	4.377	0 to 12 years	22.2
Robbery & beating	4.483	0 to 13 years	17.8	4.479	0 to 13 years	17.8
Pitbull death†	7.133	1 day to death	14.9	7.172	0 years to death	14.5
Death of baby	7.373	2 weeks to death	14.5	7.259	2 weeks to death	9.8
Death of woman†	4.405	13 years to death	55.7	5.524	10 years to death	43.9
Rape / death of child	3.749	15 years to death	66.3	3.949	15 years to death	67.8

* "0" means either no liability or liability but no punishment.

† Manipulated scenario in experimental condition.

This conclusion could derail empirical desert theory at an early stage. For empirical desert to work as Robinson imagines it, consensus is probably most important not with respect to ordinal rankings—the focus of Robinson and Kurzban's research—but with respect to specific punishments.[73] Robinson suggests that once the "endpoint" for the maximum possible punishment is decided—whether it be the death penalty, life without parole, or twenty years—the ordinal rankings he and others have discovered will establish the hierarchy that descends from there.[74] In other words, if the maximum punishment for any crime is set at the death penalty (as is true in over thirty states) or life without parole (as it is in every other state but one),[75] the crime or crimes viewed as most serious by the public would receive that penalty and the remaining crimes would be assigned lower penalties according to their rank order. Yet as signaled by the vociferous debates about the propriety of capital punishment, sentences of life without parole, and even twenty-year terms,[76] the endpoint can be the crucial factor in determining people's opinions about the criminal justice system. So it is probable that the divergence between lay views and the criminal justice system that is most likely to be noticed by and bother laypeople is the precise punishment meted out in a given case, not whether the punishment fits within some nuanced hierarchy of which they are unlikely to be aware.[77] The first type of divergence, our research shows, is likely to be quite significant.

B. *Hypothesis II: Compliance*

Given this significant range of disagreement about appropriate punishments, constructing a criminal justice system that reflects widespread societal consensus about those punishments will be very difficult, even with respect to the so-called "core crimes" that are the subject of Studies 1 and 2, much less

---

73. *See* Kolber, *supra* note 36, at 450 ("[I]t seems unlikely that a person will care very much if the penalty for grand theft is a little higher or lower than the penalty for a minor battery if penalties for all offenses are an order of magnitude higher or lower than where he thinks they should be.").

74. *See* Robinson & Darley, *supra* note 45, at 33-34.

75. *See Life Without Parole*, DEATH PENALTY INFO. CTR., http://www.deathpenaltyinfo.org/life-without-parole (last visited Dec. 19, 2012) (indicating 33 states with the death penalty, 16 states with a maximum sentence of life without parole, and Alaska with neither).

76. *See, e.g.*, ANDREW VON HIRSCH, DOING JUSTICE 139 (1976) (arguing for a five-year maximum sentence for all but the worst murders); Cullen et al., *supra* note 67, at 10, 15, 17-18, 20 (discussing widespread disagreement among the public over the death penalty and three strikes laws).

77. Robinson implicitly recognizes this point in the studies he conducted on the compliance hypothesis, which tested people's reaction to specific sentences, not discrepancies in ordinal rankings. *See infra* notes 97-107 and accompanying text.

non-core crimes such as pollution, tax evasion, and abortion.[78] According to Robinson's next hypothesis, this disconnect with societal views about punishment (which to him should be focused on desert) could have significant non-compliance effects. Deviation from desert, he states, "undermines the criminal justice system's moral credibility and thereby undermines its crime-control effectiveness" because it

> undermines [the system's] power of stigmatization, increases the chances of vigilantism, promotes resistance and subversion rather than the cooperation and acquiescence required by the criminal justice system, undermines compliance in borderline cases where the condemnatory nature of the offense may be ambiguous, and reduces the criminal justice system's influence in the public conversation by which societal norms are shaped.[79]

Robinson originally relied on the research of Tom Tyler to support these types of assertions, but doing so involved a bit of a stretch. In his seminal work, *Why People Obey the Law*, Tyler did find that people's willingness to comply with the law is related to the perceived legitimacy of the authorities promulgating the law, which in turn is based at least in part on the extent to which the law that the authorities promulgate is consistent with people's views about what the law should be.[80] But when Tyler stated that "[t]he most important normative influence on compliance with the law is the person's assessment that following the law accords with his or her sense of right and wrong" (a passage cited by Robinson),[81] he was referring principally to the person's sense of right and wrong with respect to the *particular* law in question.[82] Most other research in this area is similarly limited in its conclusions.[83]

Tyler's work, read closely, stands for a somewhat different proposition than Robinson suggests. As Tyler has stated in more recent writing, obedience

---

78. Robinson and Kurzban acknowledge, based on other studies they conducted, that "true disagreements do exist for intuitions about wrongdoing outside the core of physical aggression, unconsented-to takings, and deception or deceit in exchanges." Robinson & Kurzban, *supra* note 40, at 1880; *see also id.* at 1880-90 (reporting significant disagreement on ranking of marijuana possession, cocaine possession, prostitution, a third theft, abortion, and rape, among other crimes).

79. Robinson et al., *supra* note 46, at 1948.

80. TYLER, *supra* note 44, at 45-46, 64.

81. Robinson & Darley, *supra* note 45, at 25 (alteration in original) (quoting TYLER, *supra* note 44, at 64).

82. *See* TYLER, *supra* note 44, at 64.

83. Several of these studies are reported in Robinson et al., *supra* note 46, at 2011-16. *See, e.g.*, Herbert Jacob, *Deterrent Effects of Formal and Informal Sanctions*, 2 LAW & POL'Y Q. 61, 70 (1980) (finding that people who disagreed with speeding and marijuana possession laws are much more likely to violate those laws); Karyl A. Kinsey, *Deterrence and Alienation Effects of IRS Enforcement: An Analysis of Survey Data*, *in* WHY PEOPLE PAY TAXES: TAX COMPLIANCE AND ENFORCEMENT 259, 276, 282 (Joel Slemrod ed., 1992) (finding that news about acquaintances being treated unfairly by the IRS increased willingness to evade taxes); Matthew Silberman, *Toward a Theory of Criminal Deterrence*, 41 AM. SOC. REV. 442, 457 (1976) ("When public sentiment in general disapproves [of] a given offense, it is relatively unlikely to occur.").

to the law "is not linked to either the favorability or fairness of the decisions made or policies pursued by legal authorities. Rather, legitimacy is linked to the justice of the procedures by which the police and courts implement the law."[84] In other words, according to Tyler, overall compliance with the law may be undermined more by procedural injustice than by substantive injustice.[85] Indeed, even in his original work Tyler stated that "[p]eople generally feel that law breaking is morally wrong, and that they have a strong obligation to obey laws even if they disagree with them."[86]

Janice Nadler has conducted research that more directly supports the compliance hypothesis, which she reported in an article entitled *Flouting the Law*.[87] In one experiment, she gave two groups of subjects six newspaper stories. For both groups, three of the stories were unrelated to legal issues. But the remaining three stories differed between the two groups, by depicting legislation that either unfairly or fairly deprived citizens of property or infringed their rights.[88] She then asked the subjects to indicate on a "Likelihood of Criminal Behavior Questionnaire" the probability that they would commit eight different crimes, including parking violations, speeding, and illegally copying software.[89] She found that the group exposed to the unfair stories was significantly more likely to contemplate noncompliance with the law than the group exposed to the fair stories.[90] In a second study involving another set of subjects, she described two versions of an actual case: one in which the criminal justice system did not

84. TOM R. TYLER, WHY PEOPLE OBEY THE LAW 273 (2006) (new afterword to 2006 reprinting).

85. *See* Jaime L. Napier & Tom R. Tyler, *Does Moral Conviction Really Override Concerns About Procedural Justice? A Reexamination of the Value Protection Model*, 21 SOC. JUST. RES. 509, 509 (2008) (contesting the notion that procedural justice concerns are trumped by "moral mandates"). The third study reported by Robinson, Goodwin, and Reisig in *The Disutility of Injustice*, *supra* note 46, at 2016-23, also may have focused on the effect of attitudes about procedural, rather than substantive, justice. This study involved an analysis of data from a 2000 survey entitled *Public Opinion on the Courts in the United States*. The measure Robinson, Goodwin, and Reisig used to determine satisfaction with the criminal law was the participants' answer to the question "how well [do] you think the courts in your community handle" violent, drug, and juvenile cases? *See* DAVID B. ROTTMAN ET AL., PUBLIC OPINION ON THE COURTS IN THE UNITED STATES, 2000, app. A at 1-2, *available at* http://www.icpsr.umich.edu/cgi-bin/file?comp=none&study=3864&ds=1&file_id=892958 (subscription required). That type of question could trigger reactions based on procedural fairness as well as, or instead of, reactions based on case outcomes. Even if respondents chose to answer this question with only substantive outcomes in mind, however, the survey does not reveal the grounds on which substantive outcomes were judged—desert, utilitarian factors, or both.

86. TYLER, *supra* note 44, at 64.

87. Nadler, *supra* note 45, at 1399.

88. *Id.* at 1411-13.

89. *Id.* at 1413-14. In addition to the subjects who read the newspaper stories, Nadler also administered the questionnaire to a control group that did not read any of this material. *Id.* at 1411.

90. *Id.* at 1415-16.

impose punishment even though most people would agree that some punishment was warranted, and another in which the perpetrator was punished.[91] Nadler again found that the group exposed to the unjust hypothetical was significantly more likely to ignore or nullify unrelated criminal prohibitions than the group exposed to the just hypothetical.[92]

Nadler's studies suggest that dissatisfaction with particular cases, whether they involve excessive harshness or excessive lenience, can affect compliance with other laws. As she puts it, "specific instances of perceived injustice in the legal system can lead to diminished deference to the law generally."[93] But Nadler also reported a third study, one that controlled for gender, that came to a different conclusion; specifically, she found *diminished* deference among men who received a fair/just hypothetical compared to those who read an unfair/unjust hypothetical.[94] Nadler conjectured that this result might be due to a "moral credentialing" effect. In other words, people who perceive the law to be just might feel liberated to commit acts they would normally avoid.[95] Nadler ultimately concluded that "[i]t is undoubtedly false that perceived injustice in the legal system leads to greater willingness to break the law for all people, in all circumstances, at all times" and called for further research.[96]

In two studies reported in *The Disutility of Injustice*, Robinson, Geoffrey Goodwin, and Michael Reisig pursue that goal, using as prompts actual cases implementing punishment schemes that they conjecture are grossly disproportionate to lay views about desert, such as three strikes laws, strict liability crimes, and criminalization of regulatory violations.[97] In the first study, Robinson and his coauthors found that research subjects who were given scenarios depicting such schemes generally imposed a much lower sentence than the sentence actually meted out in the case described by the scenario.[98] More relevant to the compliance hypothesis, the first study also found that, after being told of the actual sentences that were imposed in these cases, the subjects indicated that they were significantly less willing to comply and cooperate with the law than they had been before they learned of the sentences. Specifically, Robinson et al. found that their subjects were less likely to endorse various punishments imposed by the criminal justice system, to report crimes to the authorities, or to

91. *Id.* at 1424.

92. *Id.* at 1417, 1423-24.

93. *Id.* at 1439.

94. *Id.* at 1420-21.

95. *Id.* at 1421 ("The theory of moral credentialing holds that people feel licensed to act on questionable motives when they have previously established their credentials as a person of pure motives."). In this study, male subjects who read the fair stories may have felt that they established their pure motives by agreeing with stories' content. *See id.* at 1421-22.

96. *Id.* at 1440.

97. *See* Robinson et al., *supra* note 46, at 1999-2000, app. C at 2031-32.

98. *Id.* at 2002, app. C at 2031-32.

correct injustices.[99] The latter attitudes were measured by eight queries that the authors called "Questions on the Effect of the Criminal Justice System's Moral Credibility" (which this Article will refer to as the Moral Credibility Scale) that were administered to the subjects both before and after they read the cases.[100]

Worried about the possibility that the subjects in this study changed their answers on the Moral Credibility Scale because they believed they were expected to do so after learning about the actual sentences imposed, the authors conducted a second study using a "between-subjects" design.[101] Consistent with the results of the first study, the Moral Credibility Scale responses from those exposed to the aberrant sentences indicated less willingness to endorse other criminal justice system dispositions and cooperate with authorities than the responses from a group that read cases corresponding more closely to empirically derived desert.[102]

These two studies in *The Disutility of Injustice* provide the most robust support to date for the compliance hypothesis. A strength of these studies was that the scenarios given to the experimental groups were based on actual cases.[103] But it should also be noted that the cases chosen all involved particularly draconian or abnormally lenient sentences. For instance, in one case given to the experimental groups, the court imposed a fifty-year sentence on a nineteen-year-old who reasonably but erroneously believed the girl with whom he had consensual sex was eighteen.[104] In another, the court imposed an eight-year sentence on an individual who imported seafood in plastic rather than cardboard containers, even though the individual did not know about the relevant law.[105] In a third, a serial rapist received no jail time because he was the son of a diplomat.[106] Further exacerbating the atypical nature of the prompts, the experimental groups were exposed to seven of these cases at once.[107]

---

99. *Id.* at 2002-03 & tbl.6.

100. For the content of the Moral Credibility Scale, see *id.* at 1999 tbl.5.

101. *Id.* at 2004-05. In a between-subjects design (to be distinguished from a "within-subjects" design), the manipulation is hidden from the subjects; its effect is studied by using two or more samples, ideally matched in all relevant respects, with each sample receiving a different independent variable (here, a group that reads scenarios depicting "just" sentences and a second group that reads scenarios depicting "unjust" sentences).

102. *Id.* at 2007 & tbl.7. It should also be noted, however, that while the authors found a statistically significant difference in scores between the control and experimental situations, that difference was not particularly large. The "baseline condition" (in response to "just" scenarios) for the eight Moral Credibility Scale answers averaged 6.22 on a 9-point Likert scale (with 9 indicating complete agreement with statements that grant maximum moral credibility to the law). *Id.* The two groups exposed to the seven real-world injustice scenarios averaged slightly above 5.2 and 5, respectively, which on the Scale's 9-point Likert spectrum was midway between "strongly agree" and "strongly disagree." *Id.* at 2003, 2007.

103. *See id.* at 1999.

104. *Id.* app. C at 2031.

105. *Id.*

106. *Id.* app. C at 2032.

107. *See id.* at 1999.

In Study 3, which we conducted before we were aware of the Robinson, Goodwin, and Reisig research, we employed a different methodology to test the compliance thesis. Subsequently, because the results in Study 3 raised intriguing questions and because we had in the interim learned of the studies reported in *The Disutility of Injustice*, we conducted a couple of follow-up studies, which we call Study 4. These latter studies utilized a between-subjects methodology similar to that used in the second study reported in *The Disutility of Injustice*, but relied on prompts that were less dramatic and that also injected utilitarian considerations into the calculus. Finally, in Study 5 we examined how long any noncompliance effects produced by these types of studies last.

1. *Study 3: Does dissatisfaction with the law promote noncompliance?*

Methodology. This study proceeded in two stages, using the same 530 subjects involved in Studies 1 and 2.[108] First, the participants received a list of seventeen crimes, ranging from murder to various misdemeanors and, for each, were asked what they thought the maximum sentence was in their jurisdiction *and* what they thought the maximum sentence should be. One purpose of these queries was to determine how well the participants knew the law in their jurisdiction, since an implicit assumption of empirical desert is that people are aware of the extent to which the law differs from their own preferences.[109] But for present purposes the more important aspect of this study was to obtain information about the subjects' satisfaction with the law in their jurisdiction. For some participants the answers to the "what is the law" question were identical or very similar to their answers to the "what should the law be" question. For others the divergence between these answers was quite marked with respect to many or most of the crimes.

To quantify the results of this first part of the study, we constructed a "Degree of Dissatisfaction with the Law" Scale (Dissatisfaction Scale) that measured, for each participant, the difference between what participants believed the law to be and what they thought it should be. The scale ranged from 0 (indicating no dissatisfaction with the law) to 11 (indicating a high degree of dissatisfaction with the law). According to empirical desert theory, those with a high score on the Dissatisfaction Scale should be more prone to noncompliance with

108. The subjects could not return to the prompts for Studies 1 and 2 after they accessed the prompts for Study 3.

109. *See* Slobogin, *supra* note 38, at 1193-94 (hypothesizing that "[m]ost departures from empirical desert will not be noticed by the public," in part because, as Robinson himself has noted in his criticism of general deterrence theory, "'what we know from studies is that even criminals commonly have no idea what the [criminal law's] rule is'" (emphasis in first quotation omitted) (alteration of second quotation in original) (quoting Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1093)). Without reporting results relevant to this matter in detail, we can indicate that our results show that very few people have accurate information about sentencing laws in their jurisdictions.

the law than those with a low score. While, given the mundane nature of the crimes involved in this study, even those who indicate a high degree of dissatisfaction on this scale may not have had as much antipathy toward criminal justice as those exposed to Nadler's and Robinson's more dramatic scenarios, the reaction measured by the scale probably better represents commonly held beliefs about how criminal justice works on an everyday basis. Thus, the latter beliefs may well be more likely determinants of the extent, if any, to which divergence between the criminal law and the public's views typically affects the public's compliance and cooperation with the criminal law.

The second stage of the study involved ascertaining our participants' relative willingness to comply with the law. Although we suspect that any survey measure of perceptions about the likelihood one will violate the law has minimal external validity, we decided that Nadler's Likelihood of Criminal Behavior Questionnaire would be an appropriate evaluation instrument because she had used it with some success in her own research.[110] We expected, contrary to the compliance hypothesis, that compliance rates measured by Nadler's scale would be similar regardless of whether the participant expressed great dissatisfaction with perceived punishments in his or her jurisdiction or none at all.

Results. Our expectations were not entirely borne out by the results of this study. Rather, as indicated in Table 3, we found a small but noticeable correlation between dissatisfaction with the law and noncompliance, a correlation that varied in an interesting way. In the group that in Studies 1 and 2 read Robinson and Kurzban's unchanged scenarios, there was a marginally significant correlation of 0.1 between dissatisfaction with the law and noncompliance.[111] In the group that in Studies 1 and 2 received the manipulated scenarios, the correlation was a statistically significant 0.22.[112]

Two points are worth emphasizing about these results. First, the correlation between dissatisfaction with the law and noncompliance was weak. Even a 0.2 correlation is, at best, suggestive of a relationship.[113] The second point is that,

---

110. *See* Nadler, *supra* note 45, at 1413-14. Nadler's questionnaire included eight crimes: drunk driving, parking in a no-parking zone, failing to pay required taxes, making illegal copies of software, eating a small item without paying in the grocery store, exceeding the posted speed limit, drinking alcohol under age twenty-one, and taking home office supplies for personal use. *Id.* Simply out of curiosity, we added burglary and unjustified homicide to the list. With respect to both of these crimes, willingness to comply was at or near 100% regardless of the conditions to which the subjects were exposed.

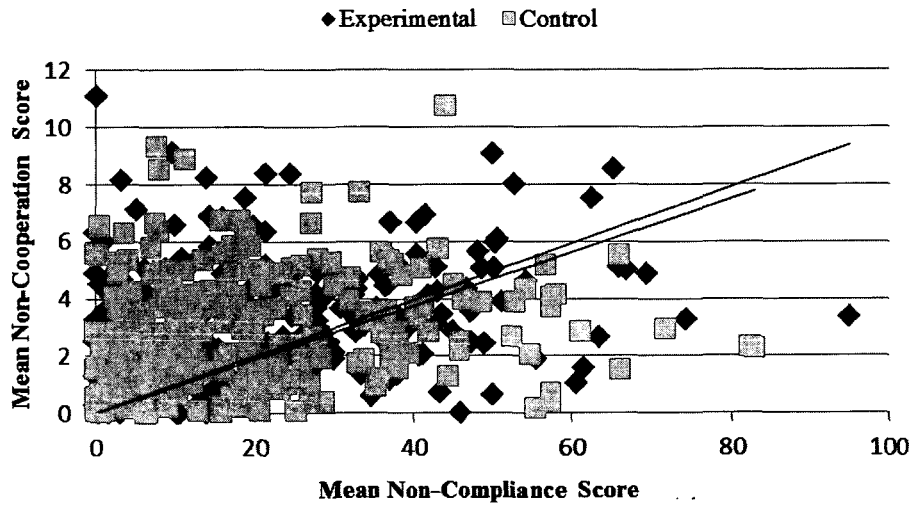
111. P-value ≤ 0.118. The correlation test used was a Pearson's R, which like most correlation measures produces a range of 0.00 (indicating no relationship between the measured variables) to 1.0 or -1.0 (indicating a very strong positive or negative relationship, respectively). *See* FAIGMAN ET AL., *supra* note 71, app. B at 245 (defining correlation coefficients).

112. P-value ≤ 0.001.

113. *See* FAIGMAN ET AL., *supra* note 71, at 223-24. Also worth noting is that the reliability of the correlation coefficient increases with its absolute value, so that a difference of 0.1 between two correlations is much less significant if the two coefficients are 0.15 and 0.25 than if they are 0.8 and 0.9. *See What Are Basic Statistics?: How to Determine Whether Two*

to the extent we did find a correlation between dissatisfaction with the law and noncompliance, it was much stronger with the group exposed to the manipulated scenarios. We were confused by this finding initially, since in this part of the study everyone was given identical prompts: both groups were asked to indicate the sentence for the seventeen listed crimes as well as what the sentence should be. But of course there was one difference between the two groups. The most noncompliant group had read, in connection with Studies 1 and 2, several scenarios providing them with information relevant to utilitarian goals such as treatability and risk. Perhaps these scenarios brought home the fact that the current sentencing regime (at least as perceived by the subjects) does not provide enough flexibility to take into account these types of factors, which in turn increased both dissatisfaction and noncompliance.

FIGURE 1
Correlation Between Dissatisfaction with the Law and Likelihood of Noncompliance



In short, Study 3 provides only minimal support for the compliance hypothesis. Furthermore, Study 3 suggests that, while failing to adhere to desert might cause some noncompliance, failing to adhere to utilitarian goals could cause even more noncompliance. Study 4, in its two variants, was designed to investigate the latter possibility further.

Correlation Coefficients Are Significant, STATSOFT, http://www.statsoft.com/textbook/basic-statistics/#Correlationsc (last visited Dec. 19, 2012).

2. *Studies 4A & B: What type of dissatisfaction promotes the most noncompliance?*

Methodology. The results of Study 3 at best mildly imply that punishment that departs from utilitarian concerns is somewhat more likely to cause noncompliance with the law than punishment that diverges from desert. In Study 4A we explored this possibility further by giving a new online sample of 236 people six scenarios from Studies 1 and 2,[114] specifically the six scenarios that were manipulated in those studies to include utilitarian considerations. However, instead of asking the subjects to rank or assign punishment as we did in the earlier studies, we provided them with the sentence that John received, one we made up but which we told the subjects was "based" on the results of "an actual case."

The sample was broken into three groups. To the first third of the sample (the control group) we gave the scenarios *without* the manipulation and told the subjects that John received a sentence identical to the average sentence assigned by the control group in Study 2 (the 264 subjects in that study who were given the unmanipulated scenarios). Given the standard deviations reported in the latter study,[115] this average sentence was only chosen by a small percentage of the subjects in that study, but it nonetheless represented a reasonably "just" sentence based on our data. To another third of the sample (the experimental-utilitarian group) we gave the subjects the scenarios with the utilitarian manipulation and told them that John received either a much harsher or much more lenient sentence than the sentence given the control group, in the direction opposite to what one would predict given the nature of the utilitarian-oriented information. For instance, in the T-shirt theft scenario where John expressed remorse and paid restitution, participants were told that John received a particularly harsh sentence (ten years as opposed to the average of two months assigned by the participants in Study 2), and in the scenario where he vowed to steal again participants were told he received a particularly lenient sentence (one year as opposed to the average of three years he received from the participants in Study 2). To the final third of the sample (the experimental-desert group) we gave the same scenarios that we gave the control group (that is, *without* the utilitarian information) but told the subjects that John received the same "unfair" sentence that we provided the experimental-utilitarian group (see Appendix B for further description).

---

114. For Studies 4A, 4B, 7A, and 7B, declining funding compelled us to use the online survey service Amazon Mechanical Turk instead of the Owen School eLab, see *supra* note 62. The Amazon service enables researchers to programmatically access an on-demand, diverse workforce that is paid per task. Amazon's service is less expensive and may be somewhat more likely to solicit inattentive participants, but on the whole has proven trustworthy. *See* Jennifer Jacquet, *The Pros & Cons of Amazon Mechanical Turk for Scientific Surveys*, SCI. AM. (July 7, 2011), http://blogs.scientificamerican.com/guilty-planet/2011/07/07/the-pros-cons-of-amazon-mechanical-turk-for-scientific-surveys.

115. *See supra* Table 2.

We then asked each of the three groups to answer questions relevant to compliance. In this Study we not only gave the subjects Nadler's Likelihood of Criminal Behavior Questionnaire (where a higher score means less compliance) but also the Moral Credibility Scale that Robinson and his coauthors used in *The Disutility of Injustice* (where a higher score means more compliance).[116] In this way, we hoped to get a sense not only of the subjects' willingness to comply with the law but also the subjects' willingness to endorse the legitimacy of the system and cooperate with authorities.

Given the results in Study 3, our hypothesis was that, because of their dissatisfaction with the sentences John received, the experimental groups would be more likely to indicate less compliance and cooperation than the control group, but only by a small margin. We further hypothesized that the experimental-utilitarian/desert group would have somewhat higher noncompliance/noncooperation scores than the experimental-desert group because in the former condition the subjects were confronted by sentences that were both strongly disproportionate to desert *and* strongly inconsistent with a prevention rationale.

Results and Further Research. As depicted in the two columns labeled Study 4A in Table 3, these hypotheses were not supported by our data. In fact, consistent with our original hypothesis in connection with Study 3, we found no statistically significant difference between any of the three groups on either the Likelihood of Criminal Behavior Scale (designated the "Composite Compliance Score") or the Moral Credibility Scale (designated the "Composite Cooperation Score").

On the theory that the reason we obtained this result was that our changes in sentence for the two experimental groups were not sufficiently aberrant to occasion any response, we ran the same study with a new sample of 289 people, except that we increased or decreased the sentences significantly in the experimental-utilitarian and experimental-desert conditions (for instance, John received a twenty-year sentence for the T-shirt theft).[117] Additionally, in this study, which we call Study 4B, we altered our methodology to conform more closely to the methodology followed in *The Disutility of Injustice* by asking the subjects to decide on a sentence for John before we provided his "actual"

116. *See* Robinson et al., *supra* note 46, at 1999. The Moral Credibility Scale uses a 9-point Likert scale (which we modified to 5 points) to gauge the extent to which subjects would: (1) believe that a person the criminal justice system sentences to life must deserve the punishment he received; (2) consider whether a law prohibiting posting false comments online criminalizes a condemnable act; (3) conclude that a sentence of twenty-eight to thirty-two years for a financial maneuver on taxes is morally condemnable; (4) report a person who has illegally taken an arrowhead from an important historical site; (5) take a handgun found in an alley to the police; (6) report a neighbor who has six dogs, three more than is legally allowed; (7) return to pay for a tank of gas they inadvertently did not pay for even though they are unlikely to get caught; and (8) return to pay a restaurant bill they inadvertently did not pay even though it is unlikely they would get caught.

117. See *infra* Appendix B for details on the manipulations.

sentence, in the hope that the harshness or leniency of the sentences we then described to the experimental groups would be more clearly brought home to them. Our hypothesis remained the same as in Study 4A, to wit, that the two experimental groups would exhibit less compliance and cooperation than the control group, and that the experimental-utilitarian group would exhibit less compliance and less cooperation than the group that received sentences disproportionate solely to desert.

TABLE 3
Relationship Between "Just" and "Unjust" Sentences and Noncompliance/Noncooperation
(no differences in values are significant at $p \leq 0.05$)

	Composite Compliance Score (on 100-pt. scale)		Composite Cooperation Score (on 5-pt. scale)	
Scenario Groupings	Study 4A	Study 4B	Study 4A	Study 4B
Control: No utilitarian factors and "just" sentences	20.83	21.19	2.62	3.21
Exp. Utilitarian: Utilitarian factors and "unjust" sentences	20.21	17.63*	2.59	3.38
Exp. Desert: No utilitarian factors and "unjust" sentences	21.4	19.5	2.39	3.27

* Significant at $p \leq 0.15$

Again, our results, reported in Table 3 as Study 4B, surprised us.[118] Contrary to Robinson et al.'s findings, none of the groups differed significantly on the Moral Credibility Scale, suggesting either that this scale is not sensitive enough in these types of surveys, or that there is simply no clear relationship between dissatisfaction with legal results and willingness to cooperate with the law or endorse its results. In looking at the average scores under the Likelihood of Criminal Behavior Scale we did find one instance of a relationship that approached significance, but not in the direction we expected. As the Composite Compliance Score for Study 4B indicates, the experimental-utilitarian group—

118. Particularly so since Study 4B was conducted two days after the jury returned a verdict of acquittal on the most serious charges in the Casey Anthony case, a decision that occasioned major outcry about the criminal justice system. *See* Amy Pavuk & Bianca Prieto, *Casey Anthony Not Guilty of Murder*, ORLANDO SENTINEL (July 5, 2011), http://articles.orlandosentinel.com/2011-07-05/news/os-casey-anthony-verdict-20110704. Under empirical desert theory, this event should have created a greater tendency toward noncompliance in our sample.

the group that received the most off-kilter sentences—was the *least* likely to register noncompliance effects (see cell marked with an asterisk). This result is contrary both to our hypothesis and to Robinson and Kurzban's.

Because this difference, with a p-value ≤ 0.15, did not meet the standard test for statistical significance and is in any event small, it could be spurious.[119] However, when we dug more deeply into the data we found a statistically significant relationship between compliance and the extent to which the sentences the subjects in this experimental group wanted to impose on John departed from the sentence John actually received.[120] In other words, the more bothered these subjects were over the sentence John received, the *less* likely they were to register a willingness to engage in criminal behavior.[121]

One possible explanation for this result could come from general deterrence theory. Perhaps people are less willing to commit the types of minor crimes listed in the Likelihood of Criminal Behavior Scale when confronted with wildly irrational sentencing practices. After all, if the government can impose a twenty-year sentence for stealing a T-shirt after the offender apologizes and pays for it (as occurred in one of the experimental-utilitarian scenarios in Study 4B), imagine what it could do in a speeding or illegal downloading case.

Of equally significant interest is the fact that we did not obtain even a marginally significant difference between the control group and the experimental-desert group, in terms either of compliance or cooperation and endorsement. This result is contrary to the results obtained by both Nadler and Robinson et al. Perhaps the prompts given to the experimental groups in Nadler's and Robinson et al.'s research were viewed as more unjust than even the enhanced changes in dispositions we used in Study 4B, and thus had a greater impact.[122] Or perhaps any anticompliance effect in the experimental groups was counteracted by fears of irrational punishment, while any procompliance effect in the control group—the group provided "rational" scenarios—was counteracted by

---

119. *See* FAIGMAN ET AL., *supra* note 71, at 212 (stating that the "most common" threshold for significance is 0.05, roughly meaning a 95% probability that the difference is not due to chance). However, higher p-values do not necessarily mean differences are nonexistent. *See* D.H. Kaye, *Is Proof of Statistical Significance Relevant?*, 61 WASH. L. REV. 1333, 1354 (1986).

120. The correlation was 0.28, $p \leq 0.023$. In no other group was this correlation significant.

121. At first glance, comparison of the composite cooperation scores in Studies 4A and 4B may seem to confirm this tendency because the participants in Study 4B, who received more "unjust" scenarios than the participants in Study 4A, had appreciably higher cooperation scores. However, since these two studies involved different samples, direct comparisons cannot be made. *See supra* note 50.

122. Nadler's excessive lenience scenario involved an individual who witnessed his friend molesting a young girl and did not report it, yet was not prosecuted or even arrested. Nadler, *supra* note 45, at 1417. Her excessive harshness scenarios were described in three newspaper stories reporting legislation that unfairly imposed taxes, permitted government confiscation of property, and infringed civil liberties. *Id.* at 1413. For a description of Robinson et al.'s scenarios, see *supra* text accompanying notes 97, 103-107.

the "moral credentialing" phenomenon described earlier, which Nadler hypothesizes might lead people to "feel licensed to act on questionable motives when they have previously established their credentials as a person of pure motives."[123]

Be that as it may, our main focus in this study was not to replicate the Nadler or Robinson et al. studies but to compare the effect on noncompliance of divergence with desert to the noncompliance effects of divergence from utilitarian considerations. The results summarized in Table 3 suggest that compliance and cooperation/endorsement effects are very similar regardless of whether the divergence is from desert or utilitarian goals. Taken together, Studies 3, 4A, and 4B do not support any particular position with respect to the compliance hypothesis; rather, they suggest that the relationship between compliance and satisfaction with the substance of the criminal law is complicated and difficult to predict, and that any relationship that does exist is not likely to be very strong.

3. *Study 5: How long does dissatisfaction with the law last?*

Methodology. Our next study raises a more fundamental concern about the compliance hypothesis to the extent that it suggests that noncompliance attitudes resulting from the failure of the criminal law to follow community views is more than fleeting. This study featured three samples from law school classes at Vanderbilt University Law School and the materials that Nadler used in her *Flouting the Law* studies.[124] Seventy-four students read her stories depicting legislation that unjustly deprived individuals of property, income, or privacy (the harsh treatment condition). A second group of seventy-two students read her story describing an individual who was not prosecuted for an act most would consider blameworthy (the lenient treatment condition). Finally, a third group of forty-seven students (the control group) did not read either type of story.

These three groups were then given Nadler's Likelihood of Criminal Behavior Questionnaire. However, rather than administer that questionnaire to everyone immediately after reading the stories, as Nadler did, we delayed giving the questionnaire to approximately half of all three groups (thirty-five who read the harsh-treatment stories, thirty-five who read the lenient-treatment story, and twenty-two who did not read the stories). The delay ranged from seven to ten days after the first two groups read the stories. Our hypothesis was that any umbrage about unfair government practices dissipates quickly.

---

123. Nadler, *supra* note 45, at 1421.

124. This sampling method was necessary because of the two-stage nature of this Study, something that cannot be easily carried out using online survey programs in which participants are anonymous.

TABLE 4

Relationship of Attitudes to Delayed Measurement of Compliance
(cells marked "a" are statistically different from cells marked "b," $p \leq 0.05$)

	Group Composite Noncompliance Score	
	Simultaneous Compliance Questionnaire	Delayed Compliance Questionnaire
Did not read stories	35.15 (N=25)[a]	35.5 (N=22)[a]
Read stories depicting harsh treatment	38.8 (N=39)[b]	36.6 (N=35)[a]
Read story depicting lenient treatment	37.9 (N=37)[b]	35.16 (N=35)[a]

Results. As indicated in Table 4, this hypothesis was sustained. Compared to the control group that did not read any stories, and consistent with Nadler's findings, the two groups that received the Likelihood of Criminal Behavior Scale immediately after reading the stories registered slight upticks (both statistically significant) in noncompliance scores.[125] But we found virtually identical noncompliance rates between the control group and the two halves of the experimental groups that completed the scale seven to ten days after reading the stories. These findings suggest that dissatisfaction with unjust legal rules and dispositions either does not last very long or does not have a long-lasting effect on willingness to flout the law.

Of course, Robinson's compliance hypothesis, stated fairly, is that reduced compliance with the law is a result of repeated or constant divergence between the criminal law and societal views, rather than the product of one-time expo-

125. Note that composite compliance scores reported in Table 4 are higher (by about fifteen points) than the composite compliance scores reported in Table 3, even though the same measure (Nadler's Likelihood of Criminal Behavior Scale) was used in both studies. There are at least three explanations for this difference. First, of course, the prompts in Studies 4A and 4B were different from those used in Study 5. *See supra* text accompanying notes 122-124. Second, the subjects in the first two studies, obtained through an online survey service, were approximately eight years older, on average, than the law students used in Study 5. *See infra* Appendix D. Social scientists have long known that age and risk averseness are related. *See generally* CHRISTOPHER SLOBOGIN & MARK R. FONDACARO, JUVENILES AT RISK: A PLEA FOR PREVENTIVE JUSTICE 19-35 (2011) (detailing developmental and neurological research indicating that juveniles are much more impulsive and risk-seeking than adults). A third difference may have to do with assumptions about anonymity. Although the online samples were guaranteed anonymity, some of the participants in Studies 4A and 4B may still have been reluctant to report willingness to commit crimes. The students in Study 5, in contrast, filled out their surveys in rooms where they could "see" that anonymity was guaranteed.

sure to one or more stories of injustice.[126] Further research will need to explore this possibility. As one of us has noted elsewhere, however, the fact that the criminal law provisions in many states consistently depart from widely held views about desert does not appear to be associated with greater noncompliance in those jurisdictions, "either because members of the public are not aware of the divergence . . . , do not care about it . . . , or . . . are more prone to obey than disobey the law even when they are disgruntled by it."[127]

C. *Hypothesis III: Crime Control*

Robinson not only argues that following desert can lead to more compliance and cooperation, but also that these positive effects could be stronger than those produced by a regime that focuses more directly on increasing compliance with the law through incapacitation, rehabilitation, and general deterrence. As noted earlier, he has stated that "strong arguments suggest greater utility in a distribution based on shared intuitions of justice than in a distribution based upon optimizing deterrence, rehabilitation, or incapacitation."[128] However, to date no research supports this view, and the research reported in Studies 3, 4A, and 4B suggests that a mixture of utilitarian and desert goals may be no worse and perhaps even better at promoting compliance and cooperation, at least as measured by Robinson's Moral Credibility Scale and Nadler's Likelihood of Criminal Behavior Scale.

More importantly, neither the research conducted by Robinson and Nadler nor the research reported up to this point in this Article gets at the core question raised by empirical desert theory: will a regime based on desert be better at crime control than a regime more directly attuned to prevention? Assuming correctional programs have any efficacy at reducing crime, the answer to this question is unlikely to be yes.

As an illustration of this point, consider the robust research concluding that Multi-Systemic Therapy (MST), a community-based treatment consisting of four to six months of intense intervention and subsequent monitoring, is significantly more effective at reducing repeat violence among juveniles than prison or other programs of longer duration.[129] If the sole goal of the criminal justice system were desert, juveniles who commit violent crime would presumably receive some incarceration, or at least some type of supervision, of far more than a half-year duration; indeed, researchers have found that people believe

126. *See* Robinson, *Doing Justice as Controlling Crime*, *supra* note 28, at 1106 ("[A] criminal justice system that *regularly* does injustice and/or fails to do justice is one that risks prompting resistance and subversion . . . ." (emphasis added)).

127. Slobogin, *supra* note 38, at 1198-99.

128. Robinson, *The Role of Moral Philosophers*, *supra* note 28, at 1836.

129. *See* Slobogin, *supra* note 38, at 1201-02 (citing SCOTT W. HENGGELER ET AL., MULTISYSTEMIC TREATMENT OF ANTISOCIAL BEHAVIOR IN CHILDREN AND ADOLESCENTS 252-54 (1998)).

adolescents who commit violent acts deserve a serious penalty.[130] But the research on MST noted above indicates that the latter disposition is unlikely to be as good as MST at reducing crime among the affected juveniles. Nor is a prolonged prison term likely to be better at minimizing crime by others, unless normally law-abiding people become so upset by the community-based treatment that they take the law into their own hands, or unless the fact that no prison is involved seriously undercuts general deterrence. The first development is improbable.[131] The second development is a greater possibility but still unlikely, for reasons Robinson himself has helped delineate.[132]

Conversely, punishment based on empirical desert may often result in dispositions that ignore the potential for recidivism reduction and even aggravate it. Consider some of the data reported in *The Disutility of Injustice.* Robinson and his coauthors found that, while the punishments imposed by their subjects were on the whole much lower than the punishments imposed in the actual cases that were used as prompts, they were still "quite punitive."[133] For example, subjects given scenarios involving theft of a microwave, an assault requiring two stitches, and a similar assault during an attempted theft (scenarios also used both by Robinson and Kurzban and by us in our first two studies)[134] assigned, on average, sentences of 2.3 years, 5 years, and 9.1 years, respectively.[135] These subjects also assigned a 17.7-year sentence to an accomplice to a felony murder, a 19.2-year sentence to a juvenile who accidentally shot a school teacher, and a 26.3-year sentence to a mentally ill mother who drowned her children, believing their deaths would save them from going to hell.[136] Even setting aside the probable criminogenic impact of incarceration,[137] confining

130. *See, e.g.*, ROBINSON & DARLEY, *supra* note 21, at 141 (reporting a study indicating that the average penalty assigned by lay participants to fourteen-year-olds who commit murder was over six years); Elizabeth S. Scott et al., *Public Attitudes About the Culpability and Punishment of Young Offenders*, 24 BEHAV. SCI. & L. 815, 823 (2006) (finding that participants assigned the same level of culpability to an offender whether they believed the offender was a fifteen-year-old or a twenty-year-old).

131. Research indicates that the public is not averse to community-based dispositions for juveniles when allowed to consider "effectiveness" rather than desert. *See, e.g.*, ELIZABETH S. SCOTT & LAURENCE STEINBERG, RETHINKING JUVENILE JUSTICE 281 (2008) ("Adult punishment and long incarceration are approved, for the most part, only as a means to protect the public from violent young criminals; . . . if other more lenient sanctions are effective, they are favored over incarceration.").

132. *See* Paul H. Robinson & John M. Darley, *Does Criminal Law Deter? A Behavioural Science Investigation*, 24 OXFORD J. LEGAL STUD. 173, 174, 204-05 (2004) (cataloguing reasons why differences in criminal dispositions have little general deterrent effect).

133. Robinson et al., *supra* note 46, at 1974.

134. See *infra* Appendix A for a description of the scenarios.

135. Robinson et al., *supra* note 46, at 1972 tbl.4.

136. *Id.*

137. *See* Martin H. Pritikin, *Is Prison Increasing Crime?*, 2008 WIS. L. REV. 1049, 1108.

such people for these extensive periods of time is unlikely to be as effective at reducing crime as more traditional utilitarian attempts at doing so.[138]

What examples like this suggest is that, from a strictly utilitarian perspective, empirical desert theory may have it backward. Rather than using desert as the linchpin of punishment and hoping that crime prevention will thereby occur, it would make more sense to prefer dispositions like MST that aim directly at preventing crime, unless they are so antithetical to desert that vigilantism or other negative effects will result. The final studies we conducted indirectly tested this proposition. More specifically, they sought to test the extent to which departing from desert is acceptable to laypeople when they are given utilitarian reasons for doing so.

1. *Study 6: To what extent are assessments of desert affected by preventive considerations?*

Methodology. In this study we gave the same 530 subjects involved in Studies 1 and 2 four scenarios depicting different offenders, all with criminal records.[139] As described in more detail in Appendix C, the four scenarios involved: an adult who has committed three separate crimes (grand theft auto, burglary, and theft of golf clubs), an adult who has just been convicted of his third drunk driving offense, a juvenile who committed three progressively more serious assaults, and an adult drug addict who committed three thefts, the last of which involved a minor assault. The control group of 264 subjects was given a description of the personal and offense history of each of these offenders, while the experimental group of 266 subjects was given, in addition, a description of innovative treatment programs like MST that were said to be effective at reducing recidivism. Specifically, the experimental group was told that, once sentenced, the adult theft recidivist would be required to participate in a one-year rehabilitative program shown to reduce offending by thieves,[140] the drunk driver would be subject to a six-month substance abuse treatment program that

138. Robinson has argued that, for at least some crimes, alternatives to prison can have as much "punitive 'bite'" as prison. Paul H. Robinson & John M. Darley, *The Role of Deterrence in the Formulation of Criminal Law Rules: At Its Worst When Doing Its Best*, 91 GEO. L.J. 949, 996 (2003). But the duration of these alternatives would have to be extreme in order to equate with the prison terms described in the text. *Cf.* Robert E. Harlow, John M. Darley & Paul H. Robinson, *The Severity of Intermediate Penal Sanctions: A Psychophysical Scaling Approach to Obtaining Community Perceptions*, 11 J. QUANTITATIVE CRIMINOLOGY 71, 86 (1995) (finding, for example, that survey participants view a six-month prison term as equivalent to eighteen months of intermediate sanctions, and that "[n]o intermediate sanctions were seen as equivalent to prison terms of 2 years or more").

139. The subjects were unable to return to the prompts for Studies 1, 2, or 3 after they accessed the prompts for this study.

140. Some research indicates that putting theft offenders on probation upon condition of undergoing some type of treatment reduces recidivism rates. *See, e.g.*, Ted Bartell & L. Thomas Winfree, Jr., *Recidivist Impacts of Differential Sentencing Practices for Burglary Offenders*, 15 CRIMINOLOGY 387, 394 (1977).

eliminates alcohol cravings,[141] the juvenile recidivist would be involved in a six-month MST program,[142] and the drug addict would undergo a one-year "drug court" program involving treatment and close supervision in the community.[143]

We asked participants in both groups to carry out two tasks. First, they were to assign punishments on the same thirty-point scale that we used in Study 2. Second, they were to indicate how much of the punishment they assigned could occur in the community. Our hypothesis was that, compared to the control group, the experimental group would be willing to reduce both the length of punishment and the amount of dispositional time spent in prison.

Results. Consistent with this hypothesis, the subjects given the treatment information were more likely to prescribe lower sentences in each of the four scenarios, as well as somewhat more likely to permit community dispositions in each. As displayed in Table 5, the sentence lengths imposed by the experimental group were 26% lower (for the juvenile) to 43% lower (for the drunk driver) than those that were imposed by the control group, and the portion of that sentence to be served in prison was from 20% to 43% shorter than for the control group. Combining these two variables, Table 5 also indicates the average prison time imposed by the two groups for each scenario. The incarceration difference ranged from 32% less prison time for the juvenile recidivist to 51% less time for the drunken driving recidivist. Furthermore, the experimental and control groups ranked the four offenders differently in terms of punishment length. The control group imposed the most punishment on the adult theft recidivist, followed by the drug addict, the juvenile offender, and the drunk driver. The experimental group also imposed the most punishment on the adult

141. Some substance abuse treatment programs appear to reduce recidivism rates. *See* William D. Bales et al., *Substance Abuse Treatment in Prison and Community Reentry: Breaking the Cycle of Drugs, Crime, Incarceration, and Recidivism?*, 3 GEO. J. ON POVERTY L. & POL'Y 383, 389 (2006) (stating, with several qualifications and notable exceptions, that "studies indicate that offenders who participate in prison-based substance abuse treatment programs experience more successful community reentry, as compared to offenders who do not participate in the programs while incarcerated").

142. For a description of MST, see *Blueprints for Violence Prevention Model Programs: Multisystemic Therapy (MST)*, CTR. FOR STUDY & PREVENTION OF VIOLENCE, http://www.colorado.edu/cspv/blueprints/modelprograms/MST.html (last visited Dec. 19, 2012) (stating that MST reduces long-term recidivism rates by 25% to 70% and reduces out-of-home placements by 47% to 64%, with the result that MST has been found to be "the most cost-effective of a wide range of intervention programs aimed at serious juvenile offenders").

143. For a description of drug court treatment programs, see Peggy Fulton Hora et al., *Therapeutic Jurisprudence and the Drug Treatment Court Movement: Revolutionizing the Criminal Justice System's Response to Drug Abuse and Crime in America*, 74 NOTRE DAME L. REV. 439, 502 (1999) ("[Drug courts] across the country have recorded substantial success in retaining participants in treatment programs, reducing recidivism rates, and saving criminal justice system resources.").

theft recidivist and the least on the drunk driver, but reversed the ranking of the juvenile and the addict.

TABLE 5
Sentences Assigned to Recidivists With and Without Treatment Information
(differences in total confinement within each scenario significant, $p \leq 0.05$)

Recidivists	Control Group			Experimental Group		
	Overall Sentence	% in Cmty.	Total Confinement	Overall Sentence	% in Cmty.	Total Prison
Recidivist Thief & Burglar	8.7	17.8	7.2	5.1	22.3	3.9
Recidivist Drunk Driver	6.3	22.3	4.9	3.6	31.8	2.4
Recidivist Violent Juvenile	6.5	16.6	5.4	4.8	23.4	3.7
Recidivist Addict Thief	6.7	15.3	5.7	4.4	26.8	3.2

The total-confinement differences, all of them statistically significant, suggest that information about a plausible treatment program both lowers and changes what people think offenders "deserve," and makes them more willing to contemplate a disposition other than confinement. In other words, consistent with the results of Study 1, people's views of desert are affected by preventive considerations. Fashioning a criminal justice system solely around desert would not be in accord with the fully informed views of the public.

The results of this Study also suggest, however, that laypeople are unwilling to abandon desert as a criterion for punishment. In each of the four cases even the experimental group, which was told that treatment interventions ranging from six months to a year could reduce the offender's risk, wanted to impose prison time beyond what was needed for the treatment. How strong is this desire to impose desert-based punishment? Our last study—carried out in three variations—begins to answer this question.

2. *Studies 7A, B & C: To what extent do laypeople prefer dispositions focused on prevention rather than desert?*

Methodology. In this study we used the four scenarios from Study 6, plus an additional two scenarios taken from one of Robinson's studies, the first in-

volving a city official who accepts a bribe and the second involving an employee who murders his boss after being fired (see Appendix C for a full description of the scenarios).[144] We gave these six scenarios to the same sample that we used in Study 4A, divided into two groups.[145] The first group (the desert group) received the scenario descriptions without treatment information. The second group (the treatment group) received the scenario descriptions with treatment information. For the four scenarios that were used in Study 6, this treatment information was identical to the information conveyed in that Study. In the first scenario borrowed from Robinson we stated that the city manager offered to return the bribe and give another $20,000 to the city, plus work in community outreach programs for the next six years. In the second borrowed Robinson scenario (involving murder) we stated that by the end of the employee's third year in prison he had experienced a religious conversion, renounced violence, and provided restitution to the family.

To get a different sense of how desert and prevention goals interact, we changed the focus of inquiry in this Study. Instead of asking the subjects to impose punishment on the thirty-point punishment scale that we used in Study 6, we gave the desert and treatment groups only two punishment options. As indicated in Appendix C, the first option was a "determinate sentence" involving a very narrow sentence range (two years), centered around the average punishment assigned by the control subjects in Study 6 for the first four scenarios and around the average punishment imposed by Robinson's subjects for the two scenarios we took from his research.[146] The second option was an "indeterminate sentence" involving a different range for each scenario, with the lowest range (for the juvenile offender) starting at six months of probation and ending at five years in prison and the highest range (for the murderer) starting at three years in prison and ending at fifty years. These ranges were selected somewhat arbitrarily, but were meant to reflect a flexible sentencing regime.

The goal of this Study was to determine the extent to which our subjects were willing to contemplate, under these various conditions, an indeterminate sentence that was not closely tied to desert. We hypothesized that the desert group would be likely to pick the determinate, desert-oriented option, since it did not receive treatment information and thus would prefer a punishment bounded by desert. We hypothesized that the treatment group would be more likely to pick the indeterminate option because the treatment information would push it toward choosing a more flexible punishment scheme that allowed utilitarian goals to operate.

Results and Further Research. Our results, reported in Table 6 as Study 7A, largely supported these two hypotheses. As hypothesized, the treatment

144. For the source of these scenarios, see Robinson et al., *supra* note 64, at 774-75.

145. The subjects could not return to the prompts for Study 4A after they accessed the prompts for this Study.

146. *See* Robinson et al., *supra* note 64, at 781 & tbl.4.

group was more likely to pick the indeterminate option and the desert group was more likely to pick the determinate option. Across the six cases, 61% of the treatment group picked the indeterminate option, while 52% of the desert group did so. Although this difference was not statistically significant in conventional terms, that was due to the fact that even the desert group picked the indeterminate option a substantial percentage of the time, with a majority of that group choosing the indeterminate option in three of the six cases and 49% of that group picking that option in two of the remaining three cases. In the sixth case, involving murder, a majority of both the desert and the treatment groups picked the determinate option, again demonstrating that people are less likely to veer from desert-based punishment in connection with serious crimes.

TABLE 6
Determinate Versus Indeterminate Sentencing
(indicating percentage preferring indeterminate sentences)

	Study 7A		Study 7B (Broader Ranges)	
	Desert Group	Treatment Group	Desert Group	Treatment Group
Recidivist Thief & Burglar	52	60	40	64
Recidivist Drunk Driver	54	52	32**	52**
Recidivist Violent Juvenile	62	64	61	57
Recidivist Addict Thief	49*	72*	45*	57*
Corrupt City Manager	49*	68*	44*	63*
Murderer	44	48	7*	20*
Average	52**	61**	38**	52**

* $p \leq 0.01$; ** $p \leq 0.2$

Although overall both groups chose the indeterminate option at least as often as the determinate one, we were concerned that the popularity of the indeterminate choice could merely mean that many people in the desert group preferred a desert-based sentence outside the determinate range we gave them and decided that the indeterminate option (which in this Study had a very low minimum but a maximum no greater than the determinate option) was the next best alternative. Put another way, disagreement with the determinate option in this Study could be explained by our finding in Study 2 that conclusions about punishment are widely dispersed over the punishment spectrum rather than a result of a real preference for the type of "indeterminate" option we gave our subjects. In an effort to test this possibility we gave the scenarios to another sample (the same sample we used for Study 4B)[147] but with determinate ranges that were

147. The subjects could not return to the prompts for Study 4B after they accessed the prompts for this Study.

somewhat broader (three years instead of two) and indeterminate ranges that were broader as well, on the theory that subjects in the desert condition were more likely to pick "determinate" sentences if the range was broader and the indeterminate range was not bounded by desert considerations.

This hypothesis was partially borne out, as the results reported as Study 7B in Table 6 indicate. While the treatment group was still likely to pick the indeterminate option in all cases except the one involving the homicide, the desert group in this Study was more likely than the desert group in Study 7A to pick the determinate option in every scenario. Even under the modified conditions, however, if we exclude the homicide case (where only 7% of the desert group and 20% of the treatment group chose the indeterminate option), anywhere from 32% to 61% of the desert group rejected the expanded determinate option and chose the indeterminate one, and 52% to 64% of the treatment group picked the indeterminate option.

Curious about how influential treatment information of the type we described in these scenarios might be, we conducted a third variant of Study 7, using the same six scenarios but this time employing a two-step process. A new 68-person sample—23 professional staff at Vanderbilt University Law School and 45 participants solicited "off the street"[148]—was first given the scenarios without the treatment information and asked to assign a punishment on a modified 15-point scale. The participants were next given the treatment information, but this time adding a statement that the treatment was either successful or not successful, and that correctional authorities thus believed either that the offender was unlikely to repeat his offenses or that he was likely to reoffend. The subjects were then asked again to assign a punishment (see Appendix C for a fuller description). We anticipated that, for the majority of the participants, the second-stage punishments would be appreciably lower or higher, depending upon the assertions by correctional authorities.

The results from this final version of Study 7, reported below in Table 7, confirmed this hypothesis, again except with respect to the case involving the homicide. The second column in Table 7 indicates for each of the six scenarios the sentence the subjects imposed at the first stage when no treatment information was provided. The data reported in the third, fourth, and fifth columns indicate that, in five of the six cases, a majority of the participants were willing to abandon the desert-based sentence they imposed during the first stage after they were provided with treatability and risk information. Even in the sixth case, involving the homicide, treatment information changed the views of 39% of the participants, who reduced the sentence by an average of 42%.

148. As indicated in Appendix D, this sample did not include any African Americans. If this lack of diversity biases the results in any direction, it is probably toward less leniency. *Cf.* PETER H. ROSSI & RICHARD A. BERK, JUST PUNISHMENTS: FEDERAL GUIDELINES AND PUBLIC VIEWS COMPARED 205 (1997) (concluding that African American ethnicity is correlated with somewhat more lenient sentence choices, although educational attainment is the strongest demographic correlate of sentencing attitudes).

TABLE 7
Changes in Sentence Resulting from Risk/Treatability Information

	Average Desert-Based Sentence	Disposition After Risk Information		
		% No Change	% Change (% Inc./ Dec.)	% Indeterminate Extension
Recidivist Juvenile (Treatment successful)	7.6	15	85 (−71)	N/A
Manager (Provides restitution and reforms)	5.3	19	81 (−44)	N/A
Recidivist Addict Thief (Treatment successful)	3.9	27	73 (−92)	N/A
Recidivist Thief (Treatment unsuccessful)	5.2	31	27 (+55)	42
Recidivist Drunk Driver (Treatment unsuccessful)	6.1	48	2 (+16)	50
Murderer (Finds religion and provides restitution)	Life	61*	39 (−42)	N/A

* Assigning life a value of sixty years

As the increasing percentages in the third column ("% No Change") shows, willingness to move away from a desert-based punishment decreased when treatment was not successful or when the crime was serious. The fourth column indicates the average amount that participants who were willing to abandon a desert-based disposition increased or decreased the sentence in light of the treatability/risk information. The final column shows that, in the two scenarios where treatment was not successful and correctional authorities considered the individual a high risk, many subjects opted for an indeterminate sentence based on decisions by experts; these choices are not reflected in the average increase to sentence noted in the fourth column because they could not be quantified.

These findings corroborate the finding in Study 6 that plausible treatment information can change views about punishment. Laypeople are willing to consider sentences that are not tied to a particular level of culpability, especially when they are given reasons for thinking that a different or flexible disposition would make sense and that the crime is not extremely serious. Where serious crimes such as murder are involved, desert appears to play a much more dominant role for a majority of people, although even in this context evidence that the offender is no longer a risk can, for a substantial proportion of individuals, have an effect on the sentence they are willing to impose.

III. IMPLICATIONS AND FUTURE DIRECTIONS

The empirical desert project's effort to assess lay views about desert and the effect of the criminal law's divergence from those views is important for several reasons. First, it provides thought-provoking data for criminal law policymakers.[149] Second, it calls into question the ability of lawmakers to represent the public's views.[150] Third, and most importantly for our project, it begins to assess the instrumental value of adhering to desert. Deontological retributivists are presumably uninterested in this kind of data. But to utilitarians this assessment is crucial, because it helps determine the extent to which desert can be ignored or downplayed in arriving at dispositional decisions.

Empirical desert theory dictates that utilitarians not only try to measure the deterrent, incapacitative, and rehabilitative impact of punishment but also try to gauge the extent to which law's allegiance to societal views about desert affect "crime control," defined broadly in terms of preventing crime, assuring cooperation with the authorities, and reinforcing societal norms. The research Robinson has helped conduct suggests a relationship between these desiderata and a criminal justice system that is based on empirically derived desert. In contrast, the studies we have reported here tend to undermine all three hypotheses underlying empirical desert theory.

The consensus hypothesis is that consensus exists with respect to the relative ranking of a sizeable subset of crimes such as homicide, assault, and theft and that this consensus is based on desert considerations. Our research confirms that consensus about the ranking of core crimes exists, but it also shows that utilitarian concerns can change that ranking in ways inconsistent with desert. Moreover, lay views about specific punishments for crime—even for core crimes—can be multimodal rather than clustered around a mean, indicating significant disagreement about punishment preferences. This disagreement is most likely to occur in connection with serious crimes, ironically the only type of crime whose punishment is predominately driven by desert considerations. Thus, the consensus that is arguably most important to empirical desert theory does not exist.

The compliance hypothesis is that punishments that fail to adhere to desert can undermine the moral credibility of the law, and thus reduce compliance and cooperation with it. One of our studies suggests, to the contrary, that noncom-

149. *See* Christopher Slobogin, *Is Justice Just Us? Using Social Science to Inform Substantive Criminal Law*, 87 J. CRIM. L. & CRIMINOLOGY 315, 325 (1996) (stating that Robinson and Darley's research in *Justice, Liability, and Blame: Community Views and the Criminal Law*, *supra* note 21, "certainly could, and probably will, make theorists, especially adherents of the [Model Penal Code], rethink their positions").

150. *See* Berman, *supra* note 34, at 1118 ("[I]f modern American democracy prompts voters and their representatives to embrace laws that produce and perpetuate obvious injustices, then it would seem our foundational political system has a fundamental flaw that runs much deeper than any problems that could flow from poorly designed criminal laws.").

pliance effects are sometimes stronger the closer punishment *conforms* to desert-based results, perhaps because people are less worried that commission of trivial crimes will result in irrational sentences. Our other studies on this topic suggest that, to the extent that noncompliance effects result from criminal justice outcomes that diverge from societal views, they are as likely when the divergence is from utilitarian-oriented preferences as from desert-oriented preferences and, in any event, will dissipate quickly over time.

Finally, the crime control hypothesis is that the crime control benefits of a punishment system based on empirically derived desert are likely to be as great as or greater than the compliance generated by adhering to dispositions focused more directly on prevention. Our research does not directly test this hypothesis, which would be hard to do given the difficulty of measuring the extent of crime control and its causes. However, the studies testing the compliance hypothesis (Studies 3, 4A, 4B, and 5) suggest that the crime control hypothesis is false to the extent that "crime control" is framed solely as the product of compliance that stems from following lay views about desert-based punishment. Furthermore, Studies 6, 7A, 7B, and 7C suggest that, for all but the most serious crimes, most people are willing to change their assessments of punishment from what empirical desert would dictate if they think that preventive goals can be achieved through different means. Laypeople appear to believe that utilitarian considerations are at least as important as desert factors in fashioning sentences in the typical criminal case.

Survey data of the type reported here are always subject to criticism on external validity grounds. In particular, studies testing the compliance hypothesis (Studies 3, 4A, 4B, and 5) are suspect. People's assertions about their willingness to commit crimes or their unwillingness to cooperate with the authorities may bear little relationship to whether they will actually commit crimes or fail to report crimes.[151] Similarly, the punishments assigned by our subjects in Studies 1, 2, 6, 7A, 7B, and 7C might differ from the punishments the subjects would impose if they were confronted with the full context of the cases described in our brief scenarios.[152] The primary response we have to these generalizability concerns is that the research with which we compare our results—specifically, the research carried out by Robinson and his colleagues and by Nadler—suffers from the same problems.

The internal validity of these studies can also be criticized on a number of grounds. We have pointed out many of these concerns throughout this Article. Future research can try to better isolate the effects of utilitarian and desert considerations on attitudes toward particular crimes and punishments, the

151. *See* Nadler, *supra* note 45, at 1416 & n.75 (stating "we cannot definitively predict behavior from [self-reports]" and further noting the questionable ethics of research that motivates illegal behavior (emphasis omitted)).

152. Research confirms that support for prison terms for all but the most serious crimes decreases markedly when individuals are provided contextual information about the costs and effectiveness of intermediate sanctions. Cullen et al., *supra* note 67, at 43-44.

relationship between these attitudes and fealty to the criminal justice system, and the extent to which people are willing to abandon desert in service of preventive goals.

With these caveats, our research suggests that, if crime prevention is the objective, adherence to empirically derived desert is not likely to be the best way of achieving it. Even the less concrete goal of bolstering the moral credibility of the criminal justice system might be better pursued by following a mixture of desert and utilitarian goals. Of course, even if further research confirms that following empirically derived desert is not a superior crime prevention mechanism or the best way of shoring up the moral credibility of the criminal justice system, policymakers could still choose to make desert the lodestar of the criminal law on deontological grounds. The purpose of this Article is solely to evaluate the utility of desert, not its normative validity.

Furthermore, in assessing the relative merits of a desert- or prevention-based system, another type of utility, one that this Article has not evaluated, must also be considered. This second type of utility might be called the political economy of criminal justice. The choice between retributive and utilitarian goals in structuring the criminal justice system can not only have an impact on crime control and legitimization, but can also be driven by budgetary considerations, concerns about implementation, or the extent to which the choice can insulate the criminal justice system from the vagaries of the political process.[153]

In *The Disutility of Injustice*, Robinson and his colleagues allude to this political economy concern. They describe a number of egregious "crime control" stories of the type alluded to earlier—life sentences for three minor offenses, extremely long sentences for possession of small amounts of drugs, tough sentences for strict liability crimes—and suggest that these punishments are not the product of reasoned judgments by the public and legislators, but rather the result of media and government distortions of the facts, unreasonable public fear of crime, and the incentives elections create for legislators to be "tough on crime."[154] We tend to agree with this assessment. But by calling these "crime control" stories, Robinson et al. also imply that a preventive approach to criminal justice plays into the hands of irrational political forces and that a system based on empirical (or deontological) desert would avoid these travesties.[155] In

153. *Cf.* Richard S. Frase, *Theories of Proportionality and Desert*, *in* THE OXFORD HANDBOOK OF SENTENCING AND CORRECTIONS 131, 135 (Joan Petersilia & Kevin Reitz eds., 2012) ("[S]pending scarce resources to bring additional charges or impose additional punishments solely because they are deserved cannot be justified given other pressing needs for these resources within and outside of the criminal justice system.").

154. Robinson et al., *supra* note 46, at 1983-94.

155. *See id.* at 1979-80 (suggesting, in the course of discussing why criminal justice provisions "are in conflict with the community's shared intuitions of justice," that "while there may be some latent crime-control concerns present in the public consciousness, . . . in

other words, the innuendo of *The Disutility of Injustice* is that, while politics may make any sensible system of criminal justice difficult to maintain, desert is more likely than other considerations to curb the worst impulses of the democratic process.[156]

This is not the place to evaluate this controversy at length. But it should at least be pointed out that desert can also be misused and abused by a dysfunctional political system. Because desert is based on backward-looking assessments of the crime and the offender's mental state, it tends to be implemented by legislators and prosecutors. Because prevention—at least at the individual level—is based on forward-looking assessments of treatability and risk, it tends to be implemented by judges and parole authorities. The first set of decisionmakers is at least as susceptible as the second to the social and political pressures of the type described in *The Disutility of Injustice*.[157]

In the end, the best way to reconcile retributive and preventive goals is probably through some sort of limiting retributivism, or what we are calling preventive justice, which allows utilitarian considerations to have significant impact within a range established by retributive principles.[158] But much rides on the breadth of that range and the rationale for its endpoints. Desert theorists probably would not be happy with the broad ranges favored by many of our subjects in the various versions of Study 7. Yet not only are such ranges apparently popular, they probably offer the best method of preventing crime.[159] In

---

fact, the public may have been merely 'riding the wave' of concern actuated by a politician's previous comments").

156. *See also* Paul H. Robinson et al., *Realism, Punishment, and Reform*, 77 U. CHI. L. REV. 1611, 1630 (2010) ("[O]ur program challenges the dominant theory of crime control in the United States for the past several decades, one based upon intentionally and regularly doing injustice in the name of general deterrence and incapacitation by its reliance upon doctrines like three strikes, high penalties for drug offenses, adult prosecution of juveniles, abolition or narrowing of the insanity defense, the felony murder rule, and the use of strict liability." (footnote omitted)).

157. *See generally* William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505 (2001) (focusing on the pathologies of legislatures and prosecutors); James Q. Whitman, *A Plea Against Retributivism*, 7 BUFF. CRIM. L. REV. 85, 88 (2003) (asking why "the renaissance of neo-retributivism [has] also been the age of epochally harsh punishment").

158. *See supra* text accompanying notes 16-20.

159. *See* Joan Petersilia, *California's Correctional Paradox of Excess and Deprivation*, 37 CRIME & JUST. 207, 252-53 (2008) ("[Under California's determinate sentencing regime, a] large percentage of Californians who are nonviolent criminals are accumulating very extensive criminal records . . . [and yet] may not be any more dangerous than offenders in other states who are left 'on the street' and successfully handled through an array of community-based intermediate sanctions. [At the same time], California's sentencing system also releases violent offenders who amass lengthy criminal records—individuals who, in a system more carefully tailored to protect public safety, probably should not have been released in the first place."); Richard Rosenfeld et al., *The Contribution of Ex-Prisoners to Crime Rates*, *in* PRISONER REENTRY AND CRIME IN AMERICA 80, 102-03 (Jeremy Travis & Christy Visher eds., 2005) (concluding that "expanded use of discretionary parole supervision" would better protect the public from the safety threat posed by released prisoners); Joanna Shepherd,

short, unless it represents a radical departure from desert, which is most likely to occur in cases involving serious crime, indeterminate punishment focused on prevention is neither likely to cause more noncompliance than it prevents (through the mechanisms of specific deterrence, incapacitation, and rehabilitation), nor likely to undermine the legitimacy of the system.

Furthermore, as one of us has argued at length, indeterminate sentencing within very broad ranges has many other advantages over determinate sentencing.[160] Even in domains normally thought to be the province of desert-based systems—including accuracy, protection of offenders' liberty interests, respect for victims, and saving money—a prevention-oriented regime may well be superior to, or at least no worse than, a system that relies on narrower ranges.[161] In other words, contrary to Robinson and Kurzban's claim, "justice" is not solely the province of desert.

CONCLUSION

There may be good reasons to base criminal sentences on assessments of societal views about deserved punishment. But the research reported in this Article suggests that optimizing crime control is not one of them. Previous studies of empirical desert have greatly advanced discourse on this issue by providing sophisticated empirical descriptions of how laypeople think about punishment. But the methodology of these studies has tended to focus subjects on desert, when in fact people's views about punishment are much more multifaceted and complicated. No single metric can capture the opinions of the public, and even within the metric of desert their views vary widely. Thus, a single-minded focus on adherence to the public's assessments of desert is probably not the best way of shoring up the moral credibility of the criminal justice system or of promoting compliance and cooperation with it.

Fortunately, the research reported in this Article suggests that a failure to track community members' views on punishment does not have a significant or

---

Blakely*'s Silver Lining: Sentencing Guidelines, Judicial Discretion, and Crime*, 58 HASTINGS L.J. 533, 574 (2007) ("The more mandatory are the guidelines, the larger is the increase in crime."); *cf.* Yan Zhang et al., *Indeterminate and Determinate Sentencing Models: A State-Specific Analysis of Their Effects on Recidivism*, CRIME & DELINQ. 1, 18 (Dec. 8, 2009), http://cad.sagepub.com/content/early/2009/12/08/0011128709354047.full.pdf+html (finding, in a multistate, pre/post study, that determinate sentencing was worse than discretionary parole release at reducing recidivism in New York and North Carolina, better than discretionary release at reducing reoffending in Maryland and Virginia, no better or worse as a recidivism-reducing mechanism in Texas and Oregon, and that much depends upon the specific back-end programs in place).

160. *See* Christopher Slobogin, *Prevention as the Primary Goal of Sentencing: The Modern Case for Indeterminate Dispositions in Criminal Cases*, 48 SAN DIEGO L. REV. 1127 (2011).

161. *See id.* at 1154-68; *see also* Frase, *supra* note 153, at 144-45 (theorizing that the persistence of indeterminate sentencing in part "may be due to the widespread support for a hybrid approach").

lasting impact on their willingness to be law-abiding citizens. With that in mind, if crime control is the objective, the criminal justice system might well be better off pursuing utilitarian goals directly rather than relying on desert as its linchpin. Empirical desert would still be relevant to fashioning punishment, but only to the extent it helps set the boundaries beyond which real de-legitimization of the government and its laws occurs.

APPENDIX A: SCENARIO PAIRS USED IN STUDIES 1, 2 & 3

The underlined portions in scenarios 2, 3, 6, 7, 9, and 11 are the utilitarian manipulations.

PAIR 1

The owner has posted rules at his all-you-can-eat buffet that expressly prohibit taking food away; patrons can only take what they eat at the buffet. The owner has set the price of the buffet accordingly. John purchases dinner at the buffet, but when he leaves he takes with him two whole pies to give to a friend.

John notices in a small family-owned music store a T-shirt with the logo of his favorite band. While the store clerk is preoccupied with inventory, John places the fifteen-dollar T-shirt in his coat and walks out, with no intention of paying for it. After John is caught, he becomes remorseful. He pays the owner for the T-shirt and apologizes.

PAIR 2

John does not have all the tools he needs for his workshop but knows of a family two streets over who sometimes leaves unlocked the door to the detached garage next to their house. When he next sees his chance, he enters the detached garage through the unlocked door and takes a medium-size electric drill, intending to keep it forever. John has previously served time for one other, similar theft.

While a family is on vacation, John jimmies the back door to their house and steps into their kitchen. On the counter, he sees their microwave, which he carries away.

PAIR 3

A record store patron is wearing a cap that mocks John's favorite band. John follows him from the store, confronts him, and then slaps him in the face hard, causing him to stumble. The man's face develops a harsh black and yellow bruise that does not go away for some time.

While attending a football game, John becomes angry as he overhears an opposing fan's disparaging remarks about John's team. At the end of the game, John sticks his face in the man's face and head-butts him, causing a black eye and a gash that requires two stitches to close. John goes home, realizes he acted rashly, and signs up for an anger-management therapy program that he successfully completes.

PAIR 4

John demands money from a man buying gas at a gas station. When the man refuses, John punches the man several times in the face, breaking his jaw and causing several cuts that each require stitches. He then runs off without getting any money. He vows to his friends that he will try again.

To force a man to give up his wallet during a robbery attempt, John beats the man with a club until he relinquishes his wallet, which contains $350. The man must be hospitalized for two days.

PAIR 5

Two vicious pitbulls that John keeps for illegal dogfighting have just learned to escape and have attacked a person who came to John's house. The police tell John that he must destroy the dogs, which he agrees to do but does not intend to do. The next day, the dogs escape again and maul to death a man delivering a package. John hides the dogs from the police and eventually moves to another state where he can continue to raise pitbulls.

John is driving to see a man about buying an illegal gun but must babysit his friend's toddler son. It occurs to him that it is too hot to safely leave the toddler in the car but he decides to leave him anyway and to return soon. He gets talking with the seller, however, and forgets about the toddler, who passes out and dies.

PAIR 6

John works out a plan to kill his sixty-year-old invalid mother for the inheritance. He drags her to her bed, puts her in, and lights her oxygen mask with a cigarette, hoping to make it look like an accident. The elderly woman screams as her clothes catch fire and she burns to death. John just watches her burn. Doctors consider him a psychopath.

John kidnaps an eight-year-old girl for ransom, rapes her, and then records the child's screams as he burns her with a cigarette lighter, sending the recording to her parents to induce them to pay his ransom demand. Even though they pay as directed, John strangles the child to death to avoid leaving a witness.

APPENDIX B: SCENARIOS FOR STUDIES 4A & 4B

The underlined portion is the utilitarian manipulation used in Studies 1, 2, and 3, and continued in these studies. The unbracketed sentence term is the average sentence imposed by the control group in Study 1 and given to the control group in Study 4A and the second experimental group in Study 4B. The first bracketed sentence term was given to the experimental groups in Study 4A. The second bracketed sentence term was given to the experimental groups in Study 4B.

John notices a T-shirt with the logo of his favorite band in a small family-owned music store. While the store clerk is preoccupied with inventory, John places the fifteen-dollar T-shirt in his coat and walks out, with no intention of paying for it. After John is caught, he becomes remorseful. He pays the owner for the T-shirt and apologizes. He is given a two-month [ten-year] [twenty-year] sentence.

John does not have all the tools he needs for his workshop but knows of a family two streets over who sometimes leave unlocked the door to the detached garage next to their house. When he next sees his chance, he enters the detached garage through the unlocked door and takes a medium-size electric drill, intending to keep it. John has previously served time for one other, similar theft. He is given a two-year [six-month] [one-day] sentence.

While attending a football game, John becomes angry as he overhears an opposing fan's disparaging remarks about John's team. At the end of the game, John sticks his face in the man's face and head-butts him, causing a black eye and a gash that requires two stitches to close. John goes home, realizes he acted rashly, and signs up for an anger-management therapy program that he successfully completes. John receives a six-month [fifteen-year] [thirty-year] sentence.

John demands money from a man buying gas at a gas station. When the man refuses, John punches the man several times in the face, breaking his jaw and causing several cuts that each require stitches. He then runs off without getting any money. He vows to his friends that he will try again. John receives a three-year [one-year] [one-month] sentence.

Two vicious pitbulls that John keeps for illegal dogfighting have just learned to escape and have attacked a person who came to John's house. The police tell John he must destroy the dogs, which he agrees to do but does not intend to do. The next day, the dogs escape again and maul to death a man delivering a package. John hides the dogs from the police and eventually moves to another state where he can continue to raise pitbulls. He receives a ten-year [five-year] [one-year] sentence.

John works out a plan to kill his sixty-year-old invalid mother for the inheritance. He drags her to her bed, puts her in, and lights her oxygen mask with a cigarette, hoping to make it look like an accident. The elderly woman screams as her clothes catch fire and she burns to death. John just watches her burn. Doctors consider him a psychopath. He receives a fifty-year [ten-year] [two-year] sentence.

APPENDIX C: SCENARIOS FOR STUDIES 6, 7A, 7B & 7C

George Smith has served sentences for grand theft auto and burglary. Ten months after completing his second sentence, Smith stole three golf clubs, priced at $399 apiece. Smith was convicted of grand theft, his third felony.

Study 6 [Experimental group manipulation]:

Once sentenced, Smith will be required to participate in a newly developed one-year rehabilitative program that has been shown to reduce offending by burglars and other types of thieves. Indicate the sentence you would impose and the percentage, if any, of that sentence that could take place in the community under close supervision by probation officers. [Thirty dispositional options available.][162]

Study 7A [Both control and experimental groups given these options]:

_____ 1 to 7 years in prison (depending on how well Smith responds to treatment)
_____ 6 to 7 years in prison (regardless of how well Smith responds to treatment)

Study 7B [Both control and experimental groups given these options]:

_____ 1 to 10 years in prison (depending on how well Smith responds to treatment)
_____ 6 to 9 years in prison (regardless of how well Smith responds to treatment)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

Smith does not complete the program, and at the end of the prison term that you elected, correctional experts conclude that he will continue to steal if released. Should Smith be released at the end of the term you imposed above? Or would you extend Smith's term and, if so, by how much? Please indicate your answer below.

Release at end of prison term indicated above _____
Detain until experts indicate no longer a threat _____
Extend term by the following number of years _____

162. *See supra* note 70 for a description of the options. The same dispositional options were used for all of the Study 6 scenarios.

In most states, driving under the influence (DUI) of alcohol or drugs is a crime. Darren has been convicted of his third DUI offense, based on overconsumption of alcohol.

Study 6 [Experimental group manipulation]:

Once sentenced, he will be required to undergo an innovative six-month substance-abuse treatment program that eliminates alcohol cravings. Indicate the sentence you would impose and the percentage, if any, of that sentence that could take place in the community under close supervision by probation officers. [Thirty dispositional options available.]

Study 7A [Both control and experimental groups given these options]:

_____ 6 months to 5 years in prison (depending on how well Darren responds to treatment)
_____ 4 to 5 years in prison (regardless of how well Darren responds to treatment)

Study 7B [Both control and experimental groups given these options]:

_____ 6 months to 5 years in prison (depending on how well Darren responds to treatment)
_____ 3 to 5 years in prison (regardless of how well Darren responds to treatment)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

Darren fails to complete the program, and his treatment team concludes that Darren will continue to drink and will probably drive while drunk again. Should Darren be released at the end of the term you imposed above? Or would you extend the term and, if so, until experts indicate he is no longer a threat, or by a certain number of years? Please indicate your answer below.

Release at end of prison term indicated above _____
Detain until experts indicate no longer a threat _____
Extend term by the following number of years _____

When Sam Jones was age sixteen, he and three other school-age youths attempted to rob a restaurant, causing a head injury requiring stitches to the manager. Jones served a twelve-month sentence, and was then released. Less than six months later, when he was seventeen, Jones participated in a home invasion robbery. He and two others forcibly entered the home of the victim, held a pistol to his chest, and ransacked the home for money, but found only a few hundred dollars. These are the only two crimes Jones has committed.

Study 6 [Experimental group manipulation]:

Once sentenced, Jones will be required to undergo an innovative four-month rehabilitation program that involves family and peers and that has been shown to reduce violent offending. Indicate the sentence you would impose and the percentage, if any, of that sentence that could take place in the community under close supervision by probation officers. [Thirty dispositional options available.]

Study 7A [Both control and experimental groups given these options]:

_____ 6 months in a community-treatment program to 5 years in prison (depending on how well Jones responds to treatment)
_____ 4 to 5 years in prison (regardless of how well Jones responds to treatment)

Study 7B [Both control and experimental groups given these options]:

_____ 6 months in the community-treatment program to 10 years in prison (depending on how well Jones responds to treatment)
_____ 5 to 7 years in prison (regardless of how well Jones responds to treatment)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

Jones successfully completes the program and correctional experts conclude that he no longer poses a risk to the community. Should he be released now (after six months in the system)? Should he serve out the remainder of the term you imposed above? Or should the sentence be something different? Please indicate your answer below.

Release Jones now _____
Detain Jones until term indicated above is completed _____
Change sentence to the following _____[163]

163. In the actual survey, subjects who chose to change the sentence were given several lines on which to provide their answer.

Donovan is a drug addict. On three occasions he has been arrested for grabbing a woman's purse. He uses the money to obtain drugs. After each previous conviction he received a short period of imprisonment and then returned to the street. After release from his third jail term he once again steals a woman's purse, this time knocking her down. The woman requires three stitches to her forehead. He is convicted and now awaits sentencing.

Study 6 [Experimental group manipulation]:

Once released from his sentence, he will be required to undergo a new "drug court" program that involves treatment and close supervision in the community and re-incarceration if he uses drugs again. Indicate the sentence you would impose and the percentage, if any, of that sentence that could take place in the community under close supervision by probation officers. [Thirty dispositional options available.]

Study 7A [Both control and experimental groups given these options]:

_____ 1 to 6 years (depending on how well Donovan responds to treatment)
_____ 5 to 6 years (regardless of how well Donovan responds to treatment)

Study 7B [Both control and experimental groups given these options]:

_____ 1 to 10 years (depending on how well Donovan responds to treatment)
_____ 5 to 7 years (regardless of how well Donovan responds to treatment)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

Donovan successfully completes the program, and drug court authorities declare that he is now eligible for release by drug authorities on the condition that he continue to undergo drug testing every month. Should he be released on this condition? Or should he complete the sentence you imposed above? Please indicate your answer below.

Released on condition of monthly drug testing _____
Detain until term indicated above is completed _____

David Johnson is the mayor of a moderately sized city and is charged with deciding between competing bids for the management of a youth-detention facility. Company A has submitted the lowest bid and is a reputable company that provides services to numerous other cities in the state. However, Company B's CEO recently visited Johnson and offered him $5000 in exchange for awarding the contract to his company, which has a long record of improper treatment of juveniles and has submitted a much higher bid. Johnson is convicted of governmental corruption.

Study 6:

Not applicable

Study 7A [Both control and experimental groups given these options]:

_____ 10 years probation in the community to 6 years in prison (depending on Johnson's conduct during the 10 years)
_____ 5 to 6 years in prison (regardless of Johnson's subsequent conduct)

Study 7B [Both control and experimental groups given these options]:

_____ 1 year probation in the community to 10 years in prison (depending on Johnson's conduct during the 10 years)
_____ 5 to 7 years in prison (regardless of Johnson's subsequent conduct)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

Once convicted, Johnson offers to return the $5000. He also pledges $20,000 to the city, commits to working in community-outreach programs for juveniles for the next six years, and swears he has learned his lesson. Assuming all of this to be genuine, would it change the sentence you imposed above? Please indicate your answer below.

Release _____
Serve out sentence imposed above _____
The new sentence should be: _____[164]

164. Subjects who chose to change the sentence were given several lines on which to provide their answer. The same is true for the final question on the following survey.

A woman at work reveals Mark Foster's misdeeds to his employer, thereby getting him fired. Foster devises a plan to get even with her. The next week he forces the woman into his car at knife point and drives her to a secluded area where he shoots her to death. Foster is convicted and sentenced to prison. By the end of his third year there he has experienced a religious conversion, renounces violence, and provides restitution to the family.

Study 6:

Not applicable

Study 7A [Both control and experimental groups given these options]:

_____ 3 to 50 years in prison (depending on whether Foster's conversion is genuine)
_____ 48 to 50 years in prison (regardless of Foster's character)

Study 7B [Both control and experimental groups given these options]:

_____ 3 to 50 years in prison (depending on whether Foster's conversion is genuine)
_____ 45 to 50 years (regardless of Foster's character)

Study 7C [All subjects given the following information and options after being given the treatment information from Study 6]:

By the end of his third year there Foster has experienced a religious conversion, renounces violence, and provides restitution to the family. Assuming Foster is genuine, should he be released, serve out the term you imposed above, or receive a different sentence? Please indicate your answer below.

Release _____
Serve out sentence imposed above _____
The new sentence should be: _____

APPENDIX D: DEMOGRAPHIC INFORMATION

Demographic Characteristics		Studies 1, 2, 3 & 6 (n = 530)	Studies 4A & 7A (n = 236)	Studies 4B & 7B (n = 289)	Study 5 (n = 141)	Study 7C (n = 68)
Gender	**Female**	193 (37%)	120 (51%)	163 (56%)	78 (56%)	38 (56%)
	Male	337 (63%)	116 (49%)	126 (44%)	63 (44%)	30 (44%)
Race	**African Amer.**	28 (5%)	12 (5%)	25 (9%)	13 (9%)	0 (0%)
	White	451 (85%)	128 (54%)	211 (73%)	98 (70%)	64 (94%)
	Latino/a	16 (3%)	12 (5%)	19 (7%)	18 (13%)	4 (6%)
	Other	35 (7%)	84 (33%)	34 (12%)	12 (8%)	0 (0%)
Age (yrs.)	**Mean Age**	46.7	32.2	31.8	24.4	37.4
Political Party	**Republican**	140 (26%)	59 (25%)	53 (18%)	43 (30%)	22 (32%)
	Democrat	174 (33%)	99 (42%)	114 (39%)	53 (38%)	28 (41%)
	Independent	170 (32%)	52 (22%)	88 (30%)	45 (32%)	18 (27%)
	Other	46 (9%)	26 (11%)	34 (12%)	0 (0%)	0 (0%)
Income ($ per year)	**0-10,000**	41 (8%)	21 (9%)	18 (6%)	85 (60%)	0 (0%)
	10,001-25,000	67 (13%)	28 (12%)	42 (14%)	35 (25%)	3 (4%)
	25,001-35,000	54 (10%)	45 (19%)	52 (18%)	15 (11%)	5 (7%)
	35,001-50,000	87 (16%)	38 (16%)	65 (22%)	6 (4%)	18 (27%)
	50,001-75,000	143 (27%)	52 (24%)	57 (20%)	0 (0%)	18 (27%)
	75,001-100,000	73 (14%)	28 (12%)	31 (11%)	0 (0%)	20 (29%)
	100,000 +	65 (12%)	24 (8%)	24 (8%)	0 (0%)	4 (6%)
Education	**No H.S. Dipl.**	7 (1%)	11 (5%)	10 (3%)	0 (0%)	0 (0%)
	H.S. Dipl.	64 (12%)	22 (9%)	24 (8%)	0 (0%)	5 (7%)
	Partial Coll.	144 (27%)	61 (26%)	102 (35%)	0 (0%)	8 (12%)
	Associate Deg.	63 (12%)	12 (5%)	17 (6%)	0 (0%)	2 (3%)
	Bachelor Deg.	156 (30%)	82 (35%)	96 (33%)	121 (85%)	38 (56%)
	Graduate or Prof'l Deg.	96 (18%)	48 (20%)	40 (14%)	20 (15%)	15 (22%)