



2019

Promoting Equality Through Empirical Desert

Ilya Rudyak

University of Pennsylvania Law School, Ilya.Rudyak@law.upenn.edu

Follow this and additional works at: <https://scholarship.law.tamu.edu/lawreview>



Part of the [Criminal Law Commons](#), [Criminal Procedure Commons](#), [Law and Philosophy Commons](#), [Law and Society Commons](#), and the [Law Enforcement and Corrections Commons](#)

Recommended Citation

Ilya Rudyak, *Promoting Equality Through Empirical Desert*, 7 Tex. A&M L. Rev. 187 (2019).

Available at: <https://scholarship.law.tamu.edu/lawreview/vol7/iss1/6>

This Article is brought to you for free and open access by Texas A&M Law Scholarship. It has been accepted for inclusion in Texas A&M Law Review by an authorized editor of Texas A&M Law Scholarship. For more information, please contact aretteen@law.tamu.edu.

PROMOTING EQUALITY THROUGH EMPIRICAL DESERT

By: Ilya Rudyak*

ABSTRACT

According to empirical desert theory, good utilitarian grounds exist for distributing criminal punishment pursuant to the (retributive) intuitions of the lay community on criminal liability. This theory's insights, based on original empirical research and informed by social science, have significantly influenced contemporary criminal law theory. Yet, ostensibly, the theory is hampered by serious limitations, which may have obstructed its progress and its potential to guide criminal justice reform. Chief among them: it draws from community intuitions, and community intuitions—as the theory acknowledges—are sometimes immoral. In addition to these “immorality objections,” (commonly illustrated by alluding to the antebellum South and Nazi Germany), critics have alleged, inter alia, that the theory is self-defeating, uses incongruous justifications, and engages in deceptive and exploitative practices.

This Article argues that these critiques are misplaced and overstated and that empirical desert theory can be safely relied on in criminal justice—and beyond. Despite the captivating historical illustrations and the intuitive appeal of immorality objections, this Article demonstrates that empirical desert theory is nearly immune to them, by virtue of previously underappreciated features of its scientific methodology. Moreover, empirical desert theory can do even better.

This Article presents an innovative proposal to reconceptualize empirical desert theory by incorporating into its scientific methodology a minimalistic normative commitment to equality and non-discrimination. It provides theoretical support and specific parameters for this reconceptualization, which imbues the theory with qualities capable of further safeguarding it from immorality objections. Furthermore, the Article explores ten additional criticisms of the theory, seriatim, and demonstrates that the proposed reconceptualization substantially strengthens the theory's ability to overcome them. In its conclusion, the Article outlines two future paths for the theory's application beyond criminal law, discussing the possibility to “export” its insights to international humanitarian law and its potential to reframe the interaction between criminal law theory and philosophy.

* I wish to thank Professors Claire O. Finkelstein, Derek P. Jinks, Leo Katz, Milton C. Regan Jr., Edward Rock, and participants in workshops at the University of Pennsylvania Law School and Georgetown University Law Center. I have greatly benefited from their comments and insights, though, of course, the mistakes are my own. A special gratitude is due to Professor Paul H. Robinson, the late Professor Harry Reicher, and Ms. Maayan Roitfarb. Research for this Article was partially supported through the generosity of the University of Pennsylvania Fox Fellowship and the Hebrew University of Jerusalem Abraham Slavitt and Sylvan Cohen Scholarship for Advanced Studies, for which I am grateful. Finally, I thank the editors of the Texas A&M Law Review for their thoughtful and thorough work.

TABLE OF CONTENTS

I.	INTRODUCTION	189
II.	EMPIRICAL DESERT THEORY AS POTENTIALLY UNJUST	195
	A. <i>Empirical Desert Theory: The Basics</i>	195
	B. <i>Immorality Objections—Why Empirical Desert Theory May Be Unjust</i> ..	200
	1. Internal Acknowledgement—Immorality Objections and Deontological Desert	200
	2. External Criticisms and Intuitive Appeal of Immorality Objections	202
	C. <i>The Value of Scientific Methodology—Why Empirical Desert Theory is Unlikely to be Unjust</i>	206
	1. Methodological Challenges	206
	2. Doing Good Science and Reducing “Noise” in the Experiment ..	208
	3. Scenario Research Methodology and Its Impact on Immorality Objections	210
	4. Empirical Desert Theory Stands Its Ground Against Immorality Objections	214
	D. <i>Limitations of Scientific Methodology—Why Doing Good Science May Not Be Good Enough</i>	214
	1. Contingent Nature of Scientific Considerations	215
	2. Unsatisfying Nature of Scientific Reasons	218
III.	RECONCEPTUALIZING EMPIRICAL DESERT THEORY	220
	A. <i>Reconceptualizing Empirical Desert: Doing the Right Thing for the Right Reasons</i>	220
	1. The Proposal: Incorporating Normative Commitment to Equality and Non-Discrimination Into Empirical Desert Theory’s Methodology ...	221
	a. <i>What Does Commitment to Equality and Non-Discrimination Mean?</i>	221
	b. <i>Incorporating the Normative Commitment Into the Scientific Methodology</i>	224
	c. <i>Benefits of Incorporation: Relying on Different Reasons, Excluding Different Elements</i>	226
	i. Continuity and Contrast: Relying on Different Reasons ..	226
	ii. Contrast and Continuity: Excluding Different Elements ..	230
	B. <i>Perfect Compatibility Between Standard and Reconceptualized Empirical Desert Theory</i>	234
IV.	STRENGTHENING EMPIRICAL DESERT THEORY	240
	A. <i>Intuitions are Biased, and Empirical Desert Theory is Cherry Picking?</i>	241
	1. Community’s Intuitions are Biased, We Should Defer to Experts	241
	2. Empirical Desert Theory is “Cherry Picking” Among the Intuitions of the Community	243
	B. <i>Mind the Gap: Do Real-Life Intuitions Undermine Empirical Desert Theory?</i>	246
	C. <i>Is Empirical Desert Theory Is Misleading and Deceiving?</i>	254
	1. Misleading Use of “Intuitions of Justice” and “Normative Crime- Control”	254
	2. Arguments from Deception	257
V.	CONCLUDING THOUGHTS AND FUTURE PATHS	262

I. INTRODUCTION

A woman enters a dark alley when, suddenly, a figure with a gun in hand emerges from a hiding place in the shadows and shoots the woman repeatedly, killing her. Empirical research suggests that by the time you have finished reading the previous sentence, you have formed a strong intuition about the appropriate punishment for the shooter. A sentence imposing five years of imprisonment is unlikely to correspond to your intuition, appearing excessively *lenient*. Conversely, consider someone stealing a clock radio from the backseat of an unlocked, parked car. Your intuition might suggest that the same sentence would be extremely severe.¹ Yet, should the law care about your intuitions?

This question polarizes criminal law scholars, social scientists, and moral philosophers.² Some argue the law should be cautious and distrustful towards such intuitions.³ Others, thinking the law should take these intuitions into account, must grapple with a host of additional, related questions. Is it laymen intuitions that should be taken into account, the intuitions of criminal law experts, or perhaps those of moral philosophers? What kind of intuitions should be taken into account: Those arising at the spur of the moment or upon some reflection? And to what extent, exactly, should they influence criminal law?⁴ Beyond their theoretical value, answers to these questions have immediate practical implications on criminal policy and the law's precise and

1. Note, both examples are modified versions of experimental scenarios used to assess people's intuitions. See Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829, 1894–97 (2006). A subsequent study found that mean punishments assigned to an ambush shooting scenario ranged from life imprisonment to the death penalty. Conversely, mean punishment assigned to a clock radio theft scenario was 1.9 years. See Paul H. Robinson, Geoffrey P. Goodwin & Michael D. Reisig, *The Disutility of Injustice*, 85 N.Y.U. L. REV. 1940, 1972 (2010). Yet, these means can serve only as approximations, *inter alia*, because the examples in the main text differ from the experimental scenarios.

2. For scholarly collections that usefully introduce the diverse views and voices partaking in this debate, see, e.g., Symposium, *Is Justice Just Us? A Symposium on the Use of Social Science to Inform the Substantive Criminal Law*, 28 HOFSTRA L. REV. 601 (2000); Symposium, *First Annual Edward J. Shoen Leading Scholars Symposium: Paul Robinson*, 42 ARIZ. ST. L.J. 1089 (2010–2011) [hereinafter *Leading Scholars Symposium*]; CRIMINAL LAW CONVERSATIONS 29–66 (Paul H. Robinson, Stephen Garvey, & Kimberly Kessler Ferzan eds., 2011).

3. See, e.g., Deborah Denno, *The Perils of Public Opinion*, 28 HOFSTRA L. REV. 741, 744, 752–58 (2000); Kenneth W. Simons, *The Relevance of Community Values to Just Deserts: Criminal Law, Punishment Rationales, and Democracy*, 28 HOFSTRA L. REV. 635 (2000); Alice Ristroph, *Terror as a Theory of Punishment*, in RETRIBUTIVISM HAS A PAST: HAS IT A FUTURE? 155, 164 (Michael Tonry ed., 2011).

4. For discussions on these questions, see, e.g., Mary Sigler, *The Methodology of Desert*, 42 ARIZ. ST. L.J. 1173, 1174–75 (2010); Julian V. Roberts & Jan W. de Keijser, *Democratizing Punishment: Sentencing, Community Views and Values*, 16 PUNISHMENT SOC'Y 474, 478–80, 486 (2014); Adam J. Kolber, *How to Improve Empirical Desert*, 75 BROOK L. REV. 433, 445–46, 448–49 (2009).

detailed provisions determining the distribution of criminal punishment.⁵

A more general question concerning the justification of criminal punishment occupied the minds (and divided the views) of philosophers and criminal law scholars from Aristotle's and Plato's time to this day.⁶ From this quest two main justifications have emerged—the utilitarian and the retributivist—most famously formulated by Jeremy Bentham and Immanuel Kant.⁷ In a nutshell, the utilitarian justification is onward looking, and is mostly concerned with future benefits resulting from criminal punishment, for instance, deterrence, rehabilitation, or incapacitation of the dangerous.⁸ Conversely, the retributivist justification is concerned with the past, and focuses on the concept of “just desert.” Just desert (often called “deontological desert”) is traditionally developed by philosophers and criminal law theorists and revolves around matters related to offenders’ *blameworthiness*. Importantly, according to traditional retributivist view, a moral mandate exists to inflict upon offenders the punishment they deserve for their crimes, regardless of the *utility* of such punishment.⁹ The utilitarian and retributive perspectives on the justification of criminal punishment appear to be in irreconcilable tension.

Both questions—whether the law should care about intuitions and whether the utilitarian or retributive view on punishment should prevail—are important, complex, and seemingly unrelated. Yet, “Empirical Desert Theory,” developed since the turn of the century by Paul H. Robinson and coauthors, offers a thought-provoking and original answer to both.¹⁰ Focusing on the *distribution* of punishment, it develops an innovative account bridging utilitarian and retributivist thought¹¹ by using *laymen* intuitions as a crucial mediating factor.¹²

5. Paul H. Robinson, *The Ongoing Revolution in Punishment Theory: Doing Justice as Controlling Crime*, 42 ARIZ. ST. L.J. 1089, 1089–90 (2010). For a brief survey of ways community views had been incorporated into various countries’ criminal policy and law, see Roberts & de Keijser, *supra* note 4, at 474–76.

6. Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 454–55 (1997).

7. *Id.*

8. *Id.* at 454.

9. *Id.*

10. Work on empirical desert theory commenced in 1993. Robinson’s first book developing this theory was coauthored with John M. Darley, a prominent social psychology scholar who subsequently contributed to numerous empirical desert theory works. Other coauthors included Michael Cahill, Geoffrey P. Goodman, Owen Jones, and Robert Kurzban, among others. For the most comprehensive account of the theory, incorporating nearly two decades of research (including most empirical desert theory works cited in this Article), see generally PAUL H. ROBINSON, *INTUITIONS OF JUSTICE AND THE UTILITY OF DESERT* (2013).

11. Robinson & Darley, *supra* note 6, at 454–55, 468, 498.

12. Note, H.L.A. Hart famously bifurcated the question of criminal punishment’s justification to questions on the general *justifying* aim of punishment and principles that should guide its *distribution* (answering them by utilitarianism and retributivism respectively). See John Bronsteen, *Retribution’s Role*, 84 IND. L.J. 1129, 1137 (2009).

In brief, empirical desert theory is a *utilitarian* theory arguing that criminal punishment should be distributed according to its novel distributive principle of “empirical desert.” It is “empirical” because it relies on sophisticated empirical studies of community “intuitions of justice”—*lay* community members’ intuitive judgments on liability and punishment matters. And it uses the retributive label “desert” because a large body of research (to which empirical desert theorists contributed substantially)¹³ established, *inter alia*, that a community’s intuitions of justice are *retributive*—sensitive to matters revolving on past blameworthiness, similarly (though not identically) to *deontological* desert.

Empirical desert theory, informed by several strands of social scientific research, asserts that substantial *utilitarian* benefits can be gained by distributing punishment according to the (*retributive*) community’s intuitions of justice (i.e., according to its distributive principle of *empirical desert*). By aligning punishment distribution with empirical desert, the criminal justice system could gain “moral credibility” within the community, harness powerful mechanisms of “normative crime-control,” and subsequently increase compliance with the law. Conversely, the system could incur substantial compliance costs by deviating from empirical desert and distributing punishment according to the prescriptions of either *deontological* desert or utilitarian principles of deterrence, rehabilitation, or incapacitation.¹⁴

Thus, empirical desert theory ingeniously answers the aforementioned questions: Yes, the law *should* care about your (and other community members’) retributive intuitions, due to *utilitarian* reasons. Yet, this immediately raises another question: should we put our *trust* in empirical desert theory?

Opinions on this question differ strongly. On the one hand, empirical desert theory’s novel theoretical perspective, coupled with its commitment to meticulous empirical research (and sound scientific methodology), has contributed to its substantial influence on modern criminal law thinking. The theory has inspired thoughtful academic discussion and writing,¹⁵ provided nuanced recommendations on criminal justice reforms, and has been relied on, approvingly, in diverse contexts.¹⁶ Even the theory’s *opponents*, committed to other utilita-

Agnostic to general justifying aims, empirical desert theory focuses on punishment *distribution*. See Robinson, *supra* note 5, at 1189–90.

13. See generally ROBINSON, *supra* note 10, at 5–94.

14. Obviously, this is a very rough sketch of the theory, and additional details will be provided in *infra* Section II.A. For the most comprehensive account, see ROBINSON, *supra* note 10.

15. See, e.g., Symposium, *Is Justice Just Us*, *supra* note 2; Leading Scholars Symposium, *supra* note 2; Paul H. Robinson, *Empirical Desert*, in CRIMINAL LAW CONVERSATIONS, *supra* note 2, at 29–66.

16. See, e.g., ROBINSON, *supra* note 10, at 208–36, 487–92 (advice on reforms); Douglas A. Berman, *A Truly (and Peculiarly) American “Revolution in Punishment*

rian distributive principles, recognized the theory findings' thought-provoking nature and its value in helping frame the retributivist-utilitarian divide in utilitarian terms.¹⁷

Moreover, despite empirical desert theory's utilitarian nature,¹⁸ arguably it should hold a special appeal for *retributively* minded philosophers. Admittedly, different methods are used for settling on "just desert" by philosophers and empirical desert theorists. The former use hypotheticals, or the method of "reflective equilibrium," contemplating on "*deontological* desert."¹⁹ The latter conduct social science experiments assessing "*empirical* desert." Yet, since empirical desert relies on lay intuitions influenced by *blameworthiness*, distributing punishment accordingly would align with retributive philosophers' views more closely than distributing punishment according to *other* utilitarian distributive principles (e.g., deterrence or incapacitation).²⁰

Furthermore, and importantly, empirical desert theory's insights are not limited to municipal criminal law. In principle, they can have much broader implications. For instance, they could be employed to bolster compliance in legal contexts where it is notoriously problematic, such as international law,²¹ and especially international *humanitarian* law.²²

On the other hand, despite the many virtues and wide-ranging potential of empirical desert theory, it appears to be hampered by serious limitations. Perhaps the most important among them could be labeled "immorality objections," and are candidly acknowledged by the theory: It draws from community intuitions, and community intuitions are sometimes morally unjust.²³ Naturally, other utilitarian the-

Theory", 42 ARIZ. ST. L.J. 1113, 1118–21 (2010) (procedural criminal law); Margaret M. deGuzman, *Harsh Justice for International Crimes?*, 39 YALE J. INT'L L. 1, 16–17, 24, 32–33 (2014) (international criminal law); Max Minzner, *Why Agencies Punish*, 53 WM. & MARY L. REV. 853, (2012) (regulatory agencies).

17. Christopher Slobogin & Lauren Brinkley-Rubinstein, *Putting Desert in Its Place*, 65 STAN. L. REV. 77, 77–79, 119 (2012); Mark R. Fondacaro & Megan J. O'Toole, *American Punitiveness and Mass Incarceration*, 18 NEW CRIM. L. REV. 477, 488 (2015).

18. Paul H. Robinson, *Reply*, in CRIMINAL LAW CONVERSATIONS, *supra* note 2, at 61; Robinson, *supra* note 5, at 1104–05; Kolber, *supra* note 4, at 439.

19. Robinson, *supra* note 15, at 37.

20. ROBINSON, *supra* note 10, at 173–74.

21. For a rare application of the theory to international law, see, e.g., Paul H. Robinson & Adil Ahmad Haque, *Advantaging Aggressors: Justice & Deterrence in International Law*, 3 HARV. NAT'L SECURITY J. 143 (2011); Margaret M. deGuzman, *supra* note 16.

22. See, e.g., Jelena Pejic, *Strengthening Compliance With IHL: The ICRC-Swiss Initiative*, 98 INT'L REV. RED CROSS 315 (2016) (describing the need and efforts to bolster IHL compliance).

23. See, e.g., ROBINSON, *supra* note 10, at 167; Paul H. Robinson, *The Role of Moral Philosophers in the Competition Between Deontological and Empirical Desert*, 48 WM. & MARY L. REV. 1831, 1838 (2006); see also Robinson, *supra* note 5, at 1108; Robinson, *supra* note 15, at 37.

ories are similarly vulnerable to immorality-based critiques.²⁴ One might devise examples in which maximizing incapacitation or deterrence requires imposing draconian sentences or punishing innocents.²⁵ Yet, immorality objections appear much more endemic to empirical desert theory, as they seem to rest directly on the intuitions of a community in a certain temporal and geographic context. The examples of the U.S. Old South or Nazi Germany, pointed out by Robinson, provide stark reminders that one needs not go too far back in history to find examples of societies with highly objectionable attitudes to justice.²⁶ Structuring a criminal justice system based on such societies' intuitions might have some utilitarian benefits; however, doing so without having some external moral check²⁷ appears to be normatively repugnant.

Moreover, empirical desert theory has been subjected to diverse *specific* criticisms by both utilitarian and retributively inclined scholars over the years. Some objected to the theory's reliance on the public's distorted intuitions instead of deference to criminal law experts' better judgments.²⁸ Others criticized it for "cherry picking" the intuitions it uses, potentially failing to reap utilitarian compliance benefits, or incongruously (for a utilitarian theory) relying on non-utilitarian justifications.²⁹ Others still claimed the theory is self-defeating because it overlooks the gap between the intuitions *it* relies on and the intuitions the *public* relies on in the real world, and discounts the implications this gap has on the theory's efficacy.³⁰ And some critics further claimed that empirical desert theory violates important moral principles such as the "publicity condition,"³¹ deceives community members, and exploits their retributive intuitions.³²

24. ROBINSON, *supra* note 10, at 172.

25. These general examples merely illustrate these critiques to which utilitarians have developed sophisticated responses. *See generally* Fondacaro & O'Toole, *supra* note 17.

26. ROBINSON, *supra* note 10, at 167.

27. *See, e.g.*, ROBINSON, *supra* note 10, at 172–74 (suggesting such an external check could have been developed by moral philosophy, yet, due to the methods contemporary moral philosophers use, it is unlikely to be forthcoming).

28. *See, e.g.*, Denno, *supra* note 3, at 752–53; Christopher Slobogin, *Some Hypotheses About Empirical Desert*, 42 ARIZ. ST. L.J. 1189, 1190–91 (2010); Julian V. Roberts, *The Future of State Punishment: The Role of Public Opinion in Sentencing*, in RETRIBUTIVISM HAS A PAST: HAS IT A FUTURE? 106–07 (Michael Tonry ed., 2011).

29. Kolber, *supra* note 4, at 441–48.

30. *See, e.g.*, Jan de Keijser, *Penal Theory and Popular Opinion: The Deficiencies of Direct Engagement*, in POPULAR PUNISHMENT: ON THE NORMATIVE SIGNIFICANCE OF PUBLIC OPINION 101, 102, 107–09 (Jesper Ryberg & Julian V. Roberts eds., 2014); Roberts & de Keijser, *supra* note 4, at 490–91.

31. In brief, this condition requires that principles underlying a moral system could be publicly announced without being undermined by their publicity. *See, e.g.*, Kolber, *supra* note 4, at 459, and text accompanying *supra* notes 324–99.

32. *See, e.g.*, Marcelo Ferrante, *Community Views and Criminal Law Reform*, 77 REV. JURIDICA U. P.R. 459, 467–70 (2008). Kolber, *supra* note 4, at 434–35, 453–60.

Immorality objections and the manifold specific criticisms strongly suggest that empirical desert theory should *not* be trusted. The theory's virtues notwithstanding, the nature, extent, and variety of its critiques may have obstructed its potential to guide criminal justice reform and exert even broader influence than it had. They may have discouraged scholars and policy makers from applying its findings and insights in criminal (or other fields of) law. Indeed, why should anyone seriously contemplate relying on what is (purportedly) a morally objectionable theory, which depends on distorted intuitions, employs incongruous justifications, operates in a self-defeating manner, and engages in deceptive and exploitative practices? If correct, such concerns substantially reduce the theory's appeal—perhaps most prominently in a democracy committed to liberal values.

Yet, I argue that these concerns are misplaced and grossly overstated. Conceptually, I advance two original theses. First, I demonstrate that by virtue of previously underappreciated methodological features, empirical desert theory, as it *presently* stands, is nearly immune to immorality objections. Second, I propose an innovative reconceptualization of empirical desert theory (entailing incorporation of a minimalistic normative commitment to equality and non-discrimination into the theory's scientific methodology). This reconceptualization imbues the theory with qualities capable of *further* safeguarding it from immorality objections and significantly strengthens and augments its ability to overcome the aforementioned specific criticisms. Consequently, I argue that, for both utilitarian and non-utilitarian *normative* reasons, empirical desert theory *should* be trusted and its findings and insights can be safely relied on in criminal law—and beyond.

The Article proceeds as follows. Section II starts with some additional essential details on empirical desert theory. It then presents the immorality objections, discusses their intuitive appeal in the context of empirical desert theory, and, drawing on historical examples, highlights their adverse implications. Next, however, it demonstrates that *in practice*, certain overlooked aspects of the theory's scientific methodology make its reliance on morally objectionable, *immoral* intuitions highly unlikely. Having demonstrated the *practical* virtues of empirical desert theory's methodology, the discussion pivots to highlight the *conceptual* limitations inherent in relying on *scientific* methodology to avoid *moral* concerns.

Section III responds to these limitations by introducing my key proposal to reconceptualize empirical desert theory by incorporating into its methodology a minimalistic normative commitment to equality and non-discrimination. It provides theoretical support for such reconceptualization, elucidates its meaning, presents its merits, and, finally, demonstrates its perfect compatibility with empirical desert theory as it presently stands.

Section IV then builds on the proposal's detailed presentation and explores, *seriatim*, the manifold specific criticisms of empirical desert theory in its light. It provides the theory's responses to criticisms, and then demonstrates how the proposal bolsters these responses, strengthening empirical desert theory.

Finally, Section V outlines two future paths for the theory's prospect to exert broader influence, addressing the potential to "export" its insights to international humanitarian law, and the theory's potential to change the interaction between criminal law theory and philosophy.

A final substantive comment: this Article aims to innovate, however—and crucially—it is *not* revolutionary in nature. It does *not* attempt to replace an existing theory with a new ground-breaking theory of its own. While such deconstructive approach is appropriate at times, this is not the purpose of the exercise undertaken here. It is rather, and quite intentionally so, a *reconceptualization* of the existing theory of empirical desert. My goal in such reconceptualization is *constructive*, and it consists of both relying on and reinforcing the foundations of empirical desert theory.³³

II. EMPIRICAL DESERT THEORY AS POTENTIALLY UNJUST

A. *Empirical Desert Theory: The Basics*

In essence, the underlying idea in empirical desert theory ("EDT")³⁴ is that good *utilitarian* grounds exist for distributing criminal punishment according to the *retributive* concept of "just desert,"³⁵ as understood by the *lay* community (rather than by criminal law experts or moral philosophers). EDT's focus on the community's assess-

33. The Article's constructive purpose (alongside space constraints) has implications on its presentation style, especially (though not exclusively) in dealing with the multiple specific criticisms of empirical desert theory in Section IV. Empirical desert theory can, and often has provided *detailed* responses to these criticisms; Section IV summarizes these responses' pertinent parts briefly (and provides full references), yet it focuses on the potential of the theory's *reconceptualized* account to augment these responses or refute the criticisms. Importantly, the brief summaries of the original responses are *not* intended to detract from their force. Additionally, the Article addresses only broad and conceptually significant critiques of the theory, bracketing "local" criticisms of experimental design or specific results. For discussions of such criticisms, see, e.g., Slobogin & Brinkley-Rubinstein, *supra* note 17; Paul H. Robinson, Joshua S. Barton & Mathew Lister, *Empirical Desert, Individual Prevention, and Limiting Retributivism: A Reply*, 17 *NEW CRIM. L. REV.* 312 (2014); Christopher Slobogin, *Empirical Desert and Preventive Justice: A Comment*, 17 *NEW CRIM. L. REV.* 376 (2014).

34. Throughout the introduction I referred to empirical desert theory by its full title, which is descriptively valuable. Yet, below I mention the theory several hundred times, and restating its full name would infuse the text with unwieldy expressions (e.g., "immorality objections to empirical desert theory"). Therefore, moving forward, I use the abbreviation "EDT" as denoting empirical desert theory.

35. Paul H. Robinson, *Competing Conceptions of Modern Desert: Vengeful, Deontological, and Empirical*, 67 *CAMBRIDGE L. J.* 145, 153–55 (2008).

ment of crime and punishment—its “intuitions of justice”³⁶—has practical, rather than philosophical foundations.³⁷ It is rooted in social science findings related to the nature of people’s intuitions of justice³⁸ and the sources of people’s obedience to society’s rules of prescribed conduct.³⁹

A large body of research (to which EDT researchers contributed substantially)⁴⁰ examined intuitions on justice, finding that people’s assessments of crime and punishment matters are *intuitive*, *nuanced*, *universal*, and *retributive* in nature.

EDT uses controlled social science experiments to assess intuitions, employing a sophisticated “scenario research” methodology.⁴¹ Researchers present subjects with scenarios of potentially criminal acts, and subjects are tasked to decide whether acts described in each scenario deserve punishment and assign an appropriate punishment (if any) to the perpetrator.⁴² The results are quite striking.

First, people’s assessments of the scenarios are indeed *intuitive*, appearing in a rapid, effortless, and unconscious manner. They resemble a “gut feeling,” phenomenologically closer to the facile processes of perception rather than the effortful processes of deliberation.⁴³ Second, contrary to common wisdom at the time of EDT’s development, the community’s intuitions of justice are not vague or general, nor can they only be operationalized at extreme circumstances arising at the edges of the punishment continuum. Rather, the community’s intuitions of justice are remarkably *nuanced*, yet predictable. Even small variations in scenarios result in large, and anticipated, changes in subjects’ responses.⁴⁴ Third, and again defying the common wisdom that there can be no agreement on matters of crime and punishment,⁴⁵ EDT studies show astonishing levels of *consensus* between subjects on the *ordinal* ranking of scenarios⁴⁶ (especially on matters related to physical aggression, takings without consent, and deception in exchanges—matters EDT defines as the “core” of criminal wrongdoing).⁴⁷ Importantly, EDT studies also show that when people consider

36. Robinson & Darley, *supra* note 6, at 454–55; Robinson, *supra* note 35, at 149; Robinson, Goodwin, & Reisig, *supra* note 1, at 1961–65, 1995–2023.

37. Robinson, *supra* note 35, at 149.

38. See generally ROBINSON, *supra* note 10, at 5–94.

39. See Robinson, *supra* note 35, at 149.

40. See generally ROBINSON, *supra* note 10, at 5–94.

41. *Id.* at 240, 421.

42. *Id.* at 240.

43. *Id.* at 5–8.

44. *Id.* at 13–16.

45. Note, both domestic and cross-cultural studies challenging the common wisdom and testing the agreement on relative blameworthiness of serious wrongdoing were conducted since the 1960s and found a significant amount of consensus. For some descriptions of these studies, see *id.* at 23–28.

46. ROBINSON, *supra* note 10, at 30–31.

47. *Id.* at 1, 34.

crime and punishment matters, they do *not* focus on utilitarian aspects such as deterrence or incapacitation,⁴⁸ but rather intuitively assess factors related to offenders' blameworthiness. These assessments are similar (though not identical), to those that might have been predicted by *retributively* inclined moral philosophers.⁴⁹

These findings have broad implications, especially when evaluated together with a related body of research, conducted domestically and abroad, suggesting that relative blameworthiness assessments are similar across cultures and demographics.⁵⁰ Whether such intuitions are innate in humans due to evolutionary biology, acquired through social learning or some other unknown mechanism, their *near universal* uniformity implies they can withstand the powerful pressures of culture and society.⁵¹ In other words, people's retributive, widely-shared, and nuanced intuitions of justice are also *robust*, and unlikely to be easily changed. This feature of community's intuitions of justice has practical implications for social reformers—it highlights that criminal law doctrines distributing punishment according to *utilitarian* principles of deterrence, rehabilitation, and incapacitation are likely to create friction with community's intuitions of justice and therefore face constant complications.⁵²

The other scientific foundation for EDT's position, namely that for good *utilitarian* reasons punishment should be distributed according to people's *retributive* intuitions of justice, is rooted in the sources of obedience to society's rules of prescribed conduct.⁵³ Social science suggests that the main reasons for such obedience are not threats of punishment but more nuanced, yet powerful, influences of social and individual moral control. These include people's concern that violating the law will bring disapproval and condemnation by their social group, coupled with perception of themselves as moral beings wishing to do the morally right thing as they perceive it.⁵⁴

EDT maintains that the criminal system can benefit greatly if it succeeds in harnessing these compliance-inducing forces of "normative crime control," which it juxtaposes to "coercive crime control."⁵⁵ The latter forces aim to increase deterrence or incapacitation and are reflected in common utilitarian doctrines such as three strikes or strict

48. *Id.* at 108.

49. *Id.* at 164–67.

50. For a survey of these studies see *id.* at 19–28.

51. For potential explanations regarding intuitions of justice's origins, see *id.* at 35–62.

52. For brief exploration of modern utilitarian crime control doctrines conflicting with intuitions of justice, see *id.* at 110–28. For a summary of research highlighting the disutility of deviating from intuitions of justice, see *id.* at 176–88.

53. Robinson, *supra* note 35, at 149.

54. Robinson & Darley, *supra* note 6, at 468, 468–70; ROBINSON, *supra* note 10, at 152–53.

55. ROBINSON, *supra* note 10, at 95, 141–62.

liability statutes⁵⁶ that commonly *conflict* with a community's intuitions of justice.⁵⁷ To employ *normative* crime control, the criminal system itself must establish "moral credibility" within the community, that is, to be perceived by the community as committed to doing justice in crime and punishment matters.⁵⁸

Importantly, the system must be mindful that such reputation is difficult to gain, easy to lose, and not only results, but *intentions* to do justice are crucial.⁵⁹ Though unintentional deviations may be overlooked by the community, deliberate ones are exceptionally damaging.⁶⁰ Therefore, the system must (while acknowledging practical limitations) "try to do justice as best it can."⁶¹ And since community's intuitions of justice are retributive in nature (constituting the default, initial *benchmark* against which its members measure the law's alignment with justice), to gain moral credibility the criminal law should follow suit. It must align itself with a community's retributive (and also nuanced, widely-shared, and robust) intuitions of justice.

Through becoming morally credible, the criminal system could harness effective "normative crime control" mechanisms.⁶² These mechanisms amplify positive societal dynamics of stigmatization, deference in border-line cases, and influence on social norms. A morally credible system increases its ability to rely on the compliance-inducing (yet essentially cost-free) force of *stigmatization*, by conjoining violations of criminal laws with defiance of social norms that triggers condemnation and censure by the community.⁶³ A morally credible system can also gain community deference in borderline cases when the conduct's criminality is not immediately apparent or unsettled (e.g., insider trading), as community members may reflexively accept the law's determinations, trusting its reputation for doing justice.⁶⁴ Perhaps most importantly, a morally credible system can play a role in shaping social norms, because it may be perceived as a highly *persuasive* source—that is, a source legitimate in authority, expert in knowledge, and trustworthy in motives (according to research on persuasion).⁶⁵ If the criminal system is perceived as a highly persuasive source, its posi-

56. For a review of these and other incapacitation and deterrence enhancing crime-control doctrines, see *id.* at 111–20.

57. For findings on such conflict, see *id.* at 120–28; see also Robinson, Goodwin, & Reisig, *supra* note 1.

58. Obviously, the alignment may be imperfect, but a rough approximation of such an alignment is expected. See *infra* Section IV.B.

59. ROBINSON, *supra* note 10, at 183.

60. *Id.*

61. *Id.*; *cf. id.* at 193–94 (explaining the stringent conditions that according to EDT may justify deviation from empirical desert).

62. ROBINSON, *supra* note 10, at 153.

63. *Id.* at 154–55.

64. *Id.* at 162–63.

65. *Id.* at 161–62.

tions on an issue can impact debates (and thoughts) on what the social norms on that issue *ought* to be.⁶⁶

Additionally, a morally credible system can ameliorate the negative societal dynamics of vigilantism, resistance to, and subversion of the law. A system misaligned with community's intuitions of justice inspires "classic" vigilantism—ordinary citizens may take the law into their own hands and engage in formally criminal, dangerous actions which, in their view, are morally justifiable.⁶⁷ Moreover, such a system may inspire similar (more subtle, but more pervasive) dynamics of resistance to and subversion of the law by various actors in criminal proceedings⁶⁸ (e.g., policemen "testilying" on search and seizure "technicalities,"⁶⁹ prosecutors overcharging without proof,⁷⁰ jurors disregarding instructions⁷¹). These dynamics spiral downward, erode the system's moral credibility, and lead to generalization of disrespect towards it.⁷² A morally credible system, in contrast, may break this cycle, inspiring acquiescence or even cooperation.⁷³

That said, EDT is not blindly committed to following a community's intuitions of justice. While cognizant of the negative impact deviations from empirical desert may generate and typically favoring *non*-deviation, it takes a complex approach.⁷⁴ EDT recognizes that some deviations (for instance, on matters outside the "core" of wrongdoing, when no consensus exists), are *unavoidable*,⁷⁵ and appreciates that *perfect* alignment between law and community intuitions of justice is sometimes unachievable—the system can only strive to maximize alignment when possible.⁷⁶ Moreover, EDT accepts that certain deviations might be *desirable*, either to maximize desert-based punishment⁷⁷ or to promote other important societal interests.⁷⁸

Finally, EDT focuses not on desert itself, but rather on the utilitarian crime-control benefits associated with it. Thus, true to its utilitarian nature, EDT is open to deviations from empirical desert when "a valuable crime-control opportunity presents itself."⁷⁹ While such op-

66. *Id.*

67. *Id.* at 155; see, e.g., Paul H. Robinson, *The Moral Vigilante and Her Cousins in the Shadows*, 2015 U. ILL. L. REV. 401, 412–14 (2015) (providing examples of such vigilantism in Harlem, New York).

68. ROBINSON, *supra* note 10, at 156.

69. Robinson, *supra* note 67, at 467–68.

70. *Id.* at 469–71.

71. ROBINSON, *supra* note 10, at 157.

72. For review of studies supporting this point, see *id.* at 158–59.

73. *Id.* at 156.

74. For a detailed discussion of such deviations, and the conditions under which they may be permitted see *id.* at 189–207.

75. *Id.* at 190.

76. *Id.* at 191.

77. *Id.* at 191–93.

78. *Id.* at 196.

79. *Id.* at 194.

portunities might be rare (due to short and long term negative impacts of deviations), if utilitarian cost-benefit analysis suggests otherwise in a particular case, EDT would support taking advantage of it.⁸⁰ EDT would not forego the possibility to reap a potential crime-control “bonanza.”⁸¹ Generally, however, EDT recommends that the criminal system actively seek opportunities to *reduce* deviation,⁸² thereby shoring up its moral credibility and earning “credibility chips,” which it should spend only very judiciously.⁸³

B. *Immorality Objections—Why Empirical Desert Theory May Be Unjust*

EDT’s innovative approach and its sophistication in reliance on social science to refute common wisdoms and provide policy recommendations contributed to its significant influence on modern criminal law thinking. As mentioned in the introduction, the theory’s value is acknowledged even by its opponents who favor other utilitarian distributive principles.⁸⁴ Moreover, despite EDT’s instrumentalist crime-control bend, its reliance on empirical desert may be appealing to retributively-minded criminal law scholars and moral philosophers, since punishment distribution according to empirical desert is likely to be more closely aligned with deontological desert than its distribution according to other utilitarian distributive principles.⁸⁵

Yet, certain objections appear to offset much of EDT’s virtues. These objections aim at EDT’s very *foundation*: its reliance on community intuitions of justice. While relying on these intuitions may generate utilitarian crime-control benefits, these intuitions can be immoral. Neither proponents nor critics of EDT ignored this crucial issue.

1. Internal Acknowledgement—Immorality Objections and Deontological Desert

EDT’s potential reliance on immoral intuitions is candidly acknowledged by Robinson, who repeatedly stated that a community’s intuitions can be unjust:

Empirical desert can tell us *only* what people *think* is just; . . . Like slave owners in the Old South or anti-Semitic Germans before World War II, one may fail to appreciate the injustice of one’s views until later, especially if one’s views at the time are shared by a large

80. *Id.* at 192, 194.

81. *See* Robinson, *supra* note 15, at 38.

82. ROBINSON, *supra* note 10, at 191, 198.

83. *Id.* at 198.

84. Slobogin & Brinkley-Rubinstein, *supra* note 17, at 77–78, 119; Fondacaro & O’Toole, *supra* note 17, at 488.

85. ROBINSON, *supra* note 10, at 173–74.

number of other people. Even a popular liability or punishment rule may be unjust.⁸⁶

Robinson even discussed a potential solution to the problem: employing *deontological* desert as a “transcendent check” on immoral intuitions of a community, because deontological desert:

[T]ranscends the particular people and situation at hand and embodies a set of principles derived from fundamental values; principles of right and good; and thus will produce justice without regard to the political, social, or other peculiarities of the situation at hand. As Henry Sidgwick famously put it, moral judgments are made “from the point of view of the universe.”⁸⁷

Yet, Robinson also expressed doubts that deontological desert (as developed by *contemporary* philosophers) can serve as a “transcendent check” on empirical desert. Contemporary philosophers’ methods typically involve testing *intuitions* on the proper resolution of cases against cleverly crafted hypothetical variations of them. Philosophers then further refine and retest hypotheticals, generating new (and sometimes conflicting) intuitions, typically attempting to reach reflective equilibrium.⁸⁸ Notably, the method of reflective equilibrium is widely used by philosophers,⁸⁹ *inter alia*, for justifying the very pro-

86. *Id.* at 167 (emphasis added); *See also* Robinson, *supra* note 23, at 1838 (“generally tracking a community’s shared intuitions of justice may well build some *credibility* within that population that the criminal justice system is doing justice, but it does *not* follow that in fact justice is being done. Shared intuitions might simply be *wrong*; they tell us only what lay persons *believe* is just”); Robinson, *supra* note 5, at 1108; Robinson, *supra* note 15, at 37.

87. ROBINSON, *supra* note 10, at 164. *See also id.* at 167 (“Only deontological desert can give us the means by which we can tell the truth of what is deserved, insulated from the vicissitudes of human irrationality.”) Note, some commentators disagree with Robinson’s deontological desert definition, claiming his position on this issue is not entirely clear. *See* Alice Ristroph, *How (Not) to Think Like a Punisher*, 61 FLA. L. REV. 727, 748 (2009). Others fault Robinson’s referencing “deontological” as expressing methodology, rather than substantive commitment, and maintain that viewing desert as a top-down philosophical conception from the “point of view of the universe” is hard to recognize as desert at all. *See* Youngjae Lee, *Desert, Deontology, and Vengeance*, 42 ARIZ. ST. L.J. 1141, 1143–45 (2010). The descriptive account in the main text, however, aims to *present* Robinson’s arguments accurately, not to *resolve* disagreements on deontological desert’s nature.

88. Robinson, *supra* note 15, at 37; *see also* EUGEN FISCHER & JOHN COLLINS, *Rationalism and Naturalism in the Age of Experimental Philosophy*, in EXPERIMENTAL PHILOSOPHY, RATIONALISM, AND NATURALISM: RETHINKING PHILOSOPHICAL METHOD 3–33, 10–11 (Eugen Fischer & John Collins eds., 2015) (mentioning references to such practice as: “analytic philosophy’s ‘standard justificatory procedure’”).

89. The method of “Reflective Equilibrium,” made prominent by John Rawls in the late part of the 20th century, entails more than simply testing competing intuitions. Yet, for present purposes the crucial part is the methods’ partial reliance on *intuitions*, as the main text below clarifies. For a more comprehensive exploration of the method, *see generally* Norman Daniels, *Reflective Equilibrium*, STAN. ENCYCLOPEDIA OF PHIL., <https://plato.stanford.edu/entries/reflective-equilibrium/> (last modified Oct. 14, 2016) [<https://perma.cc/H7DL-8BSW>]; Yuri Cath, *Reflective Equilibrium*, in THE OXFORD HANDBOOK OF PHIL. METHODOLOGY (Herman Cappelen, Tamar

ject of retributive justice and developing the concept of “deontological desert.”⁹⁰

The main problem with these methods for developing deontological desert, Robinson claims, is their own partial reliance on (potentially immoral) intuitions.⁹¹ Philosophers employing these methods cannot reliably spot the immoral *intuitions* EDT might rely on, thus the deontological desert they develop cannot be used as a “transcendent check” on empirical desert.⁹² Put differently, insofar as intuitions play a role in developing *deontological* desert, it cannot serve as an external, Archimedean point, required to keep *empirical* desert (and the intuitions it relies on) in check.⁹³

Curiously, limitations of deontological desert in spotting immoral intuitions are a mixed blessing for EDT. Initially it seems that refuting deontological desert’s ostensible superiority over empirical desert makes EDT *more* appealing. Yet, if deontological desert cannot restrain EDT’s potential reliance on immoral intuitions, EDT seems *less* appealing.⁹⁴

2. External Criticisms and Intuitive Appeal of Immorality Objections

Scholars have not failed to notice these features of EDT and their implications. Some emphasized that EDT’s increased reliance on community’s intuitions “must be set against the very real possibility of unjust outcomes.”⁹⁵ Others accentuated that congruence in a community’s intuitions carries no normative significance,⁹⁶ nor can

Szabó Gendler, & John Hawthorne eds., 2016). For a brief survey of the methods’ criticisms and responses to them, see Folke Tersman, *Recent Work on Reflective Equilibrium and Method in Ethics*, PHIL. COMPASS, June 2018, at 1, <https://onlinelibrary.wiley.com/doi/epdf/10.1111/phc3.12493> [<https://perma.cc/R4FN-APF8>]. For the methods’ presentation in EDT’s context, see Sigler, *supra* note 4.

90. See generally Alec Walen, *Retributive Justice*, STAN. ENCYCLOPEDIA OF PHIL. (June 18, 2014), <https://plato.stanford.edu/archives/win2016/entries/justice-retributive> [<https://perma.cc/7MVD-87H5>].

91. Note, while Robinson makes this point in the context of *deontological desert*, prominent critics of the method of reflective equilibrium, especially utilitarians, also criticized it for its reliance on intuitions. See Daniels, *supra* note 89 (sections 2.2 & 4.1).

92. Robinson, *supra* note 23, at 1840. For a more detailed critique see *id.* at 1838–43. For a response from the perspective of moral philosophy see Sigler, *supra* note 4. Note again, the main text aims only to *present* Robinson’s position. We will return to the question of philosophers’ reliance on intuitions in *infra* Sections III.B. and V.

93. The lack of an Archimedean point for philosophers is acknowledged by *proponents* of the method of reflective equilibrium as well. See Sigler, *supra* note 4, at 1179 n.39. Note also that for present purposes it matters not if philosophers use their *own* intuitions or their assessments (whether correct or not) of laymen intuitions. As long as intuitions figure in the equation, the problem refuses to go away.

94. ROBINSON, *supra* note 10, at 173–74.

95. Roberts, *supra* note 28, at 112.

96. Roberts & de Keijser, *supra* note 4, at 481.

normative conclusions be drawn from it.⁹⁷ Even scholars sympathetic to EDT stressed its possible reliance on immoral intuitions as *the* most compelling normative objection to the theory,⁹⁸ while others noted that EDT seems to avoid engaging with the issue fully.⁹⁹

There is something natural yet surprising in this line of criticism. Criticizing criminal justice theories for being potentially unjust or immoral is natural. Yet immorality objections are not uniquely applicable to EDT; these are *common* objections applying to *all utilitarian* theories (including theories traditionally proposing to distribute criminal punishment according to deterrence, incapacitation, or rehabilitation).¹⁰⁰ Moreover, among these competing utilitarian theories *only* EDT gives weight to aspects related to *moral* blameworthiness, thus EDT may be the *least* objectionable theory from a moral standpoint.¹⁰¹

What is it, then, that causes critics to lament *EDT's* potential immorality—although the theory is advocated by its proponents and understood by its critics as utilitarian?¹⁰² What makes immorality objections especially compelling in the case of EDT?

Ironically, the reason may be the extraordinary *intuitive* appeal of immorality objections in EDT's context. Consider the familiar examples stressing the potential immorality of *other* utilitarian distributive principles, such as framing an innocent scapegoat to increase deterrence, or not prosecuting murderers elderly and infirm enough to make considerations of incapacitation or rehabilitation irrelevant.¹⁰³ Such examples usually involve *extreme* cases testing these distributive principles' *boundaries*¹⁰⁴—because only such cases make immorality objections intuitively appealing for these principles.¹⁰⁵ In contrast, the *intuitive* appeal of immorality objections against EDT is not restricted to radical examples at the theory's outer limits. The problem appears much more pervasive. After all, EDT is founded on a community's intuitions, and no sophisticated or extreme examples are required to

97. Donald Braman, Dan M. Kahan, & David A. Hoffman, *Some Realism about Punishment Naturalism*, 77 U. CHI. L. REV. 1531, 1592, 1596 (2010).

98. Michael T. Cahill, *A Fertile Desert?*, in CRIMINAL LAW CONVERSATIONS, *supra* note 2, at 43–44.

99. Zachary R. Calo, *Empirical Desert and the Moral Economy of Punishment*, 42 ARIZ. ST. L.J. 1123, 1133 (2010).

100. Robinson, *supra* note 15, at 37.

101. ROBINSON, *supra* note 10, at 173–74.

102. Robinson, *supra* note 5, at 1104–05; Robinson, *supra* note 18, at 61; Kolber, *supra* note 4, at 439; Braman, Kahan, & Hoffman, *supra* note 97, at 1539 n.15; Calo, *supra* note 97, at 1126; Roberts & de Keijser, *supra* note 4, at 486.

103. Alice Ristroph, *Desert, Democracy, and Sentencing Reform*, 96 J. CRIM. L. & CRIMINOLOGY 1293, 1301 (2006); ROBINSON, *supra* note 10, at 142.

104. For a similar suggestion see Ristroph, *supra* note 103, at 1301.

105. See generally Fondacaro & O'Toole, *supra* note 17; Bronsteen, *supra* note 12, at 1142–44. But see *id.* at 1143 n.84 (citing Lloyd L. Weinreb, *Desert, Punishment, and Criminal Responsibility*, 49 LAW & CONTEMP. PROBS. 47, 47 (1986)).

see that such intuitions might be morally problematic.¹⁰⁶ People occasionally recognize that their own intuitions are not always perfectly aligned with moral demands. Moreover, people frequently notice the incongruence between *other* people's intuitions and moral demands. Consider the following relatively commonplace example:

Morning Commute: Sam drives to the office in regular morning traffic. Suddenly, a car driving far faster than the legal speed limit bypasses him from the right, sharply cuts into Sam's lane, speeds away, cuts into another lane, and eventually disappears without colliding with Sam or anyone else.

Some people in Sam's situation intuitively think, at least initially, that the bypassing driver deserves an exceptionally severe punishment (some of us may have even expressed, or overheard others express, such judgments in no uncertain terms). Moreover, even the admirable individuals who would not overestimate the bypassing driver's deserved punishment but are acquainted with others who would display less laudable judgment, may entertain doubts whether their *acquaintances'* intuitions should be relied on by the criminal justice system.¹⁰⁷

Commonplace illustrations aside, perhaps the most persuasive reason for immorality objections' intuitive appeal in EDT's context lies in Robinson's "compelling"¹⁰⁸ examples of the Old South or Nazi Germany.¹⁰⁹ These examples starkly remind us of relatively recent, yet abhorrent, societal attitudes towards justice, and raise doubts about the wisdom of relying on intuitions of "justice" pervading in such societies. Considering certain features of these societies' legal regimes, even briefly, is useful to appreciate the full force of immorality objections.

Racial discrimination had been an established social practice in the antebellum South. Simplistically put, racial discrimination was recognized and institutionalized by, *inter alia*, criminal laws that differentiated between whites and blacks, both authorizing whites to use force against blacks and significantly limiting blacks' right to use force in

106. Some immorality objections in this vein focused on *intuitors*, arguing that people's intuitions are biased by their *own* characteristics such as demographics, life experiences, or emotions. See, e.g., Denno, *supra* note 3, at 752. Others focused on *perpetrators*, arguing that people's intuitions are biased by *perpetrators'* characteristics such as race which should not be "considered legitimate punishment goals." See, e.g., Slobogin, *supra* note 28, at 1190–91; Denno, *supra* note 3, 752–53.

107. As one philosopher aptly put it "we do not treat . . . certain judgements about responsibility made in an agitated state as authoritative—we understand that there are circumstances in which people are tempted to blame somebody even if in a cool hour they themselves would acknowledge that nobody is to blame." Antti Kauppinen, *The Rise and Fall of Experimental Philosophy*, 10 PHIL. EXPLORATIONS 95, 103–04 (2007); see also Joshua Knobe & Shaun Nichols, *An Experimental Philosophy Manifesto*, in EXPERIMENTAL PHILOSOPHY 8 (Joshua Knobe & Shaun Nichols eds., 2008).

108. Roberts, *supra* note 28, at 111.

109. ROBINSON, *supra* note 10, at 167.

self-defense.¹¹⁰ With this context in mind, some EDT critics noted that situations we today might view as clear cases of premeditated murder or legitimate self-defense (disregarding the defendant's race), may have elicited a very different understanding in the Old South, if they involved white and black defendants, respectively.¹¹¹

The case of Nazi Germany's legal system is similar in some respects; it too relied on racially-based institutionalized discrimination. Moreover, its rhetoric explicitly relied on *community* views: the *voelkisch* conception of justice. High-ranking jurists and other regime officials commonly employed terms such as "the interest of the community" or "the demands of popular conscience" in public speeches and internal correspondence.¹¹²

Furthermore, this approach was formally incorporated into the regime's criminal law. As early as June 1935, the Reich government amended the criminal legislation, formerly codifying the legality principle, by a provision allowing to punish: "Whoever commits an act which the law declares as punishable or which *deserves* punishment according to the fundamental idea of a penal law *or the sound sentiment of the people . . .*"¹¹³ Hans Frank, at the time a high-ranking

110. Braman, Kahan, & Hoffman, *supra* note 97, at 1579–81. For an interesting discussion of approaches taken by courts in that period attempting to resolve the paradox between slaves' status as property and their personhood in the context of the use of physical coercion, see Jedediah Purdy, *People as Resources: Recruitment and Reciprocity in the Freedom-Promoting Approach to Property*, 56 DUKE L.J. 1047, 1060–68 (2007).

111. Braman, Kahan, & Hoffman, *supra* note 97, at 1579–81.

112. IAN WARD, LAW, PHILOSOPHY AND NATIONAL SOCIALISM: HEIDEGGER, SCHMITT AND RADBRUCH IN CONTEXT (1992), *reprinted in* HARRY REICHER, III LAW, THE HOLOCAUST, INTERNATIONAL HUMAN RIGHTS: MATERIALS AND COMMENTARY 456–72 (U. Penn. L. Sch., preliminary ed., 2011). The Reich's Justice minister, Franz Gürtner, explained in 1933: "The vital root of law reaches down into the depths of the *popular conscience* and then supplies it with its inner validity and affirmation." *Id.* at 472 (emphasis added). SS Brigadefuehrer, Otto Ohlendorf (who commanded Einsatzgruppen D that conducted racial extermination in southern Ukraine), disturbingly stated: "SS and Gestapo only acted according to the 'healthy intuitive feeling of the Volk.'" *Id.* at 464.

113. The Code of Criminal Procedure and the Judicature Act was amended correspondingly, stipulating in article 170a "[i]f an act deserves punishment *according to the sound sentiment of the people*, but is not declared punishable in the law, the prosecution will examine whether the underlying principle of a penal law can be applied to the act and whether justice can be helped to triumph by *analogous* application of that penal law" (i.e., Article 2 of the Penal Code). *See* Code of Criminal Procedure and Judicature Act, June 28, 1935, REICHSGESETZBLATT, TEIL I [RGL I] at 844 (Ger.), *reprinted in* HARRY REICHER, III LAW, THE HOLOCAUST, INTERNATIONAL HUMAN RIGHTS: MATERIALS AND COMMENTARY 479 (U. Penn. L. Sch., preliminary ed., 2011). Similarly, article 267a allowed analogous application when: "the trial shows that the defendant has committed an act which deserves punishment *according to the sound sentiment of the people . . .*" *See* Section 1 Article 2 of the Law Amending the Criminal (Penal) Code, June 28, 1935, REICHSGESETZBLATT, TEIL I [RGL I] at 839 (Ger.), *reprinted in* HARRY REICHER, III LAW, THE HOLOCAUST, INTERNATIONAL HUMAN RIGHTS: MATERIALS AND COMMENTARY 477 (U. Penn. L. Sch., preliminary ed., 2011); *see also* Section 1 Article 170a and 267a Code of Criminal Procedure and

official of the Nazi Party and president of the Academy of German Law, explained that by this law “the liberalist foundation of the old penal code ‘no penalty without a law’ was definitely abandoned and replaced by the postulate ‘no crime without punishment,’ which corresponds to our conception of the law.”¹¹⁴ He further emphasized: “[i]n the future, criminal behavior, even if it does not fall under formal penal precepts, *will receive the deserved punishment* if such behavior is considered punishable *according to the healthy feelings of the people*.”¹¹⁵

Even this brief presentation of the Old South and Nazi Germany’s legal regimes vividly illustrates the potential immorality of relying on the community’s intuitions. While structuring a criminal justice system based on these intuitions might still generate some utilitarian benefits, doing so with no effective external check on these intuitions may be morally repugnant.¹¹⁶

The examples above illuminate immorality objections’ unique intuitive appeal in EDT’s context and may account for the emphasis given to them by some EDT critics. Despite their appeal however, immorality objections are wide off the mark. Indeed, the dangers these examples highlight are of paramount importance for a criminal law theory. Yet, they are largely beside the point for EDT. The next Section explains why.

C. *The Value of Scientific Methodology—Why Empirical Desert Theory is Unlikely to Be Unjust*

1. Methodological Challenges

To understand why immorality objections to EDT are quite beside the point, we must look at the theory’s methodology more closely. Recall, the common wisdom when EDT developed was that there is no, and cannot be, consensus on crime and punishment matters.¹¹⁷ *Prima facie*, this is a tenable position: people commonly disagree on such issues, adopting different positions in both abstract debates and particular cases. However, appearances of disagreement can be deceiving. Some individuals are generally more punitive than others, and societies vary significantly in setting their endpoints of punishment.

Judicature Act, June 28, 1935, REICHSGESETZBLATT, TEIL I [RGL I] at 844 (Ger.), *partial translation reprinted in* HARRY REICHER, III LAW, THE HOLOCAUST, INTERNATIONAL HUMAN RIGHTS: MATERIALS AND COMMENTARY 477–79 (U. Penn. L. Sch., preliminary ed., 2011).

114. *See id.* at 474–75 (emphasis added).

115. Similarly, in 1936, Franz Schlegelberger, then State Secretary in the Reich Ministry of Justice, clarified that the law permitted one’s punishment: “even if his deed is not punishable according to the law, but *if he deserves punishment in accordance with the basic concepts of the criminal law and the sound instincts of the people*.” *See id.* at 474–75 (emphasis added).

116. *See* ROBINSON, *supra* note 10, at 12–14, 167.

117. *See supra* Section II.A.

This leads to sharp disagreements on cardinal punishment amounts, even if a consensus exists on *ordinal* ranking of offenses.¹¹⁸

Moreover, some disagreements may have little to do with people's *intuitions of justice* and may arise due to different perceptions of relevant *facts*. The case of O.J. Simpson usefully illustrates that people's perceptions of police conduct influence their trust in police testimony, and disagreements on police trustworthiness generated variance in popular judgment.¹¹⁹ Additional reasons for disagreements on criminal justice issues (e.g., the death penalty) may be related to the broader political and social context that the position one takes on such issues signifies.¹²⁰

EDT had to settle on a methodology capable of penetrating all these factors, otherwise it had no chance to discover whether a consensus on intuitions of justice exists. Yet, just asking people about front-page cases or politically contentious issues is obviously not a good option. Another methodological choice (customary for philosophers and criminal law professors) is presenting people with stylized hypotheticals of criminal conduct.¹²¹ But hypotheticals entail significant shortcomings.¹²² They are typically concise, purposefully aiming to isolate some aspect of the case.¹²³ In contrast, people usually rely on multiple factors when assessing blameworthiness and require a minimum of relevant facts.¹²⁴ When hypotheticals lack such facts, people tend to “read in” the missing details.¹²⁵ Naturally, different people may “read in” different facts (depending on their life experiences, beliefs, prejudices, etc.) and make different inferences about the hypotheticals.¹²⁶ Thus, two people with identical intuitions of justice reading the *same* hypothetical may *differ* in their liability judgments, simply because each perceives a different factual situation. Public

118. Note, practically having an ordinal ranking of offenses and defining an end point of the punishment continuum (be it fifteen years or life imprisonment) is sufficient for EDT to recommend a *specific, cardinal amount* of punishment for every offense on the punishment continuum. This transformation is possible because the number of distinguishable points on the punishment continuum is limited. See ROBINSON, *supra* note 10, at 10–11, 168.

119. Robinson, *supra* note 15, at 33–34.

120. ROBINSON, *supra* note 10, at 171–72.

121. *Id.* at 244–45.

122. For a critique similar to the one in the text below but directed at certain *empirical* studies, rather than hypotheticals, see Robinson & Kurzban, *supra* note 1, at 1860.

123. See ROBINSON, *supra* note 10, at 245.

124. *Id.*

125. *Id.*

126. *Id.* For a more detailed discussion of factors people might read into hypotheticals see *id.* at 245–47. See also Eugen Fischer & Paul E. Engelhardt, *Stereotypical Inferences: Philosophical Relevance and Psycholinguistic Toolkit*, 30 *RATIO* 411, 414 (2017) (arguing stereotypical inferences are “bound to be made” when reading hypotheticals and: “The *less informative* context these descriptions provide, the *more likely* stimulus-driven stereotypical inferences are to go through unmodified.”) (emphasis added).

opinion polling, another common methodology, also commonly fails to provide necessary information,¹²⁷ thus involving some shortcomings of hypotheticals but lacking their focus.

These methods are unlikely to discover whether a consensus on intuitions of justice exists. As Robinson noted: “*Poor testing methods* will predictably underestimate the extent of agreement. When a test scenario is written ambiguously so that different test participants perceive the facts differently, the existence of shared nuanced intuitions of justice itself will predict different judgments among the participants.”¹²⁸

2. Doing Good Science and Reducing “Noise” in the Experiment

Viewed from a purely scientific perspective, the aforementioned methodologies and their shortcomings generate (or fail to reduce) experimental “noise.”¹²⁹ Since this noise may thwart accurate assessments of intuitions, EDT required a different methodology. Instead of relying on “poor testing methods,” EDT was committed to what might be called ‘doing good science’ objective. And because EDT needed a method capable of penetrating through the noise,¹³⁰ it adopted the sophisticated and reliable empirical methodology of “scenario” research.¹³¹

Conceptually, this methodology’s experimental procedures are straightforward. Experiment subjects read short vignettes—scenarios—describing potentially criminal conduct,¹³² such as this:

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 \$15 T-shirt in his coat and walks out, with no intention of paying for it.¹³³

Subjects are asked to determine whether “John” should be criminally liable and, if so, how much punishment he should receive.¹³⁴ Next, subjects read different scenarios modified according to the theories researchers examine.¹³⁵ For instance, they may read a scenario (theorized to elicit a relatively proximate judgment)¹³⁶ about stealing a

127. Opinion polling also privileges punitive responses. See, e.g., Roberts, *supra* note 28, at 105–06; Roberts & de Keijser, *supra* note 4, at 475, 482.

128. Robinson, *supra* note 15, at 33–34.

129. Importantly, the term “noise” is *not* used in this article to express *moral* judgment, but rather as a morally neutral *scientific* term of art, similar to its use in EDT literature. See, e.g., Robinson & Kurzban, *supra* note 1, at 1840, 1861 n.145, 1869, 1874, 1875, 1880.

130. *Id.* at 1880.

131. See ROBINSON, *supra* note 10, at 171–72, 240.

132. *Id.* at 240.

133. Robinson & Kurzban, *supra* note 1, at 1895 (quoted verbatim).

134. See ROBINSON, *supra* note 10, at 240.

135. *Id.*

136. For the reasons behind such theorizing see Robinson & Kurzban, *supra* note 1, at 1900–01.

clock radio from an unlocked car.¹³⁷ Throughout the experiment subjects respond to multiple (randomized) scenarios theorized by researchers to diverge on the punishment continuum. Eventually, researchers can compare between different subjects' responses to all scenarios and determine whether a consensus exists between their (ordinal) rankings of criminal conduct.

Crucially, EDT researchers craft scenarios carefully and cautiously. It is a delicate balancing act, aimed to minimize experimental noise. First, researchers must use concrete factual descriptions, not abstract offense-labels, or skeletal fact patterns.¹³⁸ Scenarios must provide more information than typical hypotheticals—lest subjects “read in” additional, diverging details. Second, researchers must avoid incorporating numerous factors often present in real-life cases and public debates on crime and punishment—factors introducing “loaded” political or social context that might skew subjects' perceptions of the situation.¹³⁹

Researchers constantly strive to avoid scenario language that may generate experimental noise. Yet, despite their best efforts, language invoking “noise-inducing” factors might still remain.¹⁴⁰ Thus, before using their scenarios in the main experiment, researchers conduct meticulous “field testing” (also called “manipulation checks”) to ensure their scenarios would be perceived by subjects uniformly, and exactly as intended.¹⁴¹ The manipulation checks are mini-studies in which subjects are not even asked to provide liability assessments. They are rather asked about their *understanding* of what the scenarios mean.¹⁴² And if they do not understand the scenarios as intended, researchers must go back to the drawing board, modify their scenarios, and repeat the manipulation checks process.¹⁴³

137. *See id.* at 1895.

138. *Id.* at 1866. *See also* Christopher Freiman & Shaun Nichols, *Is Desert in the Details?* 82 PHIL. PHENOMENOLOGY RES. 121, 128 (2011).

139. Note also, subjects are *not* asked to provide their general punishment distribution theories, since views on such matters may depend on broad political and social context. *See* ROBINSON, *supra* note 10, at 164, 171–72; *accord* Andrew E. Taslitz, *Why Did Tinkerbell Get Off So Easy?: The Roles of Imagination and Social Norms in Excusing Human Weakness*, 42 TEX. TECH L. REV. 419, 461 (2009) (“These [scenarios] also help to reduce the passions and politics of confronting subjects with current real-world cases about which they may know little . . .”).

140. Recall, “noise” is a scientific term-of-art, and the phrase “noise-inducing factors” does *not* imply that such factors are somehow morally problematic or irrelevant—only that they inadvertently influence experimental subjects. Note also that additional potential sources for noise in the experiment exist (e.g., subjects' focus or motivation may decrease during the study, rendering their judgments less precise). For a brief discussion of such effects see, e.g., Robinson & Kurzban, *supra* note 1, at 1869, 1874–75. The procedures described in the text relate to minimizing a more fundamental source of noise, rooted in the scenarios themselves.

141. ROBINSON, *supra* note 10, at 245.

142. *Id.* at 245.

143. *Id.* For a detailed description of necessary manipulation checks and their extent, see *id.* at 453–65. Robinson aptly summarizes the required effort: “The good

Scenario crafting and conducting manipulation checks require effort and resources, sometimes requiring much more time than running the main experiment.¹⁴⁴ Nevertheless, these steps are necessary to ensure that differences in subjects' judgments result from the variables researchers aim to test, rather than from artifacts generated by each subject's idiosyncratic understanding of the scenarios.

3. Scenario Research Methodology and Its Impact on Immorality Objections

EDT's rationale for adopting scenario research methodology is *purely* scientific. This methodology enables EDT to do good science, minimizing the experimental noise inherent in "poor testing methods."¹⁴⁵ Yet, EDT's adoption of this methodology has important, heretofore overlooked, implications on immorality objections. It renders them largely moot.¹⁴⁶

Recall the key argument of immorality objections: EDT relies on the community's intuitions that can be immoral; therefore, EDT (and its recommendations) can be immoral too. Indeed, intuitions prejudiced by elements like race, color, religion, national origin, etc. would constitute a morally-objectionable foundation for a criminal justice system. Had EDT relied on such biased intuitions, these objections would have been merited. Understanding EDT's scenario research methodology, however, exposes this argument's Achilles' heel. This methodology nearly ensures that such biasing elements will be excluded, and therefore would *not* influence EDT's recommendations.

Importantly, reasons for excluding these elements have less to do with EDT's *moral* commitments, than with its dedication to *doing good science*. EDT's methodology aims to minimize *noise*, taking numerous steps to avoid ambiguity in scenarios and its negative influence on subjects.¹⁴⁷ Consequently, EDT researchers craft scenarios carefully. They use concrete factual descriptions, not offense labels; provide information sufficient to prevent subjects from "reading in" additional details; omit factors which might introduce "loaded" political or social context and skew subjects' perceptions; and conduct manipulation checks to ensure unambiguous understanding.¹⁴⁸

manipulation-check results reported in this study are not the result of either good luck or a special talent in drafting scenarios, but rather the result of dozens of manipulation-check mini-field tests together with three formal manipulation-check pilot tests. After each test, adjustments were made to the scenario's texts, which were then retested." *Id.* at 464–65.

144. *Id.* at 247.

145. Robinson, *supra* note 15, at 33–34.

146. The qualifier here (and elsewhere in this section) indicates that occasionally these objections are not mooted completely. Such rare cases are discussed in *infra*, Section II.D.

147. *See supra* Section II.C.2.

148. *See supra* Section II.C.2.

These methodological steps ensure exclusion of elements like race, religion, etc. from EDT's scenarios. Morality aside, such scenario elements are likely to generate considerable noise in the experiment. Some subjects will alter their responses according (and proportionally) to their political positions, stereotypes, or prejudices.¹⁴⁹ Others will modify their responses in the opposite direction, to compensate for their own biases. Others still might be unaffected (yet difficult to identify).¹⁵⁰ These scenario elements' disparate influences would needlessly introduce a cacophony of noise into the experiments,¹⁵¹ likely to result in a methodological nightmare. Consequently, EDT researchers exclude such elements meticulously, and contemporary EDT (and its recommendations) will not rely on intuitions affected by them, thus undermining immorality objections.

EDT's scientific methodology also undermines the special *intuitive* appeal of immorality objections in ordinary situations similar to the "morning commute" example. Clearly, EDT's experiments do not rely on real-world intuitions¹⁵² of infuriated or otherwise disturbed people.¹⁵³

Yet, one may still wonder whether EDT's scientific methodology undermines the most compelling examples of immorality objections: the Old South and Nazi Germany. The short answer is positive. In all probability, EDT would not have relied on immoral intuitions even in these societies for similar, scientific reasons.

Consider the following instructive example developed by scholars who modified a typical EDT scenario (criticizing EDT on grounds *other* than immorality objections):

149. For a brief and relatively recent overview of stereotype activation, see Fischer & Engelhardt, *supra* note 126, at 411–13.

150. These modifications might happen even on an intuitional level. For a discussion of a possibility to modify even *intuitive* responses, including those related to stereotypes, see Milton C. Regan, Jr, *Moral Intuitions and Organizational Culture*, 51 ST. LOUIS U. L.J. 941, 961–64 (2007); John Darley, *Realism on Change in Moral Intuitions*, 77 U. CHI. L. REV. 1643, 1649–50 (2010).

151. Note, some of these effects may be offset in randomized experiments. Yet, the offset might be partial if biases (or egalitarian compensatory impulses), are not normally distributed in the population. Moreover, scenario elements such as race and religion introduce *ambiguity* potentially causing experimental subjects to understand *different* scenario elements (or their *context*) differently. This may generate additional unnecessary noise in the already noisy environment in which EDT operates.

152. Relatedly, recall the immorality objections asserting people's intuitions are biased due to *individual* life experiences or demographic characteristics. See *supra* note 106. While such characteristics may have real-world influence, consensus on intuitions of justice (assessed by scientific methodology), which cuts across cultures and demographics, suggests that objections rooted in *these* biasing effects are not well grounded.

153. Note, EDT was *criticized* for disregarding such judgments, a criticism I address in *infra* Section IV.A.2.

John owns a slave who has highly offended him. As he had planned the day before, he waits for his slave to return from work and, when he appears, John shoots him to death.¹⁵⁴

These scholars pointed to the “*distinctive* understandings [of this scenario] in colonial and contemporary American communities”;¹⁵⁵ yet, the more interesting question for us is whether this modified scenario could have been used in a hypothetical EDT experiment in the Old South.¹⁵⁶ If so, it would have elicited immoral intuitions in subjects, which would have influenced EDT’s recommendations in that society. This reasoning appears to reinforce immorality objections.

Yet, this reasoning is also wrong. EDT experiments, even if conducted in the Old South, would *not* have used language such as: “owns a slave.” Recall, EDT experiments use *factual* descriptions of situations, rather than labels denoting complex legal status, such as slavery. More importantly, using the term “slave”—even in the Old South—would have likely elicited differential reactions in different subjects (corresponding to the widely *differing* attitudes towards slavery).¹⁵⁷

A similar analysis would apply, *mutatis mutandis*, to including language indicating “Jewishness”¹⁵⁸ in a hypothetical EDT scenario in

154. Donald Braman, Dan M. Kahan & David A. Hoffman, *Some Realism about Punishment Naturalism*, 77 U. CHI. L. REV. 1531, 1580 (2010). In the original EDT scenario, the victim is a woman. See Robinson & Kurzban, *supra* note 1, at 1897.

155. Braman, Kahan, & Hoffman, *supra* note 154, at 1580. These scholars’ original argument is not pertinent for the present discussion. In brief, they suggest that EDT will have difficulty in explaining such distinctive understandings, while their preferred theory, “Punishment Realism,” will not. For responses to Punishment Realism claims see Paul H. Robinson, Owen D. Jones, & Robert Kurzban, *Realism, Punishment, and Reform*, 77 U. CHI. L. REV. 1611 (2010); Darley, *supra* note 150. *But see* Donald Braman, Dan M. Kahan & David A. Hoffman, *A Core of Agreement*, 77 U. CHI. L. REV. 1655 (2010).

156. Note, I do *not* claim these scholars *argued* that it could. I merely use their scenario for illustration.

157. For a detailed description of such attitudes and their transformations from the colonial era until the Civil War see generally IRA BERLIN, *GENERATIONS OF CAPTIVITY: A HISTORY OF AFRICAN-AMERICAN SLAVES* 39–43, 67–81, 88–96, 124–40, 140–57, 163–209, 209–30 (2003). For a discussion of diverging attitudes towards slavery and slaves’ *legal* status, reflected in court decisions, see, e.g., Purdy, *supra* note 110, at 1060–68. Consider, for instance, the following passage from an 1828 Kentucky Court of Appeals opinion:

Slaves, although they are human beings, are by our laws placed on the same footing with living property of the brute creation. However deeply it may be regretted, and whether it be politic or impolitic, a slave by our code, is not treated as a person, but (*negotium*), a thing, as he stood in the civil code of the Roman Empire.

See id. at 1061 n.29 (quoting *Jarman v. Patterson*, 23 Ky. (7 T.B. Mon.) 644, 644 (Ky. 1828)). The chilling legal reasoning clearly acknowledges the highly *contentious* nature of slavery and the societal *controversies* surrounding it.

158. In discussing immoral intuitions throughout the Article, I refer to some morally (and politically) loaded elements of empirical-desert-like scenarios. When feasible, I use the element “Jewishness” for illustration because I happen to be reasonably acquainted with the sensitivities it may raise (though its interconnected meanings of religion, nationality, and race can introduce ambiguities). Yet, occasionally I must use

Nazi Germany. Even during the Third Reich, public attitudes towards Jews *varied* significantly and would have likely elicited disparate reactions among different subjects.¹⁵⁹

The main point is that even in the Old South or Nazi Germany terms like “slave” or “Jew” would have introduced unnecessary noise into experiments. Including them in experimental scenarios might have doomed researchers’ chances to assess their subjects’ intuitions. Consequently, even in these societies, researchers following EDT methodology would have excluded such terms—which undermines immorality objections even in their most compelling examples.

Yet, one may ask, what if EDT researchers in these societies, perhaps fully internalizing their prejudices, were inattentive to the *contentious* nature of terms like “slave” or “Jew”? Couldn’t *these* researchers employ the aforementioned modified scenarios, thus reviving immorality objections?

Even in this farfetched, unrealistic situation, EDT reliance on such modified scenarios is highly unlikely, due to EDT scientific methodology. First, note that prejudiced researchers may still recognize that *subjects’* responses to scenario elements like “slave” or “Jew” may vary considerably. Because this variance may introduce undesirable experimental noise, even prejudiced researchers may conclude, for scientific reasons, that using such elements is counterproductive. Second, these researchers may conclude so after conducting manipulation checks, facing evidence that these elements cause undesirable noise. Finally, even if, inexplicably, such elements are not excluded *ex-ante* from the scenarios, they are still likely to create ambiguity and noise impacting experimental results. This in turn should alert the

other terms related to slavery and the antebellum South. My familiarity with the sensitivities *these* terms may raise is, admittedly, much more limited. Therefore, two points are in order. First, my choice of particular elements to illustrate *specific* points is immaterial to the general analysis that applies to *any other* potentially immoral element based on race, religion, national origin, or other characteristics of a similar nature. Second, if, despite my best efforts, I accidentally treat sensitive subjects in a less than optimally sensitive way, I offer my sincere apologies. No offense is intended.

159. Note, the Nazi Party invested considerable resources in its anti-Semitic propaganda, indicating that it sought to *change* general public’s attitudes. Historical research also suggests that such attitudes varied greatly. For instance, purging Jews from German Civil Service and universities was largely met approvingly. See DAVID BANKIER, *THE GERMANS AND THE FINAL SOLUTION: PUBLIC OPINION UNDER NAZISM*, 69 (David Sorkin ed., 1996). Yet, the Jewish businesses boycott initiated by the Nazi Party in 1933 was met with mixed responses: “[L]arge sectors [of the public] found the form of persecution abhorrent Even Generals wearing medals came to stores owned by Jews in Berlin to demonstrate their *disapproval* of the Nazi policy.” *Id.* at 68–69. Moreover, as late as September 1941, following a decree ordering German Jews to wear the Star of David “expressions of sympathy [from German population] were not infrequent.” See SAUL FRIEDLÄNDER, *NAZI GERMANY AND THE JEWS, 1933–1945* 254 (abr. ed. 2009)); see also *id.* at 39–41. Note, these examples of *behavior* probably conceal a much deeper divide on *attitudes* towards Jews (not to mention their *distribution*, or *prevalence*). Yet, they certainly suggest these attitudes were not *uniform*.

prejudiced experimenters to reexamine their scenarios *ex-post*. For these methodological reasons, even if EDT experiments were conducted by morally-prejudiced and inattentive researchers in the Old South or Nazi Germany, in all likelihood EDT would *not* have relied on immoral intuitions.

4. Empirical Desert Theory Stands Its Ground Against Immorality Objections

Let us recap. Immorality objections were openly acknowledged by EDT proponents and stressed by its opponents as the theory's greatest critique.¹⁶⁰ These objections seem sensible and intuitively appealing.¹⁶¹ Yet, EDT's concrete scientific methodology renders them largely moot.¹⁶² Contemporary EDT researchers exclude from their scenarios elements such as race and religion which may invoke immoral intuitions, for *scientific*, methodological reasons, namely the experimental *noise* they would generate.¹⁶³ Moreover, even (hypothetical) EDT researchers operating in the Old South or Nazi Germany would most likely have excluded such elements, for the same scientific reasons.¹⁶⁴ Even the (more hypothetical) morally-prejudiced researchers in these societies would have likely recognized (or empirically discovered) this noise-inducing influence and would have excluded them.¹⁶⁵

Thus, my first thesis is upheld. EDT, as it *presently* stands, is nearly immune to immorality objections. Ostensibly doubly vulnerable to these objections—as a patently *utilitarian* theory relying heavily on community *intuitions*—EDT is effectively protected from relying on *immoral* intuitions by its *scientific methodology*. Thus, surprisingly, it appears that for all practical (and some hypothetical) purposes, EDT stands its ground against immorality objections. This conclusion corresponds well with my constructive purpose by substantially strengthening EDT in comparison to other distributive principles. Once EDT's greatest critique is undermined, relying on it becomes much more appealing (especially bearing in mind EDT's other virtues). Yet, below I argue that EDT can and should be strengthened even further.

D. *Limitations of Scientific Methodology—Why Doing Good Science May Not Be Good Enough*

Indeed, scientific methodology safeguards EDT from relying on immoral intuitions. It works in practice. But one may wonder: Does it work in theory? Witticisms aside, depending on scientific considera-

160. *See supra* Section II.B.

161. *See supra* Section II.B.2.

162. *See supra* Sections II.C.1., II.C.2.

163. *See supra* Section II.C.2.

164. *See supra* Section II.C.3.

165. *See supra* Section II.C.3.

tions to avoid moral concerns involves conceptual limitations, which are rooted in the *contingent nature* of scientific considerations and, on a deeper level, the very *reasons* they are based on. To appreciate these limitations, we must understand more clearly *why* EDT methodology happens to work so well, and *why* we find it satisfactory.

1. Contingent Nature of Scientific Considerations

Crucially, the noise-generating potential of certain elements is not inherent. Rather, it depends on (contingent) features of the *society*. When EDT experiments are conducted, the noise-generating potential of each element depends on the diversity of attitudes towards it in that society.

In the Old South and Nazi Germany where race and religion issues were historically contentious,¹⁶⁶ using scenario-elements invoking race would have been ill-advised from a purely scientific perspective; it would have generated undesirable experimental noise. Theoretically, however, the situation could have been different. Alternative histories of the Old South or Nazi Germany can be imagined, in which the population (including the oppressed minorities) is uniformly convinced that one's race or religion are *essential* characteristics for distributing liability and punishment. Yet imagining such dystopic situations based on (relatively) modern societies may appear too speculative. Contemplating a time preceding the spread of ideas about human *equality* may be more instructive, therefore the Babylonian society circa 1700 B.C.E., governed by the Code of Hammurabi, may serve as a suitable springboard for discussion.

The Code of Hammurabi famously differentiated between people according to their social status, and its punishments varied accord-

166. See *supra* Section II.C.3. Note, both societies attempted to *dehumanize* their victims. Recall the allusion to slaves as placed by law “on the same footing with living property of the brute creation.” See *supra* note 157. Yet perhaps the most notorious and graphic of dehumanization attempts was made in the Nazi film “The Eternal Jew.” The following description is illuminating: “In a particularly horrendous sequence of the film swarms of rats scurried through cellars and sewers, and, in rapid alternation, hordes of Jews moved from Palestine to the most remote corners of the world. The text was on a par: ‘Where rats turn up . . . they spread diseases and carry extermination into the land. They are cunning, cowardly and cruel; they mostly move in large packs, exactly as the Jews among the people.’” FRIEDLÄNDER, *supra* note 159, at 189 (internal references omitted). This imagery was further replicated in Nazi publications throughout Europe. *Id.* at 189–90. Note though, such dehumanization attempts subtly reveal that the dehumanizers believe their intended audience regards the to-be-dehumanized minority as *human* (which is why dehumanization efforts are required). Thus the very existence of dehumanization efforts based on a certain element (such as race) usually *indicates* disagreement in the society on that element. Accordingly, it would be noise-generating in that society. Usually—but not always. Dehumanization efforts might ultimately succeed (enduring, however, for prophylactic, or ritualistic reasons), and not indicate societal disagreement or corresponding elements’ noise-generating potential.

ingly.¹⁶⁷ For instance, the punishment of a *free* man striking one of *higher* status was public whipping; yet for striking one of *equal* status, it was a monetary fine.¹⁶⁸ In contrast, the punishment of a freed man's *slave* striking a free man was ear cutting.¹⁶⁹ What interests us in these laws is not their precise content¹⁷⁰ but the possible societal outlook they might have reflected.¹⁷¹ While difficult to fathom today, that ancient society's outlook *might* have been that legal distinctions based on social status are natural, just, and morally required.

Subsequently, questions of experimental noise might be answered differently in Babylon and more modern societies. While even in the Old South scenario elements denoting social status (i.e., slavery) were noise-generating, they may *not* have been noise-generating in ancient Babylon. Quite the contrary: scenarios *lacking* information on social status of scenario actors could have been noise-generating. Babylonian subjects might have regarded the social status of perpetrators and victims in scenarios as *essential* elements. Lacking them, different subjects would "read in" different social statuses into scenarios, and experimental noise would ensue.

The key point is this: the very same scientific considerations of minimizing experimental noise that guide modern EDT researchers to *exclude* elements such as social status from their experiments would have guided Babylonian researchers to *include* these elements in theirs. And these scientific considerations are quite *independent* from our views on the *morality* of legal distinctions based on social status.

Similar issues arise while considering the *diversity* of societal attitudes towards specific elements. EDT researchers usually operate in a noisy environment,¹⁷² striving therefore to eliminate as much noise as possible. Even if some element's noise-generating potential is small, they lose effectively nothing by excluding that element. Yet, sometimes researchers must take into account the expected *amount* of noise and engage in complex cost-benefit analyses.

167. See *The Avalon Project: The Code of Hammurabi*, YALE.EDU, <http://avalon.law.yale.edu/ancient/hamframe.asp> (last visited Aug. 3, 2019) [<https://perma.cc/QQR3-M44E>].

168. *Id.*

169. *Id.*

170. For alternative interpretation of these laws as related to "cheek slapping" offenses against personal *honor*, see Martha T. Roth, *Mesopotamian Legal Traditions and the Laws of Hammurabi*, 71 CHI.-KENT L. REV. 13, 27–36 (1995).

171. Claims about an ancient society's outlook are almost always uncertain. See *id.* at 25. The rigidly stratified nature of ancient societies might suggest that the Code of Hammurabi represented the common outlook. See also Kathryn E. Slanski, *The Law of Hammurabi and its Audience*, 24 YALE J.L. & HUMAN. 97, 103 (2012). Yet, it might have *departed* from that outlook. For present purposes, this is immaterial: these laws are used to *illustrate* that ancient society *might* have had a radically different outlook on social status (and laws using it) than we do.

172. See *supra* Section II.C.2.

For instance, consider a society in which most people regard social status as a crucial—indeed, *essential*—element for distributing criminal liability and punishment (similarly to the hypothetical Babylonians in my example). Nevertheless, a small minority in that society maintains that social status should have *no* import on these matters (similarly to us and on same moral grounds). A dilemma arises. EDT researchers crafting scenarios for a representative sample of subjects in that society must decide between competing options. They can exclude social status from scenarios and risk creating a substantial source of experimental noise resulting from the majority of subjects “reading in” different social statuses into the scenarios. Conversely, they can avoid *this* noise by including social status, yet risk creating *another* (but probably less substantial) source of noise resulting from the minority of subjects’ responses to such inclusion (and their divergence from the majority’s responses). The researchers may conclude that the second option reduces *overall* experimental noise better.¹⁷³

A different concern (rooted in the contingent nature of scientific considerations) can be illustrated by redirecting our imagination from the past towards the future. So far we examined how adoption of *specific* scientific methodology safeguards EDT from relying on immoral intuitions (and theorized instances when it might fail).¹⁷⁴ Much of the discussion revolved around a rather technical detail: noise in the experiment. We considered the noisy environment in which EDT operates, the noise-generating properties of elements, and ways to reduce experimental noise, taking for granted the current state of social science.¹⁷⁵

Yet, science evolves. Innovative use of existing techniques or new methods altogether could be brought to bear on the problem of experimental noise and impact EDT methodology. After all, EDT adopted its current methodology as a *means* to avoid “poor testing methods.”¹⁷⁶ EDT is not devoted to this *specific* methodology. Rather, it is dedicated to doing good *science*.

Now, consider a possible future development: scientists develop a novel statistical test allowing researchers to account for noise in their experiments (whatever its source), *post-hoc*.¹⁷⁷ Obviously, using this

173. The researcher’s dilemma arises because the majority of subjects regard social status as an *essential* element (thus, its exclusion would generate noise, inviting subjects to read it in). This dilemma would not arise if social status were just an element *influencing* subjects’ responses (as it would now, or even in the Old South). If it were, researchers could have *preempted* the noise-generation problem in both options by simply excluding this element.

174. *See supra* Section II.B.

175. *See supra* Section II.B.

176. Robinson, *supra* note 15, at 33–34.

177. The ways a statistical test “accounts for” experimental noise depend on the test’s precise features, with some ways possibly more problematic than others. Consider a scenario including the perpetrator’s race: any method of accounting for the noise that ultimately incorporates the influence of race on subjects’ responses into the

test would appeal to future EDT researchers. A click of a button in a statistical program may free them from much of the tedious, pre-experimental work involved in carefully crafting scenarios and conducting multiple costly and labor-intensive manipulation checks. Moreover, using this test would be entirely consistent with EDT's dedication to doing good science, indeed, even *better* science. Nevertheless, using this test may also reintroduce into EDT scenario elements eliciting immoral intuitions (e.g., race, religion, social status), resurrecting immorality objections.

Note, the aforementioned examples' value rests not in their historical accuracy or realism but in their ability to illustrate a larger point. No preordained relation exists between scientific considerations and immorality. While currently, scientific considerations safeguard EDT, providing reasons to exclude scenario elements eliciting immoral intuitions, it is a *contingent* matter. Depending on societal features or methodological developments, the very same scientific considerations may provide reasons to *include* such elements. This point also enables us to grasp a more profound matter, to which we now turn.

2. Unsatisfying Nature of Scientific Reasons

There is something deeply *unsatisfactory* in EDT's reliance on *scientific* methodology as its bulwark against *immorality*. While current EDT methodology effectively safeguards it from relying on immoral intuitions, the *reasons* we deem it satisfactory have little (if anything) to do with that methodology's scientific rationales. To see why, consider the following insightful critique of certain *utilitarian* arguments:

[U]tilitarians proffer elaborate arguments to show the instrumental value of punishing the guilty and only them; but tacitly or even explicitly, the real proof of the correctness of their arguments is that *the results are consistent with considerations of desert*, which is enough to undermine the conclusion the arguments are intended to support.¹⁷⁸

This critique's insight directly applies to the practical successes of EDT—a utilitarian theory that achieves the right *results* yet for the wrong *reasons*. We find it appealing that EDT's *scientific* methodology typically excludes immoral intuitions. Yet these results are appealing because they are consistent with considerations of *moral* (not scien-

results raises substantial moral concerns. Conversely, a method assessing the influence of race on each subject and *excluding* that precise effect from the results seems less problematic, as the results will be *identical* to those of an experiment originally unaffected by race. Yet, even in this case, initially inviting experimental subjects to rely on elements eliciting immoral intuitions *is* problematic. I will return to this issue in *infra* Section IV.B.

178. Bronsteen, *supra* note 12, at 1143 n.84 (emphasis added) (citing Lloyd L. Weinreb, *Desert, Punishment, and Criminal Responsibility*, 49 LAW & CONTEMP. PROBS. 47, 47 (1986)).

tific) nature; *these* (moral) reasons make them satisfactory. Put differently, EDT methodology enables us to avoid relying on immoral intuitions (i.e., to achieve the right results) due to scientific, morally *inert* rationales related to technical matters of noise reduction (i.e., for the wrong *reasons*).

Moreover, this predicament is more ubiquitous than the risk that EDT's methodology might not work in certain historical or futuristic situations.¹⁷⁹ It runs deeper. Even when EDT's methodology works, as it does presently, it works because of technical, morally inert, and thus unsatisfying reasons.

Is this a serious problem for a *utilitarian* theory? After all, certain utilitarian theories could respond with the “so what?” shrug. As long as a methodology maximizes the relevant utility (e.g., deterrence, incapacitation), it will serve, other moral reasons notwithstanding. Yet, such response is less available to EDT because it aims to maximize *moral credibility* and, therefore, is particularly sensitive to moral reasons. Operating for the wrong reasons may inhibit the “normative crime control” mechanisms EDT aims to activate.¹⁸⁰

Where does this leave us? We have seen that EDT's scientific methodology has much merit as a sophisticated and powerful tool that in addition to its scientific qualities is typically capable—somewhat surprisingly—to moot immorality objections. Alas, it may succeed for the wrong reasons. This should *not* be surprising. After all, EDT's rationale is “doing good science,” not “avoiding immoral intuitions,” and these different rationales can conflict. The typical success of EDT methodology to avoid relying on immoral intuitions is both unexpected and remarkable. Yet, the contingent and morally unsatisfying nature of the scientific considerations on which EDT depends and the serious conceptual limitations they entail may cause concern.

We thus reach the rather familiar impasse in criminal law theory. Yet another utilitarian theory (EDT) and its distributive principle (empirical desert), may lead to morally-problematic results in certain circumstances.¹⁸¹ Immorality objections therefore reemerge, though in

179. See *supra* Section II.D.1.

180. This point is crucially important. See discussion *infra* Sections III.A.1.c., IV.B.

181. Note, being on an *equal* footing with other utilitarian theories in this familiar impasse is already an extraordinary achievement for EDT. Even before, EDT has arguably held significant advantages over its utilitarian counterparts. See, e.g., ROBINSON, *supra* note 10, at 141–52 (describing the limitations of deterrence, incapacitation, and rehabilitation as distributive principles). Yet, the ostensible contrast between the *extreme* examples (e.g., punishing the innocent) making immorality objections evident for other utilitarian theories and these objections' *common* intuitive appeal in EDT's case (due to EDT's purportedly pervasive reliance on biased intuitions), appeared to *offset* EDT's substantial utilitarian merits, putting it at a *disadvantage* vis-à-vis its utilitarian counterparts. Yet, we now see that such disadvantage was merely illusory. As with other utilitarian theories, *extreme* examples are required to make immorality objections evident for EDT. See *supra* Sections II.C.3 and II.D. Thus, a *direct* comparison between EDT and its utilitarian counterparts is possible. Moreover, because

different shape. The familiar—for utilitarian theories, at least—trade-off between morality and utility presents itself, requiring us to compromise on either the former or the latter. Or so it appears.

Fortunately, there is a way out of this impasse. It requires a reconceptualization of EDT and involves incorporating into its scientific methodology a minimalistic *normative* commitment to equality and non-discrimination (what I will refer to as the “Proposal”). My Proposal facilitates an extraordinary “win-win” situation. It enables EDT to steer clear from immoral intuitions; retain the strengths of its scientific methodology; enhance its capacity to maximize the criminal system’s moral credibility; and strengthen and augment its responses to a plethora of additional specific criticisms directed at EDT over the years. The next Section presents this Proposal.

III. RECONCEPTUALIZING EMPIRICAL DESERT THEORY

A. *Reconceptualizing Empirical Desert: Doing the Right Thing for the Right Reasons*

The Proposal to reconceptualize EDT contains two intricately connected key aspects. The first introduces a *normative* commitment to EDT: the *commitment to equality before the law and non-discrimination*. The second incorporates this normative commitment *into* EDT’s *methodology*. Complementing these key aspects is an overarching *minimalistic* spirit, guiding the Proposal to hew close to EDT and minimize deviation from it, thereby leading to *reconceptualization*, rather than *transformation*, of the theory. The Proposal’s key aspects, coupled with its minimalism, offer an *amalgamation* of—rather than a (false) choice between—its desiderata: steering clear from immoral intuitions and enhancing moral credibility while doing good science.¹⁸² This amalgamation enables EDT reconceptualized according to the Proposal (“reconceptualized EDT”)¹⁸³ to break out of the familiar utilitarian impasse.

EDT’s utilitarian merits are *not* offset during such comparison they carry more weight on balance, which in turn bolsters the argument that EDT has the comparative *advantage* over its utilitarian counterparts.

182. Naturally, relations between these desiderata are complex. See *infra* Section III.A.1.c.ii.

183. Note, below I must differentiate between *reconceptualized* EDT and EDT as it *presently* stands. I will refer to the latter as *standard* EDT. I selected the “standard” label due to its descriptive and, more importantly, relatively *neutral* nature. Note, the term “standard” may invoke two specific connotations I emphatically eschew. One is overly positive, equating “standard” with “universally accepted,” “prevailing,” “time-tested,” etc. Since EDT is a relatively new and contested theory, referring to it as standard in the aforementioned senses may appear as a semantic trick—a (misguided) attempt to imbue the theory with that veneer of respectability we sometimes grant to matters of old vintage. This is *not* my intent. EDT has enough substantive virtues; it requires no false façades. The other, overly negative connotation of “standard” makes it synonymous to “unremarkable,” “orthodox,” or “obsolete.” Obviously, invoking *such* connotation is not my intent either. EDT is a modern theory offering original

Below I detail the Proposal and discuss its merits.¹⁸⁴ I start by clarifying the meaning of the commitment to equality before the law and non-discrimination for our purposes. Next, I specify the procedure for incorporating this normative commitment *into* EDT's scientific methodology. Subsequently, I present the Proposal's main benefits, addressing both the different *reasons* it allows researchers to rely on as well as the different (potentially immoral) *elements* it enables researchers to notice (and exclude). Then, I address a fundamental objection to the Proposal, which may be raised by EDT proponents. This objection revolves around the idea that amalgamation of normative considerations and empirical desert theory is almost oxymoronic and raises doubts about the compatibility of reconceptualized and standard EDT. I debunk this objection, demonstrating that both EDT versions *substantively* agree on the application of normative considerations, and that despite *procedural* differences between them, reconceptualized and standard EDT are perfectly compatible.

1. The Proposal: Incorporating Normative Commitment to Equality and Non-Discrimination into Empirical Desert Theory's Methodology

- a. *What Does Commitment to Equality and Non-Discrimination Mean?*

The *idea* behind introducing a normative commitment to EDT is straightforward. Here is the problem: EDT's successes in eluding immoral intuitions are mere byproducts of its scientific methodology. This roundabout method may occasionally fail and is unsatisfactory. The solution: using a *normative* safeguard *directly* addressing EDT's reliance on immoral intuitions.¹⁸⁵

arguments and innovative insights. Moreover, my purpose in reconceptualizing "standard" EDT is *constructive*—aiming to rely upon and reinforce EDT. Thus, invoking a disparaging connotation to it would be odd and counter-productive. Ultimately, although I distinguish between "standard" and "reconceptualized" EDT to explain the unique features of each, these distinctions must not mask a fundamental fact: "Reconceptualized" EDT is *not* a separate theory—it is a variation on, or a version of, "standard" EDT.

184. Note, the term "Proposal" below refers to the proposed reconceptualization of EDT and its details, and "reconceptualized EDT" is simply standard EDT reconceptualized according to the Proposal.

185. For an earlier, brief, and very interesting argument suggesting that a solution to this problem may be found by embedding empirical desert *within* the philosophical approach of political liberalism, see Matthew Lister, *Desert: Empirical, not Metaphysical*, in CRIMINAL LAW CONVERSATIONS, *supra* note 2, at 51–53. I believe that despite the theoretical differences between this suggestion and the Proposal, they are somewhat alike, at least in spirit. Unfortunately, space constraints do not permit exploring the potential connections between them in detail.

The normative safeguard that can achieve this goal is a commitment to *equality before the law and non-discrimination*,¹⁸⁶ for reasons which will be explicated below.¹⁸⁷ First we must answer a preliminary question: What does the phrase “commitment to equality before the law and non-discrimination” actually mean for our purposes? *Prima facie*, this phrase sounds quite nebulous and may raise concerns because of its focus on *equality*—a complex, multifaceted, and actively disputed concept.¹⁸⁸ Does the proposed commitment rely on some idiosyncratic, or vigorously contested account of equality? Or, perhaps, does it mean nothing *specific*, strategically employing vague general terms, open to divergent and inconsistent interpretations?

These are valid concerns. It would be unfortunate if by attempting to solve an existing problem (reliance on immoral intuitions), we would introduce new ones into EDT (contested accounts of equality or interpretive indeterminacy). The Proposal, however, does nothing of that sort.

The meaning of the proposed commitment to equality before the law and non-discrimination, in keeping with the Proposal’s general spirit, is intentionally *minimalistic*, limited to widely accepted and uncontroversial understanding of its plain words. Most generally, it represents the simple recognition that law should treat people *similarly*, not *differently*, even when people differ from each other in virtue of certain *characteristics*. Importantly, the Proposal eschews subscribing to idiosyncratic or vigorously contested views about what these characteristics are. It only recognizes characteristics enjoying wide contemporary consensus, such as race or religion.¹⁸⁹ Besides, the Proposal

186. Note, I refer to this commitment throughout using the label *normative*, not *deontological*. The latter label may be confusing, because its use in the concept “*deontological desert*” is disputed in EDT literature. See *supra* note 87. The proposed commitment is certainly *different* from a commitment to deontological desert. Therefore, (and for additional reasons explained in *infra* Section IV.C.1.), I prefer the label *normative*. Note also that below I often contrast the terms “utilitarian” and “*normative*”: by the latter I always mean “*non-utilitarian normative*.”

187. See *infra* Section III.A.1.c.

188. For concise presentation of multiple perspectives on equality and related debates, see generally Stefan Gosepath, *Equality*, STAN. ENCYCLOPEDIA OF PHIL., <https://plato.stanford.edu/archives/spr2011/entries/equality> (last modified June 27, 2007) [<https://perma.cc/ED8R-WRTN>].

189. Note, these characteristics are viewed as fundamental at least in liberal democratic societies. For instance, John Rawls maintained that:

[R]eligious toleration is now accepted, and arguments for persecution are no longer openly professed; similarly, slavery is rejected as inherently unjust, and however much the aftermath of slavery may persist in social practices and unavowed attitudes, no one is willing to defend it.

John Rawls, *Justice as Fairness: Political Not Metaphysical*, 14 PHIL. PUB. AFF. 223, 228 (1985). Similarly, and more recently, commentators have noted that:

This fundamental idea of equal respect for all persons and of the equal worth or equal dignity of all human beings is accepted as a *minimal* standard by all leading schools of modern Western political and moral culture. Any political theory abandoning this notion of equality will not be found plausible today.

suggests adopting the normative commitment which the *community itself* supports.¹⁹⁰

Thus stated, the normative commitment's meaning leaves little space for conflicting interpretations about which characteristics cannot serve as a basis for treating people differently.¹⁹¹ This space is even further constrained by the legal context in which the Proposal operates. Legal systems typically address equality, providing specific answers as to which characteristics cannot serve a legal basis for differential treatment. They often contain explicit equalitarian prohibitions on differential treatment due to specific characteristics such as race or religion.¹⁹² It is the specific characteristics that the legal system deems discriminatory—rather than some idiosyncratic views as to what these characteristics ought to be—that serve the basis of the characteristics recognized by the Proposal.¹⁹³

Putting the point concretely in EDT context, the Proposal recognizes that basing criminal liability on intuitions influenced by characteristics such as race or religion is inconsistent with a normative commitment to equality before the law and non-discrimination. And the chief practical implication of this recognition for reconceptualized EDT is that it would exclude from its experimental scenarios, ele-

Gosepath, *supra* note 188 (section 2.3) (internal citations omitted) (emphasis added).

190. Note, this suggestion is not a *sine-qua-non* part of the Proposal. In *theory*, the Proposal can operate in a society that does not support a normative commitment to equality and non-discrimination. In *practice*, most contemporary societies *do*, and this support entails substantial benefits (though I postpone discussing them until *infra* Section IV). Presently, it is also useful to clarify that this suggestion focuses on a community's support of the normative commitment in *principle*, not on perfect practical compliance. A community can be regarded as supporting a commitment to equality before the law and non-discrimination, even if some of its *practices* are inconsistent with such commitment (as long as these inconsistencies, if pointed out, will be clear to its members). The Proposal's suggestion *is* fulfilled in most post-Enlightenment (and arguably many pre-Enlightenment) societies, including contemporary ones, since (discriminatory practices notwithstanding), *in principle*, these societies support a normative commitment to equality before the law and non-discrimination.

191. Some difficulties regarding particular characteristics exist. See *infra* Section III.A.1.c.ii, for a discussion of the difficulties regarding particular characteristics and the ways to resolve them.

192. Such prohibitions can be expressed in constitutional norms, such as (1) the Equal Protection Clause, see U.S. CONST. amend. XIV, § 1 (“nor shall any State . . . deny to any person within its jurisdiction the equal protection of the laws.”); (2) international legal norms, see U.N. Charter art. 1, ¶ 3 (“The Purposes of the United Nations are: . . . 3. To achieve international cooperation in . . . promoting and encouraging respect for human rights and for fundamental freedoms for all without distinction as to race, sex, language, or religion”); and (3) ordinary domestic legislation. The larger point beyond these particular examples is simple: although no consensus exists on every philosophical or legal interpretation of equality, there *is* a consensus about certain specific characteristics such as race or religion. The proposed normative commitment refers to *these* characteristics.

193. Such characteristics can include race, religion, skin color, national origin, language, etc. Below I refer to race and religion as archetypes, but my analysis applies to additional characteristics as well.

ments inconsistent with this commitment, including, for instance, elements implying race or religion.

Familiarity with the specific meaning of the commitment to “equality and non-discrimination,” which the Proposal introduces into EDT, improves one’s grasp on the Proposal’s first key aspect.¹⁹⁴ To understand it fully, one must turn to the Proposal’s second key aspect—incorporating this commitment *into* EDT.

*b. Incorporating the Normative Commitment Into
the Scientific Methodology*

The Proposal’s second key aspect is incorporating the normative commitment to equality and non-discrimination *into* the methodology of EDT.¹⁹⁵ EDT’s methodology involves multiple phases. In a typical experiment, researchers must, *inter alia*, decide which criminal law question to examine, select a suitable experimental design, create stimuli, recruit participants, conduct “manipulation checks,” run the main experiments, analyze results, and draw conclusions. At which phase is it proposed to incorporate the normative commitment?

In keeping with its general spirit of minimalism, the Proposal aims at *one* discrete methodological phase: the stimuli creation phase and, more precisely, the stage of crafting experimental scenarios. No changes are proposed to any other methodological phase.

How would incorporating a normative commitment during experimental scenario crafting work in practice? The answer is straightforward. According to the Proposal, the normative commitment should come into play *early*. While researchers craft their scenarios’ language yet temporally *prior* to their present consideration of whether a particular element in the scenarios’ language is noise-generating, they would be required to conduct an *additional* inquiry. Researchers would have to consider whether their scenarios’ language includes elements inconsistent with a commitment to equality and non-discrimination. If such elements are included, researchers should exclude them from scenarios whenever feasible (similarly to their present exclusion of noise-inducing elements).

This *modus operandi* requires some further unpacking on three fronts. The first refers to a substantive issue: *What question*, exactly,

194. Below, I refer to the commitment to “equality and non-discrimination,” and omit the words “before the law” for two reasons. First, their meaning is sufficiently captured by the term “non-discrimination.” Second, when “equality” is followed by “before the law” its meaning is explicitly restricted. The Proposal, however, interacts with other, *non-legal* facets of equality. (E.g., the Proposal treats intuitions of *all* experimental subjects as *equal*, abstaining from “sanitizing” some, but not other, intuitions post-experiment. See *infra* Section IV.B). While these facets of equality are entirely consistent with, and do not change the meaning of, the commitment to “equality *before the law*,” the unrestricted term “equality” captures them better.

195. The commitment’s incorporation into the *existing* methodological structure reflects the Proposal’s minimalistic spirit, aiming to minimize deviation from EDT.

should researchers ask themselves? The second refers to the temporal stipulation: Why should researchers ask this question *early*? The third refers to the matter of exclusion: Why should elements just be *excluded*, not included?

The substantive issue is complex. Researchers considering the relation between a specific scenario element and the commitment to equality and non-discrimination are required to take into account a *normative* question, not an *empirical one*. The question they must ask is *not* the question standard EDT researchers ask presently: “Would including this element influence subjects’ intuitions in a manner generating undesirable experimental *noise*?” Nor is it the more general question: “How would including this element *influence* people’s intuitions?”¹⁹⁶ The answers to these questions directly relate to empirical facts about people’s intuitions.¹⁹⁷

Rather, the Proposal requires researchers to ask: “*Is this element consistent with the commitment to equality and non-discrimination?*” The answer to *this* question is not directly related to people’s intuitions about particular elements, but to normative considerations of what is consistent with that commitment.¹⁹⁸

The Proposal’s temporal stipulation is not complex. It simply demands asking and answering the aforementioned normative question *prior* to conducting empirical inquiries. By functioning this way, the Proposal requires researchers to focus on the *normative* question *first*, and at an *early stage* of the process.

The exclusion matter is also simple. The Proposal is limited to elements’ *exclusion* due to its minimalistic spirit. First, the Proposal aims to minimize deviations from EDT, and elements’ *exclusion* (yet not *inclusion*) is an integral part of EDT methodology. Second, and relatedly, the Proposal’s focus on excluding elements addresses reasonable concerns that adding elements may increase the potential for conflict with EDT’s *scientific* considerations (e.g., adding elements to scenarios might generate additional experimental noise). Finally, the Proposal’s suggestion to exclude elements for *normative reasons* is minimalistic in another sense: it entails nothing radical. In fact, such exclusion is common in multiple criminal law related contexts (e.g., rules of evidence, or jury instructions) and for reasons related to the

196. These questions are distinct and may generate different answers for scenario elements that *influence* people’s intuitions without causing undesirable *noise*. See discussion *infra* Section III.A.1.c.ii for more on this issue.

197. In rare circumstances, these questions may resurface later on. See *infra* Section III.A.1.c.ii.

198. I suspect that the normative commitment to equality and non-discrimination is not entirely orthogonal to people’s *intuitions*. While I discuss closely related issues in *infra* Section III.A.1.c.ii, developing this point comprehensively is well beyond this Article’s scope.

proposed normative commitment (i.e., reducing inappropriate biases).¹⁹⁹

The Proposal's mode of operation, while minimalistic, triggers crucial alterations in the *reasons* for excluding certain elements from experimental scenarios and the *nature* of the excluded elements. These alterations, which we now turn to explore, entail substantial benefits and are instrumental in achieving the Proposal's goals to avoid relying on immoral intuitions, without giving up on EDT's scientific strengths.

c. *Benefits of Incorporation: Relying on Different Reasons, Excluding Different Elements*

i. Continuity and Contrast: Relying on Different Reasons

The Proposal's *modus operandi* requires researchers to start by focusing directly on *normative* considerations of equality and non-discrimination before asking empirical questions on people's intuitions. This organization of normative and empirical inquiries simultaneously facilitates continuity with—and contrast from—standard EDT. The continuity is manifested by the fact that experimental scenarios resulting from reconceptualized EDT's normative inquiries and standard EDT's empirical inquiries are typically *identical*. Identity notwithstanding, there is contrast: Reconceptualized EDT scenarios are influenced by different (normative) *reasons*. The causes for both continuity and contrast and their benefits will be considered in turn.

Because the Proposal incorporates the commitment to equality and non-discrimination and defines it narrowly, the vast majority of scenario elements presently passing standard EDT empirical inquiry's scrutiny would also pass the test of reconceptualized EDT's normative inquiry. For instance, the normative inquiry would not recommend excluding vital scenario elements (e.g., mental states, actions, means used, or harms caused) nor incidental yet useful scenario elements (e.g., general background, time of day, location) because these elements are *consistent* with the normative commitment to equality and non-discrimination.²⁰⁰

Crucially, the recommendations of the normative and the empirical inquiries would commonly match even for scenario elements obvi-

199. See, e.g., FED. R. EVID. 403 (“The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: *unfair prejudice*. . . .”) (emphasis added); see also BARBARA E. BERGMAN ET AL., 4 WHARTON'S CRIMINAL EVIDENCE § 27:1 (15th ed. 1999 & Supp. 2016) (citing the federal rule and asserting: “The military and 34 states have rules that are essentially the same as this federal rule.”).

200. For examples of scenarios containing incidental elements see Robinson, Goodwin, & Reisig, *supra* note 1, at 207–28. Note, researchers must examine *all* scenario elements for consistency with the normative commitment; none are *exempt* from the normative inquiry. Yet, the resulting recommendations for the aforementioned elements are easy and will match the empirical inquiry's recommendations.

ously inconsistent with the normative commitment (e.g., race, religion). The underlying reasons will differ. The normative inquiry would recommend excluding these elements due to their inconsistency with the normative commitment; the empirical inquiry would recommend excluding them due to scientific noise-reduction considerations. Yet, regardless of whether a scenario element (e.g., “Jewishness”) is deemed normatively inconsistent or empirically noise-inducing, the resulting recommendation (for exclusion) would be the same.

This similarity between the normative and empirical inquiries’ results, regarding both elements that will and will *not* be excluded, accords with the Proposal’s minimalistic spirit of minimizing deviation from EDT, whenever possible. Moreover, the typical identity between the scenarios crafted according to the recommendations of standard and reconceptualized EDT beneficially creates a substantial overlap between them.

The simple *practical* effect of the typical similarity in recommendations is that reconceptualized EDT researchers could typically *preserve* the language of the sophisticated, meticulously crafted, and experimentally tested standard EDT scenarios and utilize these established scenarios in future studies. For instance, reconceptualized EDT researchers could design studies using novel combinations of such previously developed scenarios or profitably employ them in conjunction with new ones.²⁰¹

More importantly, this practical effect also has profound theoretical implications. It means that reconceptualized EDT can rely heavily on the extensive body of research conducted by standard EDT researchers over the years and on their findings. In cases of overlap, we can already, at present, speak of estimated “findings” of not-yet-conducted reconceptualized EDT “experiments” and draw conclusions from them. These “findings” would be essentially identical to the concrete findings of already-completed EDT experiments. Put differently: in cases of overlap, we can regard past EDT experiments as if they followed reconceptualized EDT methodology.²⁰² Subsequently, we can speak of and draw conclusions from actual findings of past experiments as if they were the findings of reconceptualized EDT experiments.

This general point is vital. The overarching purpose of the Proposal (i.e., reconceptualization of EDT) is constructive. Reconceptualized EDT aims to rely on standard EDT foundations. The practical and

201. See, e.g., *id.* at 1961–75 (explaining that researchers used subjects’ responses to typical EDT scenarios as “milestones” for subsequent study, contrasting them with responses to vignettes presenting the facts of real cases in which modern crime-control doctrines were applied).

202. Note, while reliance on past EDT research is possible (and beneficial) in cases of overlap, reconceptualized empirical desert theorists should carefully examine past experimental scenarios, ensuring they contain no elements inconsistent with considerations of equality and non-discrimination.

theoretical capability of reconceptualized EDT researchers to rely on standard EDT researchers' extensive body of work furthers this goal. Indeed, this capability is the key benefit of the continuity between standard and reconceptualized EDT.

The *contrast* between the recommendations of standard and reconceptualized EDT is similarly beneficial. Even if identical elements are excluded from scenarios, they are excluded for different *reasons*.²⁰³ According to the Proposal, elements are excluded because of a normative commitment to equality and non-discrimination; not as a by-product of scientific methodology for noise-reduction. This contrast in reasons for exclusion is beneficial because occasionally *reasons* matter.

Recall, for instance, the limitations of standard EDT's "doing good science" rationale, specifically the unsatisfactory nature of reliance on contingent scientific reasons.²⁰⁴ The limitation was that while doing the "right things" (by regularly excluding potentially immoral intuitions), standard EDT was doing it for the "wrong reasons" (i.e., experimental noise considerations). Reconceptualized EDT overcomes this limitation: By providing recommendations similar to those of standard EDT, it will continue doing the "right thing." Yet, it will be doing it *not* for "wrong" reasons rooted only in the empirics of people's intuitions, but for "right" reasons rooted (also) in *normative* considerations.

Importantly, these normative considerations are directly germane to immorality objections. Superficially, it might seem that reconceptualized EDT's normative commitment was selected arbitrarily. After all, if one ventures to introduce a normative commitment to EDT, why choose this specific one? Yet, the connection is straightforward. Recall immorality objections' most persuasive examples. They pertain to *inequality* and the *discriminatory* attitudes of the Old South and Nazi Germany. The Proposal's normative commitment to equality and non-discrimination copes with these examples head on.

Contrasting the proposed normative commitment with an alternative may elucidate this point from a different angle. Consider, for example, the option of incorporating into EDT a commitment to *liberty*. Putting aside practical complications inherent to such incorporation,²⁰⁵ the crucial difficulty is this: incorporating the commitment to

203. Note a similar difference in reasons concerning identical elements that will *not* be excluded according to both standard and reconceptualized EDT. Though below I focus on benefits pertaining to the manifest act of elements' *exclusion*, the argument applies, *mutatis mutandis*, to the less visible act of *non-exclusion*.

204. See *infra* Section II.D.2. Recall also the importance of the *intention* to do justice for gaining moral credibility reputation. See *supra* Section II.A.

205. I hedge the possibility of incorporating "liberty" into EDT. I believe it involves inherent (though perhaps not insurmountable) difficulties. Liberty is obviously an important value, protected, *inter alia*, by criminal law's general principles. Nevertheless, by focusing on substantive criminal law provisions (and corresponding EDT

liberty will fail to address the main problem requiring solution. The concerns underlying immorality objections are not directly pertinent to liberty. Yet they are directly related to equality and non-discrimination.²⁰⁶ Thus, the *proposed* commitment also enables reconceptualized EDT to do the right thing for the right reasons among normative ones (i.e., for reasons of equality and non-discrimination).

Additionally, the temporal sequence of considering normative questions *prior* to scientific ones ensures that these *normative* reasons come into play first, activating the exclusion mechanism. This further emphasizes the normative commitment's *primacy*, highlighting that reconceptualized EDT excludes identical elements—that could have been excluded for scientific reasons (e.g., race or religion)—for normative reasons.²⁰⁷

scenarios) we focus precisely on the conduct which may justify society in imposing limitations on individual liberty. In this context, I find incorporating liberty into EDT's methodology difficult to fathom. While the option should not be ignored, I suspect considerations of liberty may function better as *external* constraints on EDT's recommendations. See *infra* Section III.B; Cahill, *supra* note 98, at 43–44.

206. Could other or additional normative commitments be incorporated into EDT? This possibility should not be discounted, but should be viewed with a degree of apprehension, especially in light of the Proposal's general spirit of minimalism preferring to make only *absolutely* necessary changes to standard EDT. Incorporating even a single unsuitable normative commitment (not to mention multiple), may derail EDT's *constructive reconceptualization* and cause *destructive transformation*. For instance, consider the option of incorporating a normative commitment to *deontological* desert into EDT. Initially, such incorporation even appears aligned with the Proposal's minimalistic spirit, due to similarities between deontological and empirical desert. See ROBINSON, *supra* note 10, at 173–74. Yet, this incorporation raises considerable problems. First, it introduces into EDT indeterminacies inherent in deontological desert, as deontological desert theorists are divided on multiple issues, including whether resulting *harm* should matter for desert. See Robinson, *supra* note 23, at 1832–33. Second, such incorporation may be patently counterproductive. Assume that the “correct” deontological position is that harm is *irrelevant* to desert. This position is at odds with robust empirical findings that lay people consider harm as highly *relevant*. See ROBINSON, *supra* note 10, at 90 (“People overwhelmingly have a strong intuition that resulting harm matters . . . documented across demographics and cultures.”). Disregarding these findings and incorporating the “correct” deontological position into EDT will *directly* conflict with people's intuitions of justice, fail to harness normative crime control mechanisms, and may generate the adverse effects on moral credibility that EDT was designed to avoid in the first place. In short, incorporating a commitment to deontological desert into EDT may create a *transformed* theory regrettably incapable of capitalizing on EDT's strengths. Despite these challenges, future research may discover normative commitments that can replace or, perhaps more plausibly, operate alongside the commitment to equality and non-discrimination. For practical purposes, however, the latter commitment most directly pertains to immorality objections and is organically suited for incorporation into EDT methodology.

207. Ostensibly, requiring researchers to consider an additional normative question increases the amount of work they must put into scenario crafting. Curiously, this is not necessarily so. The new requirement might *ease* researchers' burden. Presently, researchers must investigate empirically whether elements are noise-inducing (e.g., by surveying relevant social-science literature or conducting experiments). Occasionally, the proposed *normative* inquiry (preceding the empirical), may indicate straightaway that elements are inconsistent with a commitment to equality and non-discrimination,

To summarize, two kinds of benefits stem from the Proposal's incorporation of the normative commitment to equality and non-discrimination into standard EDT. Typically, the recommendations for and against excluding scenario elements according to the Proposal will be similar to those of standard EDT. The benefit of this substantial overlap is that adopting the Proposal requires no alterations in standard EDT scenario language. Moreover, it allows speaking of, and drawing conclusions from, the extensive research already conducted by standard EDT researchers, as if it were conducted according to the Proposal. However, despite the identity between elements excluded from scenarios according to standard and reconceptualized EDT, these elements are excluded for different reasons. The benefits of excluding these elements for *normative* reasons lie in overcoming the unsatisfactory nature of reliance on contingent scientific reasons and using the appropriate normative commitment (at the right time)—allowing reconceptualized EDT to do the right thing for the right reasons.

ii. Contrast and Continuity: Excluding Different Elements

The Proposal's *modus operandi* entails another vital benefit concerning the *different nature* of the elements it enables researchers to notice and exclude. Typically, standard and reconceptualized EDT's recommendations will be similar; but requiring researchers to focus on the *normative* question during scenario crafting casts a different net than focusing on the *scientific*, noise-oriented question would. By switching their focus to the normative question, researchers may discover potentially problematic elements hitherto ignored by the noise-oriented inquiry.

Recall that the existence and amount of noise generated by particular scenario elements depends upon the (contingent) features of the society at hand.²⁰⁸ Consequently, researchers' ability to discover whether certain elements are noise-inducing depends on the diversity and distribution of intuitions regarding such elements in the society. The difficulty arising from this dependence is that elements tapping into pervasive and uniformly distributed societal biases might influence intuitions without inducing easily discoverable noise. By focusing on the normative question, however, researchers may uncover such elements' problematic nature by and large independently from the state of diversity and distribution of intuitions that happen to be elicited by such elements in a given society.

Discovering elements that will be retained by the scientific inquiry but excluded by the normative one is challenging due to the substantial overlap between them. Nevertheless, rare cases when the norma-

and obviate the need for empirical inquiry, reducing researchers' overall efforts. Other times, the opposite may be true. See *infra* Section III.A.1.c.ii.

208. See *supra* Section II.D.1.

tive and scientific inquiries may yield different recommendations do exist. Recall the scenarios in which “John” committed potentially criminal acts. As is standard for meticulously crafted EDT scenarios, these scenarios contained *no* elements specifying John’s race or religion. Yet, they strongly suggested John’s *gender*.²⁰⁹

Clearly, whether indicating perpetrators’ gender in scenarios *has* noise-inducing (or other) influence on experimental subjects’ judgments is a purely *empirical* question, contingent upon features of the relevant society.²¹⁰ In contrast, whether perpetrators’ gender *should* be relevant to criminal liability and punishment is a different, *normative* question.

The question standard EDT researchers must ask is whether including elements indicating perpetrators’ gender in scenarios will generate noise. The answer hinges on the empirical state of affairs within the society researchers examine.²¹¹ And if researchers do not expect or observe experimental noise from using gender-related elements in their scenarios, they will have no reason to exclude them.²¹² Con-

209. What exactly experimental scenarios suggest by using proper names like “John” and third-person pronouns like “he” can be disputed. Is it gender, sex, or both? This dispute is immaterial for the analysis below. I will use “gender” because it captures nicely some of the *societal* aspects inherent in that analysis. The analysis below would remain unchanged if I used the other term(s).

210. See Denno, *supra* note 3, at 755–56 (suggesting perpetrators’ gender in EDT scenarios might have had such influence).

211. Note, whether a scenario element *influences* intuitions or generates experimental *noise* are different questions. If there is a pervasive and uniform bias in a society, an element tapping into that bias may *influence* subjects (causing them to adjust their responses similarly) yet generate *no noise*. Consider the following illustrations. In the contemporary U.S., it will be obviously controversial for scenarios to classify perpetrators as Jewish or black explicitly, or imply so by naming them “Moses” or “Draymond” (for anecdotal evidence that the latter name is highly suggestive of skin color, see Dave Chappelle, *Dave Chappelle: Equanimity — Draymond Green Clip — Netflix is a Joke*, YOUTUBE (Jan. 4, 2018), <https://www.youtube.com/watch?v=-1VhigA7o88>.). Scenarios mentioning these classifications or names will likely lead to *differential* judgments (driven by subjects’ stereotypes or, conversely, attempts to compensate for them) and will both *influence* subjects and generate experimental *noise*. Scientific considerations will recommend excluding such elements. Consider now a society in which stating or implying perpetrators’ *gender* in scenarios is similarly controversial. In *this* society, invoking perpetrators’ gender will also *influence* subjects and generate experimental *noise*. Scientific considerations will, again, recommend exclusion. Finally, consider a society in which *all* are uniformly and consistently more *lenient* towards one gender. See, e.g., Denno, *supra* note 3, at 756 (referring to research suggesting that offenses committed by males may be judged more *leniently*). In this society, the gender element *influences* subjects (per hypothesis). However, this *influence* is uniform and consistent; thus, it will not necessarily generate *noise*. Scientific considerations alone would *not* suggest exclusion in this society. Yet, even if indicating gender is not noise-inducing (or if results can be “sanitized” from noise) and no *scientific* reason exists to exclude gender from scenarios, there might still be *normative* reasons to do so.

212. This may explain why standard EDT scenarios include elements indicating gender (e.g., “John”), though there may be *additional* reasons, which I will address shortly below.

versely, the *normative* question which researchers must ask according to the Proposal is whether including elements indicating perpetrators' gender is consistent with a *commitment to equality and non-discrimination*. The answer to *this* question will not hinge on the empirical distribution of intuitions that gender invokes in the society at hand but rather on normative grounds.

In this particular case, the normative question also appears to yield a different answer from the empirical one. Indicating perpetrators' gender (just as indicating race or religion, and for the same reasons) is inconsistent with a commitment to equality and non-discrimination. Therefore, elements indicating perpetrators' gender should be *excluded* from (reconceptualized) EDT scenarios—on normative grounds.²¹³

Thus, in rare cases, the Proposal will recommend excluding an element such as gender from experimental scenarios for normative reasons, even if it generates no experimental noise. This initial recommendation is important, and it should be followed whenever feasible. Yet, excluding normatively problematic elements may still involve thorny issues related to empirical questions that must be considered.

For instance, occasionally excluding perpetrators' gender from a research scenario is empirically *impossible*. The limitation may be linguistic. While English can use gender-neutral scenarios, other languages may simply lack gender-neutral options. Alternatively, sometimes using gender-neutral phrasing (regardless of language) may introduce ambiguity into scenarios, causing some (but not other) subjects to perceive perpetrators' gender differently, and generate experimental noise.²¹⁴

Moreover, it may be that using gender-neutral options in EDT scenarios will *necessarily* invoke a stereotypical gender.²¹⁵ Consider again the example opening this Article, in which the perpetrator shoots a woman in a dark alley. Can you recall—and if not, guess—what the perpetrator's gender was in that example? If I were to tell you the perpetrator was a woman, would that surprise you?

Actually, explicit reference to the perpetrator's gender was intentionally omitted. Nevertheless, the context may have implied that the perpetrator was a man.²¹⁶

213. While the main text focuses on *perpetrators'* gender, *victims'* gender may also influence subjects' judgments without causing experimental noise and should be excluded on normative grounds.

214. Recall, EDT aspires to avoid ambiguities. See *supra* Section II.C.1 & Section II.C.2.

215. For a survey of research on stereotypical features' activation when non-stereotypical features are not indicated, see Fischer & Engelhardt, *supra* note 126, at 412–13.

216. In other contexts, stereotype activation may lead to opposite conclusions. For instance, research indicates that upon hearing the phrase: "the secretary greeted

These complications illustrate that while asking the normative question may occasionally ease the burden on researchers by obviating the need to consider thorny empirical questions, this may *not* always be the case. For instance, researchers might be required to compare between gender-neutral scenarios and scenarios explicitly utilizing different genders. If the results are identical, the Proposal would recommend using gender-neutral scenarios (for normative reasons). If researchers do find differences between gender-neutral and gendered scenarios, additional matters must be considered.²¹⁷

In any event, the key benefit of the Proposal's new focus does *not* depend on whether asking the normative question reduces or increases researchers' burden. The key benefit is that asking the normative question increases the methodology's capabilities and sophistication by enabling researchers to *notice* elements that might have been entirely overlooked if they were limited only to the scientific question.²¹⁸

This matter is fundamental. It also allows us to reflect on the present discussion from a slightly different perspective, resonating with our discussion of immorality objections. Excluding a specific element from a scenario due to scientific considerations of experimental noise hinges on the effect that element has on the intuitions of community members. That is, it depends on the existing state of affairs in that community. Conversely, excluding a specific element from a scenario due to its inconsistency with a normative commitment to equality and non-discrimination depends on certain normative standards of the community. That is, it depends less on the empirical state of affairs as it is with regard to intuitions. It depends more on the normative state of affairs that the community aspires to and on that community's (philosophical) conception of what a commitment to equality and non-discrimination entails. This difference in underlying reasons accounts for the different nature of the elements the Proposal helps un-

Jane," people "leap to the conclusion that Jane was greeted by a woman," although the secretary's gender is unspecified. *See id.* at 412. Thus, researchers must be mindful to elements (or interaction between them) that may indirectly indicate gender and influence subjects.

217. In more complicated circumstances, when researchers find substantial differences between gender-neutral and gendered scenarios (or when gender-neutral options are unavailable), the normative considerations will require them to consider methods to counter-balance or account for the biases gender may cause. Such methods may vary, and some (specifically those accounting for elements' bias *ex-post*) might be partially subject to criticisms similar to those I will present in *infra* Sections III.B., and IV.B. Finding a definitive solution to these questions is beyond this Article's scope.

218. Besides, the relationship between normative and scientific considerations may be mutually reinforcing. Normative considerations may heighten researchers' attention to different scenario elements (e.g., gender), which may in turn raise awareness to previously unnoticed *scientific* questions (e.g., activation of default stereotypes), thereby enabling researchers to do even *better* science.

cover—elements that may be discerned only if we focus not only on the empirical, but on the normative question.²¹⁹

B. *Perfect Compatibility Between Standard and Reconceptualized Empirical Desert Theory*

A discussion of one fundamental objection to the Proposal to reconceptualize EDT was deferred to this point. This objection revolves around the idea that the proposed amalgamation of *normative* considerations and *empirical* desert theory is almost oxymoronic. The Proposal's normative foundations are incompatible with EDT's empirical foundations. Fusing them together in one "mixed" theory is precisely the wrong way to go.²²⁰

This fundamental objection could be raised not by EDT's opponents but rather by its *proponents*. The objection focuses on the nature of the appropriate connection between EDT and various normative commitments. One might surmise, for instance, that EDT proponents may be opposed in principle to applying any kind of normative commitment in conjunction with their utilitarian theory. Alternatively, empirical desert theory proponents might be specifically opposed to applying a normative commitment not directly related to desert but, rather, to equality and non-discrimination.

Yet, both conjectures would be mistaken. First, EDT does not *oppose* applying normative commitments in conjunction with it. Rather, it explicitly welcomes them. For instance, in discussing why real cases in general—and publicly known cases specifically—are inadequate for EDT's purposes, Robinson, EDT's founder, explains: "People's views on such cases are commonly *skewed* by political or social context or by other factors, such as *race*, that *all would agree have no proper role in setting principles of justice*."²²¹ Obviously, the statement above refers to some normative standard of agreement *outside* the confines of EDT itself.

219. Note, I do not claim that the Proposal provides a "transcendent" solution to overcome immorality objections in any society or time period. Nevertheless, and though it is ultimately an empirical matter, I think that a methodology relying on the contemporary state of a community's *normative aspirations* can notice the problematic nature of elements temporally *earlier* than a methodology relying on the contemporary state of a community's *intuitions*.

220. Presently, any potential objections offering alternatives to the Proposal are hypothetical since this Article introduces the Proposal for the first time. So, the argument below is *preemptive*.

221. ROBINSON, *supra* note 10, at 164 (emphasis added). In a similar vein, while conducting a study on extralegal punishment factors ("XPF"), Robinson mentioned: "[W]hile some decision makers might in fact be influenced by an offender's race or physical attractiveness, *no one seriously argues that such practices should be followed*. The goal here is in providing guidance to reformers, so only those XPFs that have at least some *colorable claim to legitimacy* were considered." *Id.* at 495 (emphasis added).

What about the alternative hypothesis that the normative standard to which Robinson refers is exclusively the standard of *deontological* desert? That would be mistaken too. First, recall that Robinson has considerable reservations regarding *deontological* desert's ability to provide a valuable "transcendent check" on empirical desert.²²² Second, and more importantly, similarly to *any* desert-oriented theory, EDT is not "monomaniacal" about desert.²²³ This is all the more true for *EDT* because it focuses not on desert itself but, rather, on the utilitarian crime-control benefits associated with it.²²⁴ EDT openly recognizes that other important values exist and occasionally these values might be more important than empirical desert.

Robinson, for instance, clearly stated: "[I]t is simply *not* the case that crime control is the *only* or even the *most* important societal interest" and stressed that considerations like privacy, fairness, and others may outweigh it.²²⁵ This was neither a solitary nor a vague statement. Time and again, Robinson stated that specific important societal interests (such as the legality principle, for instance) may introduce constraints that will conflict with EDT.²²⁶ Yet, EDT fully accepts such limitations, recognizing that distribution of liability and punishment according to empirical desert is just part of a larger scheme of crime-control mechanisms operating within and outside of criminal law.²²⁷

Importantly, EDT does not view these external standards as some unwelcome obstructions of justice. These standards are not merely acknowledged but also valued and even occasionally employed to appraise EDT itself. Perhaps the best example of such appraisal—which is also most closely related to the normative commitment to equality

222. See *supra* Section II.B.1.

223. See Michael T. Cahill, *Retributive Justice in the Real World*, 85 WASH. U. L. REV. 815, 841 n.78 (2007) (citing MICHAEL MOORE, *PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW* 172 (1997) ("The retributivist like anyone else can admit that there are other intrinsic goods, such as the goods protected by the rights to life, liberty, and bodily integrity. The retributivist can also admit that sometimes some of these rights will trump the achieving of retributive justice. . . .")); *id.* at 186 ("[H]ere again we need to remind ourselves that retributivists are not monomaniacal about the achieving of retributive justice . . . there are other intrinsic goods besides giving culpable wrongdoers their due and sometimes these other goods override the achievement of retributive justice.").

224. See *supra* Section II.A.

225. ROBINSON, *supra* note 10, at 189 (emphasis added). Robinson goes on to say:

Of course, articulating beforehand exactly which of these interests will trump doing justice and its crime-control benefits is an enormous project. It is a significant portion of the agenda of moral philosophy, constitutional theory, and political theory. The point here is simply to acknowledge that those debates are indeed relevant to the distribution of criminal liability and punishment.

Id. at 196. Cf. Robinson, *supra* note 5, at 1110–11.

226. See, e.g., Robinson, *supra* note 15, at 38–39.

227. Robinson, *supra* note 5, at 1110–11.

and non-discrimination adopted by the Proposal—can be found in the following statement:

Indeed, empirical desert provides a *unique opportunity* for developing *color-blind* principles of justice. If the empirical study subjects are not told the race of the offender during testing, even subconscious biases can have no effect. Thus, sentencing guidelines based upon such principles, for example, can be *color-blind*.²²⁸

This statement is significant not for acknowledging that EDT's scenarios would exclude factors such as race. That is a commonplace by-product of EDT's scientific methodology. The crucial point is that "developing color blind-principles of justice" is viewed by EDT as a *positive* outcome, indeed a "unique opportunity."²²⁹ And this outcome's appraisal as positive is done by EDT in accordance with some *external normative* standard.

These examples show that EDT proponents are neither opposed to applying normative commitments in conjunction with EDT, nor to applying normative commitments *not* directly related to deontological desert or desert in general. With these objections debunked, it is easy to see that standard EDT proponents are likely to recognize that both standard EDT and reconceptualized EDT are perfectly compatible. Such compatibility should not be at all surprising in light of reconceptualized EDT's constructive purpose and its intentional efforts to achieve such compatibility, manifested by its general spirit of minimalism and its specific efforts to deviate as *little as possible* from standard EDT.²³⁰

There is, of course, an unmistakable difference between these versions of EDT concerning the nature of their appropriate connection to normative commitments. Yet, importantly, the source of difference is not a *substantive* matter—it is a matter of *methodology*. Both versions of EDT agree that normative considerations can and *should* be applied. The real difference between the two versions of EDT is in the *procedure* used to apply normative considerations.

Standard EDT calls for applying normative considerations as *external* side-constraints. That is, after standard EDT's empirical studies are completed and corresponding recommendations are finalized, a separate stage, external to the theory itself, begins. During *that* stage all the pertinent *normative* considerations are considered (for instance

228. Robinson, *supra* note 18, at 64 (emphasis added). This specific statement was part of a response to EDT's criticism the details of which are not pertinent to present discussion. What *is* pertinent is that the statement's content is entirely consistent with the pattern of interaction between EDT and other, *external* normative considerations, according to which the theory's color-blindness is appraised.

229. *See id.*

230. See, e.g., discussions on the commitment's narrow meaning in *supra* Section III.A.1.a.; the incorporation of the commitment *into* standard EDT's existing methodology in *supra* Section III.A.1.b.; or the substantial overlap in the recommendations of both versions of EDT in *supra* Section III.A.1.c.i.

by criminal law theorists or philosophers), and standard EDT's recommendations may be set aside, if need be. Reconceptualized EDT concurs in general. Yet, it calls for an *earlier* consideration of *one* specific normative commitment and its incorporation into EDT's methodology.

Standard EDT's view of normative considerations as external constraints on its recommendations is rationally sound. EDT is a *utilitarian* theory aiming to provide crime-control recommendations based on scientific methodology designed to assess laymen intuitions. It is not in the business of weighing normative considerations and, therefore, is not methodologically equipped to do so. EDT is content having normative questions considered externally, by others, recognizing that consequently its own recommendations may be set aside.²³¹

Indeed, using the external constraint approach can work well in many (probably most) cases. For instance, it makes good sense to have EDT crime-control recommendations finalized and then consider whether they are consistent with a normative commitment such as a commitment to *privacy*.²³² Yet difficulties arise when the intuitions (on which EDT's recommendations are based) and the normative issues under consideration are more closely connected. This happens to be the case with the normative commitment to equality and non-discrimination, as the following example illustrates.

Imagine philosophers reviewing EDT's crime-control recommendations and considering whether they are consistent with the normative commitment to equality and non-discrimination. Assume further that these philosophers suspect that the intuitions on which EDT's recommendations are based have been "tainted" by elements inconsistent with equality and non-discrimination. How could these philosophers apply the external constraint approach to ameliorate the situation?²³³

One option would be to take a "categorical" stance, prohibiting *any* reliance on crime-control recommendations based on empirical data suspected to be "tainted." This solution, however, comes at a heavy cost. Since the "categorical" stance is applied *after* EDT experiments have been conducted and rejects the recommendations based on these experiments, it results in a substantial waste of information and resources.

An alternative option would be to take a more moderate "sanitizing" stance, endeavoring to partially retain the crime-control rec-

231. See ROBINSON, *supra* note 10, at 189, 196; Cf. Robinson, *supra* note 5 at 1110–11 (noting that the issue of punishment distribution EDT focuses on, important as it is, may be *not* the most important one even for the narrow purposes of crime-control, which may be affected by other important factors within and outside of the criminal justice system).

232. For a similar suggestion see ROBINSON, *supra* note 10, at 189.

233. Note, as we have already discussed, using deontological desert as a "transcendent check" on the intuitions EDT relies on is not an available option. See *supra* Section II.B.1.

ommendations by sanitizing the potentially “tainted” data from intuitions inconsistent with a commitment to equality and non-discrimination. Yet, this solution, too, involves serious difficulties. For our present purposes, suffice it to say that it depends on philosophers’ ability to prove they are equipped with a procedure capable of “sanitizing” the “tainted” data consistently and reliably. This is a heavy burden. In fact, decades of scientific research on pervasive and persistent influences of bias on perception and judgments suggest that once introduced, biases are extremely difficult to eliminate.²³⁴

To be sure, the aforementioned difficulties primarily challenge the philosophers (or criminal law theorists) wishing to align EDT’s recommendations with the normative commitment to equality and non-discrimination. Being external to EDT, these are not challenges *of* EDT. Nevertheless, they present challenges *for* EDT. The more difficult it becomes for experts to align EDT’s crime-control recommendations with the normative commitment to equality and non-discrimination, the less attractive they may find relying on EDT’s recommendations²³⁵ or taking them as a starting point for deliberation.

These are real challenges, which *reconceptualized* EDT helps overcome by incorporating the normative commitment to equality and non-discrimination *into* the existing methodology of standard EDT.²³⁶ By operating as early as it does—excluding elements inconsistent with a commitment to equality and non-discrimination from experimental scenarios during the scenario crafting stage—the Proposal effectively de-biases subjects’ intuitions *ex-ante*.²³⁷ Thus, it addresses the challenges presented by the external constraint approach, avoids the waste of information and resources of the “categorical” stance, and preempts the difficulties of the *post-hoc* sanitizing efforts of the “sanitizing” stance. By resolving *these* challenges internally, reconceptualized EDT makes relying on its recommendations more attractive.

234. For a brief overview of studies casting serious doubts on the possibility to eliminate biasing influences in the *legal* context see, e.g., Thomas Nadelhoffer, *Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Juror Impartiality*, 9 PHIL. EXPLORATIONS 203, 212–14 (2006). For a review of numerous studies demonstrating *philosophers’* susceptibility to various biases, see generally Stephen Stich & Kevin Tobia, *Intuition and Its Critics*, in THE ROUTLEDGE COMPANION TO THOUGHT EXPERIMENTS (Michael T. Stuart, Yiftach Fehige, & James Robert Brown eds., 2017). For a study suggesting that factors such as extensive training, reflection, or considering alternatives fail to mitigate certain biases’ influence on philosophers (compared to similarly educated non-philosophers), see generally Eric Schwitzgebel & Fiery Cushman, *Philosophers’ Biased Judgments Persist Despite Training, Expertise and Reflection*, 141 COGNITION 127 (2015).

235. For a similar argument maintaining the limitation of deontological desert in spotting unjust intuitions is a mixed blessing, making *EDT* less appealing, see *supra* Section II.B.1.

236. In a way, it can be said that the Proposal uses an *internal* normative commitment approach, rather than using an *external* normative constraint approach.

237. For a similar claim about the de-biasing effect of excluding scenario elements in the context of standard EDT, see Robinson, *supra* note 18, at 64.

Having said that, it is important to see this difference between the two versions of EDT for what it really is. It is not rooted in some principled objection of standard EDT to applying normative considerations. Nor is it a result of some diametrically opposite position taken on this matter by reconceptualized EDT. The two versions of EDT simply rely on different procedures for taking into account one particular, narrowly-defined normative commitment.²³⁸ In all other respects, standard and reconceptualized EDT are very similar and perfectly compatible.

The conceptual difference regarding normative commitments between both versions of EDT is small. Nevertheless, as we have seen throughout this Section, it has many implications, and incorporating the normative commitment to equality and non-discrimination into EDT according to the Proposal entails substantial benefits. So far, we mostly focused on these benefits as they pertain to immorality objections. Yet advantages of reconceptualized EDT extend much further.²³⁹ The next Section demonstrates that reconceptualized EDT can strengthen and augment standard EDT responses to a plethora of criticisms directed at the theory from its inception.

Before closing this Section, however, it is worth our while to reflect on one valuable feature of reconceptualized EDT: it is truly “color-blind” in a very broad sense of the term. As standard EDT, it is blind to race, religion, and any other noise-inducing elements. But it is also blind to gender, social status, and additional elements inconsistent with a commitment to equality and non-discrimination. And it is blind to *all* of the aforementioned elements *because* of its commitment to equality and non-discrimination, hearkening back to fundamental conceptions of impartial justice. Figuratively speaking, reconceptualized EDT adds a normative layer on top of the (scientific) blindfold placed by standard EDT on Femida’s eyes, further ensuring that morally objectionable elements of a case will not impact criminal liability and punishment.

238. Importantly, reconceptualized EDT’s spirit of minimalism also steers it clear from being bogged down by vague moral (or moralistic) considerations, by restricting the elements that would be excluded for normative reasons. Recall, the meaning of the normative commitment to equality and non-discrimination is limited. The Proposal eschews subscribing to idiosyncratic or highly contested views about characteristics that are inconsistent with a commitment to equality and non-discrimination. It only recognizes characteristics enjoying wide contemporary consensus, such as race, religion, etc. *See generally supra* Section III.A.1.a.

239. One of such advantages is related to potential connections between reconceptualized EDT, and the novel criminal theory of reconstructivism. *See generally*, Joshua Kleinfeld, *Reconstructivism: The Place of Criminal Law in Ethical Life*, 129 HARV. L. REV. 1485 (2015). Space constraints do not permit exploring these connections here, yet I wish to acknowledge that Professor Kleinfeld’s writing on reconstructivism has contributed to my thinking on reconceptualized EDT.

IV. STRENGTHENING EMPIRICAL DESERT THEORY

Now that we have a sound understanding of the Proposal to reconceptualize EDT, and its advantages regarding immorality objections are clearer, we are in a good position to examine its additional benefits. Over the years, standard EDT was subjected to a barrage of criticisms. As mentioned in the introduction, commentators criticized it for manifold reasons, asserting, *inter alia*, that EDT: relies on the distorted intuitions of the public instead of deferring to the better informed judgments of criminal law experts; “cherry picks” the intuitions it uses—potentially failing to reap utilitarian compliance benefits or incongruously (for a utilitarian theory) relying on non-utilitarian justifications; overlooks, in a self-defeating manner, the gap between the intuitions it relies on and the intuitions the public relies on in the real world, discounting the implications this gap has on the theory’s efficacy; and violates important moral principles such as the “publicity condition,” deceives community members, and exploits their retributive intuitions.²⁴⁰

Yet, these very criticisms provide a perfect vehicle for showcasing reconceptualized EDT’s benefits. The exploration below will reveal that some of these apparently separate claims are interconnected and have grown from the same roots that gave rise to immorality objections. More importantly, it will demonstrate that reconceptualized EDT can either augment standard EDT responses to these criticisms or overcome them entirely.

Two short comments on the presentation style in this Section are in order. First, it groups the major strands of EDT critiques in its individual subsections. The discrete critiques within each subsection (numbered and italicized) are interspersed into the argument to emphasize their interconnectedness and streamline the presentation.

Second, distinct critiques were raised by different scholars at different times. Some critiques’ meanings are not always perfectly clear, others are difficult to comprehend outside their original context, and others, still, are phrased as neutral observations or even praises of EDT rather than as its critiques. To provide the best test cases for the Proposal, I will attempt to distill these critiques’ essence, provide necessary context, and phrase them as presenting the best case *against* EDT.

240. See *supra* notes 28–32 and accompanying text.

A. *Intuitions are Biased, and Empirical Desert Theory is Cherry Picking?*

1. Community's Intuitions are Biased, and We Should Defer to Experts

Claims that community intuitions are biased underlie immorality objections to EDT. These claims are periodically followed by recommendations to defer, instead, to the better judgments of professional elite²⁴¹ or experts.²⁴² As we have seen, reconceptualized EDT addresses immorality objections by incorporating a normative commitment into standard EDT's methodology and requiring *researchers* to think from a different perspective about which elements they should exclude from their scenarios. This further safeguards the intuitions of experiment participants and provides a more robust response to claims that standard EDT relies on (ostensibly) biased community intuitions. Interestingly, reconceptualized EDT's normative perspective, coupled with its attention to the respective roles of researchers and people participating in experiments, is also instrumental in addressing the related standard EDT's criticism holding that: *in criminal law matters, it is preferable to defer to experts* {1}.

This criticism assumes experts to be more knowledgeable than the public in considering factors relevant to criminal punishment.²⁴³ Experts' advantages could be manifested in various manners. For instance, as one commentator noted, if irrelevant factors influence public "attitudes about deserved punishment," then using "orderly, dispassionate conceptions of desert" might help mitigate this influence.²⁴⁴ Such conceptions could rely on carefully specified "factors a *decision maker* may *legitimately* consider when determining desert."²⁴⁵ Unlike members of the public, experts appear to be well-equipped to specify such factors.

Have experts been successful in this endeavor? This is debatable. For instance, the same commentator also argued that desert theorists were never able to specify such factors "in terms very much more specific than desert itself."²⁴⁶ The factors they specified, such as blameworthiness, crime seriousness, harm, culpability, wrongdoing, and

241. Roberts, *supra* note 28, at 106–07.

242. Tom R. Tyler, *Public Attitudes on Criminal Justice*, 7 CRIM. L. F. 697, 700 (1996).

243. See, e.g., Denno, *supra* note 3, at 743–44, 752–55; Roberts, *supra* note 28, at 107.

244. Ristroph, *supra* note 3, at 162.

245. *Id.* (emphasis added). Note, the commentator raising this possibility was *skeptical* about experts' potential success in such an endeavor, as will be further elaborated in the main text. Yet it is worthwhile to explore this possibility, because it highlights a particularly interesting manner in which experts' advantages *could* be manifested.

246. *Id.* at 162–63.

autonomy, are (claimed to be) cross referring, interchangeable, inexact, and elastic.²⁴⁷

For us, that debate is immaterial; the key point is that reconceptualized EDT approaches the matter from a different angle. Rather than attempting to specify the factors which *may* be legitimately considered in determining *desert*, it focuses on the factors which may *not* be legitimately considered in the criminal context, according to a commitment to *equality and non-discrimination*. Reconceptualized EDT's approach displays two advantages. First, the list of factors which may not be legitimately considered according to reconceptualized EDT's normative commitment is fairly well defined.²⁴⁸ Second, it is not focused on the factors which may be relied upon by the "*decision maker*."²⁴⁹ Rather, it aims directly at the factors which may *not* be legitimately relied upon by the *people* participating in the experiments, that is, by actual community members.

The second advantage is key. It highlights that reconceptualized EDT suggests a new "division of labor" between experts and community members, balancing two aspects of the argument for deferring to experts {*I*}. That argument must assume not only that experts are more *knowledgeable* than community members on crime and punishment matters but also that they can better *decide* on such matters.

The assumption's first (epistemic) part is correct in many respects.²⁵⁰ Experts clearly *do* enjoy certain advantages over community members, including better familiarity with their subject matter and better training in relevant research methodologies. However, the assumption's second (decision-making) part is highly questionable for multiple reasons.²⁵¹ Among them (and closely related to EDT) is the risk that decision-making by experts will lead to formation of elite notions of crime and punishment divorced from those of the community, which in turn would undermine the criminal law system's "normative crime-control" mechanisms.

Reconceptualized EDT offers a compromise allowing us to capitalize on the aforementioned advantages while mitigating the risks. It still asks experts to play a vital role in criminal law policy. Yet, according to reconceptualized EDT, the experts' role is not to *decide* on mat-

247. *Id.* at 163. The same commentator also examined other potential ways to rationalize desert, but concluded that it is likely that any desert determination will not be dispassionate but rather influenced by emotions. *Id.* at 162–64.

248. See *supra* Section III.A.1.a.

249. Ristroph, *supra* note 3, at 162.

250. This assumption is also corroborated by research. See Roberts, *supra* note 28, at 107–08.

251. For a brief summary of scholarly opinion *disagreeing* with this assumption, and discussing reasons related to non-technical, but rather *societal* aspects of punishment, see *id.* at 107. For a more comprehensive discussion evaluating the value of public opinions for reasons of legitimacy and compliance, the censuring function of criminal sanctions, and the societal elements in proportionality judgments, see *id.* at 110–14.

ters of crime and punishment *for* the community. It is to *facilitate* such decisions *by* the community. By using their expertise, experts can design, administer, and interpret the proper experiments for measuring community intuitions of justice. Moreover, these experiments will also exclude factors which, according to the normative commitments of the *community* itself, may not be legitimately considered. Thus, while recognizing experts' vital role and capitalizing on their expertise, reconceptualized EDT also gives primacy to both the intuitions and the normative commitments of the community itself.²⁵²

2. Empirical Desert Theory is “Cherry Picking” Among the Intuitions of the Community

Another criticism faulting the intuitions EDT utilizes, alleges that EDT is “cherry picking” intuitions.²⁵³ To unpack this criticism, it is important to distinguish between “predictive” and “justificatory” modes in assessing relevance of community intuitions for the law.²⁵⁴ In the *predictive* mode, knowing the content of community intuitions is important for *predicting* the community's reactions to laws and policies related to these intuitions.²⁵⁵ These intuitions themselves have no justificatory force.²⁵⁶ Yet a prediction based on these intuitions may enter our cost-benefit analysis when we consider laws and policies likely to implicate them (though we settled on these laws and policies for *other* reasons).²⁵⁷ Conversely, intuitions can be used for *justifying* moral claims. Yet, using intuitions in the *justificatory* mode is a much more difficult endeavor due to, *inter alia*, dissensus on whether and under which conditions intuitions can serve as justifications as well as on ways to resolve conflicts between intuitions.²⁵⁸

The cherry picking criticism rests on the claim that EDT, while operating in the predictive mode,²⁵⁹ chooses to focus only on intuitions elicited by scenarios drafted at an intermediate abstraction level and administered in a dispassionate experimental environment. Oddly, EDT ignores intuitions the community may have on broad issues such as the criminal system's general aims;²⁶⁰ ignores intuitions elicited by specific cases, such as that of O.J. Simpson, which may be laden with racial, political, or social biases;²⁶¹ and ignores heat-of-the-moment in-

252. See *supra* note 190. This point will be significantly expanded upon in *infra* Section IV.B.

253. Kolber, *supra* note 4, at 441–48.

254. See *id.* at 435–37.

255. *Id.* at 435.

256. *Id.*

257. *Id.*

258. *Id.* at 436–37.

259. *Id.* at 441–42.

260. *Id.* at 442.

261. *Id.* at 443–44.

tutions.²⁶² EDT's choice to limit itself to such intuitions leads to the cherry picking criticism's two-fold charge: EDT is either *deficient from a consequentialist standpoint* {2} or *illicitly slips into the justificatory mode* {3}.²⁶³

The "consequentialist deficiency" charge {2} is predicated on the claim that if EDT operates in the predictive mode and seeks to increase compliance with the law, it should *defend* its choice to use the intuitions it uses.²⁶⁴ It can do so by demonstrating that *these* (and not other) intuitions are the most useful to promote compliance.²⁶⁵ This is an empirical claim that must be proved—yet EDT proponents have not provided a sustained defense of such a claim.²⁶⁶ Instead, EDT inexplicably ignores heat-of-the-moment, racially, politically, or otherwise biased intuitions that may have a significant effect on compliance, potentially missing valuable crime-control opportunities.²⁶⁷

The "illicit slip" charge {3} has a different focus. It faults the *reasons* for EDT's selectivity reflected in its proponents' statements such as the following: "People's views on such [publicly known] cases are commonly skewed by political or social context or by other factors, such as race, that *all would agree have no proper role in setting principles of justice.*"²⁶⁸ The claim is that by excluding certain intuitions for *such* reasons, EDT undermines its consequentialist foundations.²⁶⁹ It no longer operates in the predictive mode, but seems to "*illicitly*" slip into the justificatory mode.²⁷⁰

EDT has a response to both charges. As for the "consequentialist deficiency" charge {2}, the reasons it chooses to focus on specific intuitions (but not others) are *empirical* and related to the theory's "doing good science" rationale. It excludes certain kinds of intuitions because they may generate experimental noise.²⁷¹ Such noise may in turn reduce the chances of discovering or usefully measuring the shared intuitions of justice concealed by the already noisy environment of disagreements on crime and punishment matters.²⁷² Thus, for instance, even if racially-biased intuitions are indeed relevant to compliance, a scenario including elements eliciting them might be of limited (if any) usefulness because of the noise such elements will generate.

262. *Id.* at 445–46.

263. *Id.* at 441–48.

264. *Id.* at 448.

265. *Id.* at 441.

266. *Id.* at 443–47.

267. *Id.* at 443–46.

268. *Id.* at 443–44. We have considered this statement in *supra* Section III.B.

269. Kolber, *supra* note 4, at 444–45.

270. *Id.* Other commentators similarly criticized EDT scenarios for being deliberately designed to ignore factors EDT acknowledges can shape desert judgments like "racial prejudice or political considerations." Alice Ristroph, *Third Wave Legal Moralism*, 42 ARIZ. ST. L.J. 1151, 1166 (2010).

271. *See supra* Section II.C.

272. This point is partially acknowledged in Kolber, *supra* note 4, at 443.

This response's essence is clear. Perhaps, *theoretically*, from a consequentialist standpoint, inability to account for certain kinds of intuitions may involve crime-control opportunity loss. Yet, *in practice*, including such intuitions in experimental scenarios is counter-productive. Bearing in mind empirical limitations in the real world, standard EDT *does the best it can* to discover and measure intuitions of justice.

Reconceptualized EDT, however, has different, *normative* reasons for excluding racially-biased intuitions (even if they are relevant for compliance). Reconceptualized EDT excludes them not because it operates in a noisy environment and is *limited* in what it *can* do. It excludes them because it views exclusion as something *it should do*. For reconceptualized EDT, failing to include within its ambit the potential compliance-relevant effects of these intuitions (like a good consequentialist presumably would wish to do) is not a critique at all, even theoretically. For reconceptualized EDT, such exclusion is not a missed opportunity, but a virtue. Compliance based on intuitions influenced by racial biases is not something reconceptualized EDT seeks to maximize—it is an undesirable feature it aims to eliminate.

Standard EDT's response to the "illicit slip" charge {3} seems to be more complicated. At first glance, its appeal to matters "such as race, that all would agree have no proper role in setting principles of justice" appears to be inconsistent with the theory's purely utilitarian stance.²⁷³ But this impression probably stems from what is ultimately a confusion. Standard EDT is utilitarian *internally*, yet it accepts *external* normative side constraints and welcomes their use when appropriate.²⁷⁴ There is nothing "illicit" in using them as far as EDT is concerned.

Moreover, even if relying on normative considerations internally and "slipping" into the justificatory mode is "illicit" for standard EDT, it is certainly *not* the case for reconceptualized EDT. On the contrary, this is reconceptualized EDT's typical *modus operandi*. There can be no confusion here. Reconceptualized EDT is specifically designed to use the normative commitment to equality and non-discrimination internally to exclude elements from experimental scenarios. The fact that reconceptualized EDT follows through cannot possibly serve as a *critique* that it does something "illicit." Rather, it is simply an accurate description of (and arguably a praise to) the way reconceptualized EDT is designed to function.

Thus, reconceptualized EDT simply moots the "illicit slip" charge. And it should come as no surprise. To return to the distinction presented at the outset, that critique is predicated on the notion that EDT operates mainly in the *predictive* mode and ought not slip into the *justificatory* mode. Whatever merit this claim may have for stan-

273. ROBINSON, *supra* note 10, at 164.

274. *See supra* Section III.B.

dard EDT, it is clearly inapposite for its reconceptualized version. As our discussion clarifies, reconceptualized EDT is *not* limited to the predictive mode. Rather—and this is key—it is intentionally designed to engage (also) in the *justificatory* mode.²⁷⁵

B. *Mind the Gap: Do Real-Life Intuitions Undermine Empirical Desert Theory?*

Another major line of EDT criticisms is related to the *gap* between the intuitions it relies on and the real-world intuitions of the public (“experimental” and “ordinary” intuitions, respectively). Underlying these criticisms is the observation that experimental intuitions elicited by a scenario of a crime will differ from ordinary intuitions elicited by a real case of crime (or media reports about it). This gap between the two kinds of intuitions presents several challenges for EDT, and each “gap-related”²⁷⁶ criticism below focuses on a different one.

One criticism is that it may be *self-defeating for EDT to rely on experimental intuitions due to the gap* {4}.²⁷⁷ The challenge is intrinsic. EDT seeks to increase the criminal system’s *moral credibility*, which in turn depends on the alignment between law and *ordinary* intuitions. Yet, EDT insists on using *experimental* (not ordinary) intuitions.²⁷⁸ The result, as some commentators noted, may be that: “the very concept of shared community intuitions of justice is not captured by Robinson’s scenario methodology,” and therefore “*no* beneficial effect in terms of *moral credibility* is to be expected from empirical desert.”²⁷⁹ In other words, EDT is hamstringing itself by relying on *experimental* intuitions, according to the critique. Due to the gap between experimental and ordinary intuitions, EDT’s reliance on the *former* will not increase the system’s moral credibility. Rather, it will undermine it.²⁸⁰

275. Reconceptualized EDT’s ability to engage in a justificatory debate is vastly important, and I will explore its implications in depth. See, e.g., *supra* Sections IV.B., IV.C.1., & IV.C.2.

276. Note, the term “gap” refers only to differences between ordinary and experimental intuitions. As the main text below explains, the gap is *related* to the alignment between the *law* and intuitions. Yet it is *not* the same.

277. See generally de Keijser, *supra* note 30, at 102, 107–08. For specific application of this critique to EDT see *id.* at 103, 109, 115–16; Roberts & de Keijser, *supra* note 4, at 490–91.

278. See, e.g., Roberts & de Keijser, *supra* note 4, at 490–91.

279. *Id.* at 493.

280. Similar criticisms focusing more narrowly on the gap between *ordinary* intuitions on *cardinal* (absolute) amounts of punishment and EDT’s expected recommendations on such punishment (which are based on a transformation of *experimental* intuitions’ *ordinal* rankings) were also made. See Kolber, *supra* note 4, at 449–50; Slobogin & Brinkley-Rubinstein, *supra* note 17, at 80, 89–90, 96. The text in this Section refers to the general “moral credibility” oriented form of the criticism, but is applicable to its specific “cardinal” version as well.

EDT has several responses to this criticism. The first points out that EDT's goal is *not* eliminating the gap between experimental and ordinary intuitions. It is rather to improve the criminal system's overall moral credibility by aligning the *law* with ordinary intuitions. And *this* goal can be furthered (even assuming a gap between ordinary and experimental intuitions) by recommending a rule that is *better* aligned with the ordinary intuitions than current law.²⁸¹ As EDT researchers observed:

Even if community views on an issue were expressed by a bell curve, rather than (say) a single point, the law's moral credibility could still be improved by moving from a point on the tail of the curve to a point nearer the center, since this would reduce the number of people with whose intuitions the law conflicts and the extent of that conflict.²⁸²

This response is further bolstered by the fact that experimental and ordinary intuitions are *correlated*, whereas current law may have been designed to promote utilitarian goals, such as deterrence or incapacitation that are largely *unrelated* to ordinary intuitions.²⁸³ Therefore, by relying on *experimental* intuitions (which *approximate* ordinary intuitions), EDT usually recommends rules *reducing* the divergence between current law and *ordinary* intuitions.

Yet, even if we align the law *closer* to experimental intuitions, *improving* the system's moral credibility, some gap between experimental and ordinary intuitions may still remain and lead to negative effects.²⁸⁴ A second set of responses helps EDT address this concern. It maintains that notwithstanding the gap, negative effects might not occur because people recognize that punishments embodied in law may differ from their judgments in specific cases. Perfect alignment is *not* required for sustaining the law's moral credibility. Generally, people will accept the criminal system's prescriptions if they "find the rank ordering of sentences to be *in accord* with their own rank orderings and the magnitudes of the sentences within a *latitude of acceptance* from their own sentences."²⁸⁵

In other words, experimental intuitions underlying EDT do not have to match ordinary intuitions precisely.²⁸⁶ It is sufficient if they

281. Robinson, *supra* note 15, at 34.

282. Robinson, Barton, & Lister, *supra* note 33, at 331. While the quoted paragraph refers to community disagreements on *ordinal* ranking, a similar point was made concerning disagreements on *cardinal* punishments as well. *See id.* at 336–37. *But see* Slobogin, *supra* note 33, at 393 (questioning this argument if the distribution is not normal but rather bi- or tri-modal).

283. *See, e.g.*, ROBINSON, *supra* note 10, at 173–74; *supra* Section II.A.

284. For a similar observation, *see* Kolber, *supra* note 4, at 450.

285. John M. Darley, *Morality in the Law: The Psychological Foundations of Citizens' Desires to Punish Transgressions*, 5 ANN. REV. L. & SOC. SCI. 1, 17 (2009) (emphasis added).

286. As Robinson stated:

are “close enough,” within some acceptable *range* to not be self-defeating. These responses are even further bolstered by the crucial impact the system’s *intentions* have on its moral credibility. As long as the system is genuinely trying to do justice, a certain extent of deviation from what every community member deems appropriate in any particular case might still be perceived as acceptable.²⁸⁷

In general, the aforementioned EDT responses undermine the gap-related critiques substantially. Yet, they might not suffice in some highly challenging cases when ordinary intuitions are substantially affected by passions arising in a highly polarizing context. Consider, for example, a widely publicized murder case highlighting the offender’s race. The racial element is likely to influence ordinary intuitions pertaining to the case. Yet, EDT’s recommendations on murder will be based on *experimental* intuitions elicited by scenarios in which racial elements are *absent*.

This example brings us to a further EDT critique. *In such (predictable) cases of a likely gap between experimental and ordinary intuitions, EDT provides limited guidance as to which set of intuitions should be adhered to {5}*.²⁸⁸ This critique applies in the specific cases in which a gap is likely to be widest and presumably present the most acute challenge for EDT. Its thrust is that in *these* cases, EDT is in a particularly difficult position. Should it counsel adhering to experimental intuitions and risk bringing about public discontent in the real world? Or vice versa?²⁸⁹

Indeed, from EDT’s utilitarian standpoint, such cases are the most challenging. These cases also enable us to see most clearly how *reconceptualized* EDT can contribute to standard EDT’s responses in overcoming these challenges. Reconceptualized EDT can bring its *normative* perspective to bear, providing a very different kind of response (which would apply also to criticism {4}).

Let us look at the challenges presented to standard EDT by the aforementioned case in which ordinary intuitions are substantially influenced by passions pertaining to a particular characteristic of the

Happily for the criminal law, minor subtle discrepancies between the legal rule and the community’s views are not likely to have a serious effect in undermining the law’s moral credibility. The law is seen as providing a rough approximation of what people see as just, and that may well be tolerated as adequate.

ROBINSON, *supra* note 10, at 111. Though this was stated while discussing alignment between community intuitions of justice and the law on attempts and accomplice liability, the statement’s logic is not limited to that discussion.

287. See *supra* Section II.A.

288. Slobogin, *supra* note 28, at 1195–96. Originally this critique was made in a slightly different and limited context, yet its rationale applies more generally.

289. *Id.* at 1195–96. Note, though this criticism closely resembles criticism {4} it addresses the most *challenging* cases and focuses less on the *results* of EDT’s reliance on experimental intuitions and more on EDT’s difficulty to provide *guidance* regarding *which* set of intuitions should be followed.

offender (such as race). Standard EDT recognizes that a gap between experimental and ordinary intuitions in this case may be inevitable. This gap is also undesirable. Since standard EDT is ultimately in the crime-control maximizing business, this case presents a real dilemma, and the aforementioned criticisms point to its different aspects. EDT's reliance on experimental intuitions in this case is self-defeating and may reduce the system's moral credibility, according to critique {4}. Moreover, it is *especially* unclear whether in this case EDT should recommend adhering to experimental intuitions (as it typically does) or follow ordinary intuitions, according to critique {5}.

Standard EDT has the conceptual tools to solve the dilemma even in such challenging cases. These tools are, naturally, strictly *utilitarian*. In principle, standard EDT is *indifferent* as to whether experimental or ordinary intuitions should be followed. Typically, for good scientific reasons, standard EDT recommends aligning the law with *experimental* intuitions (measured by carefully crafted scenarios that exclude “noise-inducing” elements such as race). Yet, standard EDT is open to deviation from its typical recommendations if such deviation may achieve a crime-control “bonanza,” as it may in the most challenging cases.²⁹⁰ Thus, if aligning the law with *ordinary* intuitions influenced by elements such as race in the case at hand achieves better crime-control, standard EDT should recommend it, at least in theory.²⁹¹

Yet, standard EDT's utilitarian solution of the dilemma encounters a familiar problem. In the case at hand, aligning the law with ordinary intuitions means aligning it with intuitions influenced by the *race* of the offender. This, in turn, resurrects (a version of) immorality objections.²⁹²

Reconceptualized EDT possesses different conceptual tools to address the dilemma. In contrast to standard EDT's utilitarian perspective—according to which neither experimental nor ordinary intuitions are normatively preferable—reconceptualized EDT's preference is

290. See Robinson, *supra* note 15, at 38; see also *supra* Section II.A.

291. Note, though EDT's preference for ordinary intuitions is *theoretically* possible, it is *highly* unlikely in practice. EDT's typical function is to guide *legislation*, and its recommendations depend on *overall* effects on moral credibility. Even in special, *atypical* cases following experimental intuitions may result in loss of moral credibility, laws deviating from them might be viewed by the public as too draconian or too lenient in the majority of *typical* cases. This may lead to an aggregated credibility *loss*, and the standard non-deviation from EDT's recommendations should be preferred—even if occasionally it might conflict with ordinary intuitions. For an earlier version of this basic argument, and more sustained discussion of the long-term effects of deviations from desert in a slightly different context, see, e.g., ROBINSON, *supra* note 10, at 191–95. Note, additionally, the exclusive focus in the main text on standard EDT's *initial* recommendation. Standard EDT clearly recognizes that maximizing moral credibility may give way to other important societal interests. See *supra* Sections II.A & III.B.

292. See *supra* Section II.B. While the “resurrected” version of immorality objections is theoretically concerning, practically it is limited. See *supra* note 291.

clear and unequivocal. Even in the most challenging cases for standard EDT, reconceptualized EDT recommends relying on *experimental* intuitions.

The reasons for this preference are, emphatically, not utilitarian but *normative*. While standard EDT regards the influence of race on ordinary intuitions as noise-inducing (and *therefore* excludes such intuitions from its scenarios), reconceptualized EDT regards such influence as inconsistent with the commitment to equality and non-discrimination. Reconceptualized EDT regards intuitions influenced by race not as merely noisy but as inappropriately *biased*. Therefore, it will recommend against aligning the law with such intuitions even in the (presumably rare) cases when, from a strictly utilitarian perspective, doing so might generate crime-control benefits.²⁹³

The normative perspective of reconceptualized EDT enables it to resolve the dilemma presented by the most challenging cases without reviving immorality objections. Furthermore, this normative perspective generates *additional* crime-control benefits. It allows reconceptualized EDT to reframe the gap-related criticisms not as a *challenge* for the criminal system's moral credibility, but as an *opportunity* to enhance it.

Recall again the aforementioned distinction between the descriptive and justificatory modes of operation. While standard EDT operates in the predictive mode, reconceptualized EDT is intentionally designed to operate (also) in the justificatory mode, which offers it valuable advantages.²⁹⁴ For instance, as mentioned above, standard EDT does not (and cannot) claim that experimental intuitions are normatively better than ordinary intuitions. Its reasons for typically not relying on ordinary intuitions are scientific and thus lack justificatory force. In contrast, reconceptualized EDT offers compelling *normative* reasons to justify its preference of experimental intuitions. It asserts that experimental intuitions *are* normatively better than ordinary intuitions because they are unbiased. Moreover, and importantly, according to reconceptualized EDT, experimental intuitions are not *unbiased* as an *accidental* byproduct of complex scientific techniques to reduce experimental "noise" (techniques the community is unlikely to care about). They are *intentionally de-biased* according to and because of a normative commitment to equality and non-discrimination, which the *community itself supports*.

This is more than a *theoretical* point. It can have important *practical* consequences. The justificatory force of reconceptualized EDT's rea-

293. One may wonder whether reconceptualized EDT sacrifices significant crime-control benefits in preferring experimental intuitions even in these cases. Not so. Practically, such cases are rare, and even standard EDT will almost always prefer following *experimental* (rather than ordinary) intuitions for scientific, non-normative reasons.

294. See *supra* Section IV.A.2.

soning may help to *reduce*, and perhaps even *eliminate*, the potential adverse effects of the gap between experimental and ordinary intuitions. This is because both standard and reconceptualized EDT are situated on the same side of the gap, but they prefer to be on opposite sides. Both rely on *experimental* (not ordinary) intuitions, yet for standard EDT this is a *concession*. For reconceptualized EDT it is a *success*.

For standard EDT the gap represents a challenge, because conceptually, from a strictly utilitarian standpoint, relying on *ordinary* intuitions may be preferable (though impractical, for various reasons). Thus, standard EDT settles for “doing the best it can” by relying on the second-best proxy for ordinary intuitions—experimental intuitions. In contrast, for reconceptualized EDT the gap does not represent the same challenge. From its normative standpoint, relying on ordinary intuitions is *not* preferable because they can be biased. Thus, reconceptualized EDT doesn’t compromise but relies on the normatively better option—experimental intuitions.

While the gap between ordinary and experimental intuitions is equally wide for standard and reconceptualized EDT, reconceptualized EDT frames it differently and provides a persuasive *explanation* for its preference to rely on *experimental* intuitions. This explanation is normatively compelling and thereby can constitute a *reason* for the community to accept and support reconceptualized EDT’s recommendations even when they differ from community members’ ordinary intuitions. This, in turn, offers reconceptualized EDT an opportunity to *enhance* the system’s moral credibility.

Significantly, this aspect of reconceptualized EDT also provides a clear response to another EDT criticism, which maintains that *relying on experimental intuitions might be perceived as offensive by the general public* {6}. According to that criticism, EDT researchers selectively rely on the intuitions of a carefully informed *subset* of the public to provide their recommendations. Yet, such selective reliance might not be perceived sympathetically by the *general public*, which may find researchers’ dismissal of its views offensive.²⁹⁵ The situation may deteriorate further if even *informed* public opinion is subsequently curtailed by other external side constraints (deriving from moral principles or experts’ views).²⁹⁶

Reconceptualized EDT’s response to this criticism is straightforward. Its reasons for relying on experimental (rather than ordinary) intuitions demonstrate utmost respect for the general public. Unlike some experimental methods (e.g., opinion polls), it refuses to first measure ordinary intuitions which it expects (or worse, causes) to be

295. de Keijser, *supra* note 30, at 108, 115.

296. *Id.* at 115.

biased²⁹⁷ and then—lamenting about the unprincipled and punitive public—apply external constraints in an attempt to “sanitize” or otherwise restrain these intuitions.²⁹⁸ Doing so would have presented a conceptual problem for a theory wishing to gain *moral credibility* with the community, as it may be perceived by lay people as disregard and contempt for their intuitions, resulting in a *negative* effect on the criminal system’s moral credibility.²⁹⁹ Reconceptualized EDT takes a different route, aiming to assess *actual* intuitions of the public while safeguarding them from being distorted by questionable methods or “sanitized” *post-hoc*. In this, it gives deference and respect to the actual intuitions people have (as evidenced by their actual judgments).

Importantly, reconceptualized EDT’s safeguarding function is *respectful* of the community in another way. Since its main criterion for such safeguarding (i.e., the commitment to equality and non-discrimination) is a normative standard the *community itself* subscribes to,³⁰⁰ it excludes the elements community members find inconsistent with *their own* normative commitments. There is nothing dismissive of or offensive to the community in such an approach. The opposite is true.³⁰¹

297. For a short survey of studies suggesting that “punitive public opinion” is an artifact of the methodology used to assess them, see, e.g., *id.* at 105.

298. Recall the discussion of the “sanitizing” stance of the external constraint approach in *supra* Section III.B. This Section presents additional, conceptual, difficulty with that approach.

299. Note, lay people may support the *results* of such a procedure, agreeing that morally problematic intuitions must not be relied upon but still find the *procedure* itself to be disrespectful, since it intentionally invites them to rely on morally problematic intuitions.

300. Recall also that for its recommendations on cardinal scales of punishment it uses the end-point *accepted by the community*. See, e.g., ROBINSON, *supra* note 10, at 10–11, 164, 166–67.

301. One may wonder whether de-biasing can go too far, and I am grateful to Professor Leo Katz for helping me realize that addressing this question here would be beneficial. In general, the methodology of reconceptualized EDT seems to invite even further scenario de-biasing. For instance, (elitist) experts might be tempted to de-bias scenarios according to their (aspirational) interpretations of the community’s normative commitments, rather than according to the *community’s actual* normative commitments, rendering the results of the process foreign to community members. This, in turn could raise the critique that even reconceptualized EDT is dismissive of the community. Reconceptualized EDT’s minimalism, however, minimizes this concern, by accurately specifying the normative commitment to be employed. See *supra* Sections III.A.1.a & III.A.1.c.i (discussing the limited and commonly accepted meaning of the commitment to equality and non-discrimination, and the reasons for choosing this particular commitment). Moreover, the theory’s minimalism guides the commitment’s interpretation. What matters for reconceptualized EDT is the *community’s* views on what a commitment to equality and non-discrimination *actually* entails rather than some *expert’s* idiosyncratic views of what it *should* entail. Illustrating this point might be useful. Consider for instance, the (plausible) moral arguments that offenses towards persons and animals (or certain kinds of animals) should be treated equally by the criminal law (for a brief overview of the spectrum of approaches towards the rights of animals, including approaches suggesting animals’ personhood, see generally, Cass R. Sunstein, *The Rights of Animals*, 70 U. CHI. L. REV. 387 (2003)). If this posi-

Finally, note that by operating in this manner, reconceptualized EDT's methodology is both aligned with standard EDT foundations in relying on actual intuitions *and* it embraces another facet of *equality* that echoes its *own* normative commitment to equality and non-discrimination. The latter point is relevant to the present discussion, yet its significance extends beyond it. The facet of equality embraced by reconceptualized EDT exists on a fundamental level. Reconceptualized EDT gives *equal* opportunity and assigns *equal* weight to the intuitions of *all* subjects participating in reconceptualized EDT experiments, including subjects who *would have* relied on distorted or biased intuitions. And it does so without first “entrapping” these (and only these) subjects to rely on distorted or biased intuitions to provide negatively prejudiced (or over-compensatory egalitarian) biased judgments and subsequently “sanitizing” these subjects’ intuitions. Rather, the Proposal’s methodology enables treating the actual judgments of *all* subjects as *equal* from the start.³⁰²

tion is accepted as part of the commitment to equality and non-discrimination, it would require the reconceptualized theory to exclude scenario elements indicating whether the victim of an offense is a person or an animal (perhaps by using the term “creature” in its scenarios, or excluding elements implying human personhood) and recommend, for instance, that killing a human being and an animal be punished equally. While such interpretation of the commitment to equality and non-discrimination may be suggested by some experts and even if their reasons are compelling, presently (at least in the U.S.), such interpretation of the commitment is unlikely to match most communities’ understanding of it. Thus, presently, reconceptualized EDT’s minimalism would *preclude* it from de-biasing its scenarios from elements indicating human personhood. Yet, present-day results of reconceptualized EDT’s minimalism are themselves contingent. It is entirely possible that in some contemporary or future communities, commitment to equality and non-discrimination will indeed encompass the moral equivalence of humans and animals (or certain kinds of animals) for criminal law purposes. For an interesting poll surveying contemporary approaches towards animals’ rights in the U.S. examining, *inter alia*, whether “animals deserve the exact same rights as people to be free from harm and exploitation” and finding a modest increase in support for this statement (from 25% in 2003 and 2008, to 32% in 2015) see Rebecca Rifkin, *In U.S., More Say Animals Should Have Same Rights as People* GALLUP (May 18, 2015), <https://news.gallup.com/poll/183275/say-animals-rights-people.aspx> [<https://perma.cc/4F3Q-6EAN>]. In these (and only these) communities, reconceptualized EDT will indeed need to de-bias human personhood from its scenarios, and recommend adapting the law accordingly. For a suggestion in this vein see Sunstein, *supra*, at 401 (“I believe that in the long run, our willingness to subject animals to unjustified suffering will be seen as a form of unconscionable barbarity—not the same as, but in some ways morally akin to, slavery and the mass extermination of human beings.”).

302. For a different, yet not entirely dissimilar, concept of equality on a fundamental level (in a philosophical context), see Samuel Freeman, *Original Position*, STAN. ENCYCLOPEDIA OF PHIL., <http://stanford.library.sydney.edu.au/entries/original-position> (last modified April 3, 2019) [<https://perma.cc/ZH2X-RQWX>]. Note also that this internal coherence between the Proposal’s methodology and its normative commitment to equality and non-discrimination does not mean the latter subscribes to some idiosyncratic or highly contested accounts of equality. It still remains the same commitment to equality and non-discrimination as described in *supra* Section III.A.1.a.

C. *Is Empirical Desert Theory Misleading and Deceiving?*

To recap, subsection A above addressed critiques pertaining to the nature of community intuitions. Subsection B followed suit by exploring permutations of a possible gap between experimental intuitions that EDT relies on and the ordinary intuitions that community members have.

Subsection C addresses criticisms of a different sort. Their common theme is that EDT does something deceptive and morally inappropriate. Yet, I will demonstrate that reconceptualized EDT is especially well suited to bolster standard EDT's response to these deception-based critiques.

1. Misleading Use of "Intuitions of Justice" and "Normative Crime Control"

Perhaps the most general criticism to be discussed in this subsection maintains that EDT habitually uses terms suggesting normative underpinnings which they lack. Namely, *EDT incorrectly and misleadingly uses the labels "intuitions of justice," "normative crime control," and "moral credibility" {7}*.

The so-called "intuitions of justice," according to the criticism, may have little, if anything, to do with *justice*. After all, EDT is concerned with *empirical facts* of people's judgments rather than the underlying *rationale* for such judgments, which may be entirely unrelated to *justice*.³⁰³ Even Robinson himself argues that "people hold strong intuitions of justice even though the *reasons* for their holding those intuitions are inaccessible to them" and these intuitions of *justice* are no more than "behavioral phenomena."³⁰⁴ Thus, EDT relies on what may best be described as intuitions of perceptions of desert, not truly intuitions of "justice."³⁰⁵

A similar point is made concerning EDT's use of the term "normative crime control." It criticizes EDT's claim that "traditional crime control mechanisms of general deterrence and incapacitation ought to be replaced with the 'normative crime-control' principle by which the criminal law tracks community views that result from empirical research."³⁰⁶ Specifically, the critics are disturbed by what they see as merely "Populist sentencing, rebranded as 'normative crime-control'"³⁰⁷

303. Ristroph, *supra* note 3, at 157.

304. *Id.*

305. Some critics also object to EDT's use of the words "justice" and "desert" interchangeably, noting disapprovingly: "Indeed, some commentators already equate desert with justice." *Id.* at 156 n.5. See also Christopher Slobogin, *Empirical Desert and Preventive Justice: A Comment*, 17 *NEW CRIM. L. REV.* 376, 379–80 (2014).

306. Roberts & de Keijser, *supra* note 4, at 487–88.

307. *Id.*

A similar criticism may apply to another key EDT concept. The term “moral credibility,” so regularly employed by EDT, might be viewed as incorrect and misleading as well. As one commentator noted, law will be morally credible if it happens to be morally “true” according to some objective factor (e.g., utility maximization or God’s will) and “the reason why citizens believe that their own moral beliefs are true is *not* the fact that they are their own beliefs, but the relation between the beliefs and one such *objective* factor”³⁰⁸ EDT’s “moral credibility,” arguably, does not match this definition. While a criminal law system following EDT might be *credible* in the public’s eyes (since it will fit its intuitions), there will be nothing “morally true” about it. The term “*moral* credibility” is therefore, inapt.

The common thread among these criticisms appears to be that EDT’s use of the normative-sounding terms “intuitions of justice,” “normative crime-control,” and “moral credibility” is incorrect because it is *misleading*. The allegation is that EDT does something improper. While EDT is strictly utilitarian and does not rely on claims from the moral standpoint, its use of normative-sounding terms may misleadingly suggest that it does.

In a crucial sense, these critiques miss the mark. They *assume* a *philosophical* perspective about “justice,” “morality,” and “normativity.” However, EDT uses these terms in a *sociological* sense,³⁰⁹ as they are perceived from the relevant community’s perspective.³¹⁰ Far from misleading, EDT *explicitly* explains and *precisely* defines the terms it uses. Briefly, “intuitions of justice”³¹¹ are simply the intuitions people have about justice. A criminal system’s “moral credibility” refers to no more than the concordance between community intuitions of justice and the system’s prescriptions.³¹² “Normative crime control” is the in-

308. Ferrante, *supra* note 32, at 470 (emphasis added). Though this argument originally supported a slightly different, related critique (which will be discussed in *infra* Section IV.C.2.) it is pertinent to the present discussion.

309. For a useful explanation of the distinction between “normative” and “sociological” desert (that is roughly equivalent to the distinction between “deontological” and “empirical” desert), see Ristorph, *supra* note 3, at 157.

310. For instance, in responding to a related criticism asserting that EDT exploits people’s commitment to retributive justice (which I will address in *infra* Section IV.C.2), Robinson noted: “[F]rom the layperson’s point of view, empirical desert is deontological desert. . . . An empirical desert distribution of punishment to them is exactly what true justice requires.” Robinson, *supra* note 18, at 62 (emphasis in original).

311. See Robinson & Darley, *supra* note 6, at 454–55; Robinson, *supra* note 35, at 149; Robinson, Goodwin, & Reisig, *supra* note 1, at 1961–65, 1995–2023.

312. Moreover, it appears that when referring to “moral credibility,” EDT uses the word “moral” in the *descriptive* (not *normative*) sense. For a general discussion of the distinction between these two senses, see Bernard Gert & Joshua Gert, *The Definition of Morality*, STAN. ENCYCLOPEDIA OF PHIL., <http://plato.stanford.edu/archives/spr2016/entries/morality-definition> (last modified Feb. 8, 2016) [<https://perma.cc/B6PU-Z47M>]. Note also that “[a]lthough most philosophers do not use “morality” in any of the above descriptive senses, some philosophers do.” *Id.* (section 1).

fluence on behavior through the powerful forces of social and individual control,³¹³ which is both correlated with moral credibility and juxtaposed to the more traditional utilitarian deterrence or incapacitation-based modes of “coercive crime control.”³¹⁴ There is nothing *misleading* about EDT’s use of these terms when they are properly understood.

Yet, in another sense, some unease related to EDT’s adoption of normative-sounding terms may still linger. As there is nothing necessarily *moral* in the theory’s “moral credibility” nor particularly *normative* in its “normative crime control,” these labels may appear as inappropriate misnomers. Arguably, these terms may mislead non-experts unfamiliar with EDT intricacies and also may improperly invoke the powerful rhetorical force associated with a common *normative* sense of words like “justice” or “moral credibility,” although these terms are technically well defined by EDT and such misleading effects were never intended.

Reconceptualized EDT addresses such concerns on two fronts. First, it explicitly adds the normative dimension to the theory. To be sure, reconceptualized EDT still relies on *community* intuitions. Yet, it does not rely on just any kind of intuitions it can measure by good scientific methods. It relies only on intuitions debiased according to a certain *normative* commitment. Unlike standard EDT, reconceptualized EDT is intentionally imbued with *normative* features and seeks to engage in a *normative* justificatory discourse. Therefore, it is entirely appropriate for it to use normative-sounding terms.

Second, reconceptualized EDT chooses to focus specifically on the normative commitment to *equality and non-discrimination*. This choice is important because, in theory, community intuitions could be debiased according to normative commitments *not* supported by the community.³¹⁵ Even this alternative could introduce *some* normative component into standard EDT, thus partially responding to the concern that its use of normative-sounding terms is misleading. Yet, reconceptualized EDT does something very different. It relies on (and coheres with) normative principles of equality and non-discrimination which the community *itself* supports.³¹⁶

313. To be sure, these forces may have genuine *moral* dimensions. They may influence people because people want to be viewed as good and moral in their own and their community’s eyes. See *supra* Section II.A. Yet, here as well, the morality of these forces’ influence may be contingent on the society at hand. Consider, for instance the potential impact they may have on the behavior of a child born and socialized in a family of devout SS officers during the heyday of Nazi Germany.

314. ROBINSON, *supra* note 10, at 95.

315. Recall, for instance, the previously discussed option of incorporating into EDT a version of deontological desert insisting on irrelevance of harm for desert. See *supra* Section III.A.1.c.

316. The ability of reconceptualized EDT methodology to gain coherence between intuitions of justice and deeply held principles may have vast normative implications, yet their comprehensive exploration is beyond the scope of this Article.

This choice allows reconceptualized EDT to engage in a special kind of justificatory dialogue. By incorporating the community's normative reasons—instead of reasons supported by (perhaps only some) moral philosophers—reconceptualized EDT goes beyond adding a normative meaning to the normative-sounding terms used by standard EDT. Reconceptualized EDT also ensures that this normative meaning *mirrors* (at least part of) the community's own understanding of normativity. As a result, even non-expert community members unfamiliar with standard EDT's definitions of normative-sounding terms are not misled by reconceptualized EDT's use of these terms. Rather, the opposite is true. Reconceptualized EDT appropriately and justifiably invokes the normative undertones associated with the *community's* normative understanding of these terms.

2. Arguments from Deception

A closely related but more specific and developed set of criticisms focuses on the notion that EDT engages in or incentivizes deception. These criticisms come in many stripes. One argues, most generally, that *EDT incentivizes concealing information from the public* {8}. Another focuses on exploitation, maintaining that *EDT exploits people's commitment to retributive justice* {9}. Yet another asserts that *EDT entails a violation of the publicity condition* {10}.

While these critiques emphasize the conceptual connection between EDT and deception, each highlights a different aspect. The “concealing information from the public” critique {8} points to a key feature of EDT—people are more likely to comply with the law if they *perceive* it as just (apart from whether it is *indeed* just).³¹⁷ This feature, according to the critique, may have perverse consequences at times; it may incentivize EDT to conceal information about legal injustices from the public. As one commentator noted: “Government should, in principle, expend resources to deceive people into believing that the system is just (e.g., by suppressing information about injustices) if this would be cheaper than expending resources to improve the actual justice of the system.”³¹⁸ As a purely utilitarian theory, EDT appears to require taking such consideration into account.

The exploitation of retributive intuitions critique {9} focuses on a different facet of EDT's alleged deceptiveness. Specifically, it maintains that EDT exploits people's commitment to “true” retributivism³¹⁹ by misleading them to think that it, too, is committed to “true”

317. As Robinson aptly put it (in a slightly different context): “It is the community's perception that justice is being done that pays dividends, not the system's actual success as measured by a deontological conception of desert.” Robinson, *supra* note 23, at 1838.

318. Simons, *supra* note 3, at 660.

319. The critique's “true” retributivism appears identical to the concept of “deontological desert.” See, e.g., Kolber, *supra* note 4, at 454. We will, therefore, use both

retributivism while in fact it is only committed to *utilitarian* goals of increasing compliance with the law.³²⁰ One commentator's illustration of this point is instructive. Consider a group of officials contemplating several alternatives of a criminal rule. If these officials are following EDT they should (barring exceptional cases) adopt the alternative best aligned with community intuitions on the matter. Importantly, they would adopt this alternative only because they are interested in increasing compliance; their *reasons* are *unrelated* to the alternative's moral value.³²¹ Yet, it is argued, what generates the (desired) increase in compliance is not only the factual alignment between community's intuitions and the legal rule but also the community's *belief* that criminal law rules are enacted for moral reasons, reasons related to *retribution*. The officials in the example allegedly exploit this (uninformed) belief and have reasons to sustain or even reinforce it.³²² They only aim to promote compliance yet wish people to believe that they aim to promote "true" retribution.³²³

The violation of the publicity condition critique {10} is also closely related to the aforementioned deception-oriented critiques, yet it anchors the discussion in the plane of moral theory most explicitly. The publicity condition, in brief, "requires that a system of morality be based on principles that can be announced publicly without thereby undermining those same principles"³²⁴ and according to this condition "no acceptable moral theory can advocate its own secrecy."³²⁵ The critique is that EDT's deceptive aspects, that is, the incentives it has for concealing information from the public, *violate* the publicity condition.³²⁶ Such violation is also claimed to be particularly challenging to EDT (but not to other utilitarian theories), since EDT "seem[s] to have already endorsed a version of the publicity condition"³²⁷ by accepting non-consequentialist limitations on methods of obtaining compliance.³²⁸

EDT has several responses to these critiques. First, EDT literature does *not* encourage deceiving the public about its nature, nor does it attempt to conceal EDT features. In contrast, EDT's most prominent proponents explicitly oppose concealing information about the operation of criminal law from the public and appear to accept the publicity condition. For instance, while EDT's co-founders Robinson and

labels interchangeably—though some commentators expressed doubts about such identity. *See supra* note 87.

320. *See generally* Ferrante, *supra* note 32, at 467–70; Kolber, *supra* note 4, at 453–59.

321. Ferrante, *supra* note 32, at 468.

322. *Id.* at 468–70. *See also* Kolber, *supra* note 4, at 453–59.

323. Ferrante, *supra* note 32, at 469; *see also* Kolber, *supra* note 4, at 454.

324. Kolber, *supra* note 4, at 459.

325. *Id.*

326. *Id.* at 434–35.

327. *Id.* at 459.

328. *Id.* at 460.

Darley mention that information regarding the criminal system injustice may be concealed (preventing the negative effects EDT warns against), they regard this option *disapprovingly* and note that concealing such information would be difficult “without breaching notions of press freedom and government transparency to which liberal democracies aspire.”³²⁹ Likewise, Robinson staunchly opposes suppressing information about criminal law and has criticized attempts of the Model Penal Code drafters to conceal certain features of it from the public view.³³⁰

Second, and relatedly, EDT has no *need* to deceive the public. An important reason can be gleaned from the following statement, clarifying EDT’s perspective:

From the layperson’s point of view, empirical desert *is* deontological desert, both in its distribution and its motivation. [Laypeople] will see no difference between the two. An empirical desert distribution of punishment to them is *exactly* what true justice requires.³³¹

This statement’s point is not that people *mistake* empirical desert for the more *correct* deontological desert (which, in theory, could have incentivized deception). The point is rather that people view *empirical* desert as correct and deontological desert (insofar as it conflicts with empirical desert) as *mistaken*. Thus EDT has no incentive to deceive people into thinking it follows a mistaken theory (i.e., deontological desert) or otherwise conceal its nature.

This line of reasoning can be further illuminated from the angle of the exploitation critique {9}. Much of it hinges on the empirical claim that people are “true” retributivists. Obviously, if this claim is false and people are *not* committed to “true” retributivism, such (non-existent) commitment cannot be exploited by EDT.

Are people “true” retributivists? This claim is difficult to test empirically because in many cases people’s intuitions appear to *match* the requirements of “true” retributivism (i.e., deontological desert). Yet, cases in which people’s intuitions appear to *conflict* with “true” retributivism suggest this claim is, in fact, false.

Consider, for instance, the aforementioned conflict between people’s intuitions and (some) “true” retributive approaches regarding harm as irrelevant for desert.³³² There is no empirical evidence that people would prefer the “true” retributive approach. If anything, the evidence suggested the opposite conclusion. People in general are not (this kind of) “true” retributivists. Rather, people appear to be retributivists of a somewhat different kind, a kind that (by definition) would be captured by *empirical* desert.

329. Paul H. Robinson & John M. Darley, *Intuitions of Justice: Implications for Criminal Law and Justice Policy*, 81 S. CAL. L. REV. 1, 1–2 (2007).

330. ROBINSON, *supra* note 10, at 97, 101–03.

331. Robinson, *supra* note 18, at 62 (emphasis added).

332. See *supra* Section III.A.1.c.

The point can be put more generally. Any instance of conflict between “true” retributivism and empirically-assessed *people’s* intuitions would reinforce the notion that *people* are not “true” retributivists. People are rather “empirical” retributivists. Therefore, EDT simply relies on the empirical *truth* concerning people’s (non-deontological) retributivism—it does not *exploit* people’s (non-existent) “true” retributivism.

Nevertheless, arguably, the deception-oriented critiques do not go away entirely. Perhaps EDT eschews concealing information from the public, does not oppose the publicity condition, and does not exploit people’s “true” retributivism because *typically* it has no incentive to do any of this. However, deception may be incentivized in atypical instances when concealing information about the system’s operation can yield substantial compliance benefits, and EDT, as a utilitarian theory, must consider them.³³³ For instance, concealing information about the system’s inner workings might be especially effective for gaining compliance from one specific group of people. While the empirical claim that people *in general* are “true” retributivists is probably false, *some* people are “true” retributivists indeed (or at least view themselves as such). EDT’s possible interaction with *this* group of people presents the best target for the deception-based critiques.

The residual deception-oriented argument can be stated along the following lines: EDT is engaged in the persuasion business, seeking to enhance a criminal system’s credibility by assisting it to gain reputation for doing justice. The main tool to achieve this goal is aligning criminal law with community intuitions of justice. For some (probably most) people, the law’s correspondence with their intuitions of justice will be sufficiently persuasive—the law will be seen as (and gain reputation for) doing justice. Yet, for “true” retributivists, this may not suffice. If they fully understand the way EDT works, they are unlikely to be persuaded by the theory’s reliance on what one commentator aptly named “aggregate data about other people’s intuitions.”³³⁴ The inability to persuade these “true” retributivists might (at least in theory) incentivize EDT’s proponents to deceive these retributivists about that theory’s inner workings. Resorting to such deception will also be inconsistent with the publicity condition. Conversely, if these “true” retributivists think (either mistakenly, or due to deception) that EDT promotes “true” retributivism and *therefore* are inclined to comply with EDT-based law, the exploitation critique becomes applicable.

Reconceptualized EDT has a good response to the deception-oriented critiques even in the case of “true” retributivists.³³⁵ As noted

333. For a suggestion in a similar vein, see Kolber, *supra* note 4, at 459.

334. *Id.* at 455.

335. One might hypothesize that the group of “true” retributivists consists of just a few moral philosophers and legal theoreticians too small to be given significant con-

above, “true” retributivists are unlikely to be persuaded by standard EDT’s aggregation of other people’s intuitions. After all, “true” retributivists might say, even if standard EDT does a terrific job at aggregating people’s intuitions of justice, there is no guarantee these intuitions match what “true” retributivism requires.³³⁶

The argument about a lack of guarantee is correct yet proves too much. After all, there is also no guarantee that what “true” retributivists *believe* to be “true” retributivism’s requirements matches what “true” retributivism actually requires. Being the fallible humans that we are, we may never have such guarantee.³³⁷

The more suitable claim “true” retributivists can make is that the *method* they use for accessing “true” retributivism’s requirements is better than the one used by EDT. This method (presumably “reflective equilibrium”) has already been mentioned above,³³⁸ and for the limited purposes of our discussion a brief description—one provided by an EDT critic—might suffice.

Reflective equilibrium is a complex philosophical process which, *inter alia*, involves: “critical reflection on and systematic revision of one’s considered convictions in *terms* of the *values of the relevant political community*. The resulting judgments are . . . aiming at a coherent account of our *deepest commitments* and their *normative implications*.”³³⁹ While intuitions also figure in the process of reflective equilibrium,³⁴⁰ it is *qualitatively* different from standard EDT’s aggregation of intuitions. Very broadly stated, “true” retributivists may claim that standard EDT engages in a merely *empirical* inquiry, while the process of reflective equilibrium engages in a *normative* one.³⁴¹ It is apples and oranges, and it is rather clear why “true” retributivists are likely to prefer the method of reflective equilibrium to assess “true” retributivism’s requirements.

Yet, and crucially, *reconceptualized* EDT is *not* exclusively engaged in an empirical inquiry. It is engaged in a normative inquiry as well.

sideration by standard EDT proponents. The reduction in compliance of members of *this* small group is negligible, and the possibility of gaining their compliance is unlikely to incentivize standard EDT to conceal its nature. This argument stands to reason (though ultimately, this is an empirical question). Yet, the critique is still worth considering. First, its major point is theoretical and doesn’t hinge on the amount of “true” retributivists in existence. Second, even if the numbers of “true” retributivists are small, their intellectual influence (and their criticisms of EDT) may be significant. For instance, there may be a much larger group of people who would prefer deferring to the opinions of the smaller, expert group of “true” retributivists, and their *deontological* conception of desert. We will return to this point shortly. See *infra* note 343.

336. Kolber, *supra* note 4, at 455.

337. For a similar point on epistemic fallibility of universalistic or moral-realist approaches to retribution and *community* views importance, see Simons, *supra* note 3, at 657.

338. See *supra* Section II.B.1.

339. Sigler, *supra* note 4, at 1174 (emphasis added).

340. *Id.* at 1179–80.

341. *Id.* at 1181, 1186.

Recall, *reconceptualized* EDT purposefully engages in a *justificatory* mode of discussion. Somewhat similarly to reflective equilibrium, by debiasing intuitions according to considerations of equality and non-discrimination, it also takes into account the pertinent “values of the relevant political community” and relies on intuitions that cohere with (some) of our “deepest commitments and their normative implications.”³⁴²

Of course, this cursory comparison between reconceptualized EDT and the method of reflective equilibrium cannot do justice to either. For our limited purposes, however, it highlights the vital point: reconceptualized EDT’s *normative* aspect makes this comparison *feasible*, as it is no longer between apples and oranges but rather between apples and apples.

The upshot is that unlike standard EDT, reconceptualized EDT at least has a *chance* to persuade “true” retributivists. Whether it can avail itself of such a chance depends on the specific merits of the argument. Eventually, “true” retributivists might remain unpersuaded and doubt that reconceptualized EDT offers a suitable alternative to their preferred method of assessing “true” retributivism’s requirements. The key point is that reconceptualized EDT at least uses a similar (normative) vocabulary to that of “true” retributivists, making the persuasion enterprise at least theoretically possible.

This possibility of persuasion is valuable for reconceptualized EDT’s response to the deception-oriented critiques. While for standard EDT the true retributivists’ existence (at least in theory) may present an incentive for concealing the theory’s inner workings, the exact opposite is true for reconceptualized EDT. It has both practical and theoretical incentives to explain its inner workings, avoiding any kind of “noble lies.” The wider it is publicized and better understood, the better chance it has to capitalize on the opportunity to convince even “true” retributivists.³⁴³

V. CONCLUDING THOUGHTS AND FUTURE PATHS

EDT is an innovative and influential criminal law theory that ingeniously answers key criminal law questions. Yet, since its inception, critics doubted whether EDT (and its recommendations) should be

342. *Id.* at 1174.

343. This persuasion opportunity might generate even broader moral credibility benefits. As briefly mentioned above (*supra* note 335), there might be a much larger group of people who are not “true” retributivists themselves but may still prefer deferring to “true” retributivism experts. For reconceptualized EDT, members of this larger group present an *opportunity* (perhaps even more important than that presented by “true” retributivists). They may find reconceptualized EDT’s exclusion of biasing factors from the intuitions it measures and its insistence to do so for *normative* reasons appealing. The opportunity to persuade the members of this group (and its potential moral credibility gains) incentivizes reconceptualized EDT even further not to conceal but, rather, to popularize its inner workings.

trusted in criminal law and beyond. They stressed immorality objections and criticized EDT, *inter alia*, for relying on community intuitions instead of experts' better judgments; cherry picking intuitions, thus squandering utilitarian benefits or illicitly slipping from descriptive to justificatory discourse; overlooking the gap between experimental and ordinary intuitions, thus being self-defeating, inadequately guiding in most challenging cases and being disrespectful to the public; misleadingly using normative-sounding terms; deceptively incentivizing information suppression; violating the publicity condition; and exploiting people's commitment to "true" retributivism.

This Article aimed to undermine these doubts. It first showed that *standard* EDT is nearly immune to immorality objections. It then demonstrated that incorporating a minimal normative commitment to equality and non-discrimination into EDT methodology *further* safeguards the theory from immorality objections, enabling it to rely on different reasons and notice different elements. Finally, it showed that reconceptualized EDT significantly strengthens standard EDT's ability to overcome the aforementioned litany of criticisms.

Crucially, the continuity, overlap, and compatibility between reconceptualized and standard EDT, conjoined with the capacity to rely on a community's normative commitments and to engage in a justificatory discourse, correspond to the *constructive* purpose of the reconceptualization: relying on and reinforcing the foundations of empirical desert theory. With immorality objections and specific criticisms weakened, these features of EDT reinforce the broader argument: EDT *should* be trusted and its findings and insights *can* be safely relied on in criminal law—and beyond.

Developing this broader argument in detail is beyond our scope. Yet, to hint at possible future paths for EDT's impact, I will present (in very broad brushstrokes) two interesting avenues of investigation beyond the narrow confines of criminal law, in which EDT insights could be brought to bear. The first addresses the potential to "export" EDT insights to a different legal regime, namely international humanitarian law ("IHL"). The second addresses EDT's potential to change the interaction between criminal law theory and philosophy.

The notoriously low levels of compliance with IHL rules have long been a source of enduring concern in the international community, leading to frequent, diverse, and tenacious efforts to improve IHL compliance.³⁴⁴ These sustained efforts, however, have been isolated from and unaffected by EDT—possibly due to concerns with criti-

344. See, e.g., James D. Morrow, *When Do States Follow the Laws of War?*, 101 AM. POL. SCI. REV. 559, 562–68 (2007) (providing data on low levels of compliance with certain IHL rules throughout the twentieth century); Pejic, *supra* note 22 (emphasizing that inadequate IHL compliance poses great challenge to IHL's "continued *credibility*" and describing efforts undertaken by numerous states and the International Committee of the Red Cross to strengthen IHL in 2012–2015).

cisms of EDT. This is regrettable because EDT methods and insights may be exceptionally valuable in this context.

EDT methods, for instance, could be used to assess the intuitions of justice of the community whose adherence to IHL is most vital—the community of service members actively engaging in hostilities (“combat soldiers”). EDT insights could then be applied for *predicting* (and ultimately influencing) combat soldiers’ IHL compliance as a function of alignment between their intuitions of justice and IHL. To see why, recall EDT’s fundamental argument; aligning law with intuitions of justice can generate substantial benefits by harnessing “normative crime control” mechanisms amplifying positive social dynamics (enhancing stigmatization, compliance in borderline cases, and the law’s ability to shape new social norms). Conversely *misalignment* entails dangers, inspiring vigilantism, resistance, and subversion.³⁴⁵ Arguably, both benefits and dangers intensify when we consider the potential impact of these mechanisms on combat soldiers on and off the battlefield.

Acts of vigilantism due to *misalignment* between community intuitions and applicable legal regimes illustrate the dangers most vividly. Vigilantism by ordinary people even in domestic contexts (and despite typical ability of criminal enforcement to curb it)³⁴⁶ can be disconcerting. Yet, if it is trained and armed combat soldiers who take the law into their own hands rather than ordinary citizens, vigilantism can be utterly terrifying, and its consequences exceptionally grave—jeopardizing combatants and civilians on and off the battlefield.

Vigilantism aside, misalignment between IHL and combat soldiers’ intuitions of justice may engender more mundane risks. For instance, combat soldiers participating (as defendants, witnesses, or jurors) in proceedings concerning violations of misaligned IHL rules may be more likely to resist and subvert these proceedings (rather than acquiesce and cooperate with them), which in turn may further decrease the moral credibility of (and compliance with) IHL. Conversely, when IHL and combat soldiers’ intuitions of justice are aligned, these dangers could be mitigated by harnessing the positive dynamics of normative crime control.³⁴⁷

345. See *supra* Section II.A.

346. Robinson, *supra* note 67, at 404.

347. Note, I make *no* claim that combat soldiers intentionally engage in vigilantism due to misalignment between their intuitions of justice and IHL. I only claim such misalignment *may* entail risks, and assessing combat soldiers’ intuitions may assist in predicting and mitigating them. Naturally, interactions between such misalignment and compliance could be complex and indirect. For instance, misalignment may increase the *probability* for intentional non-compliance or the probability that soldiers will *unintentionally* endeavor less to comply. Obviously, additional reasons may influence soldiers’ unlawful behavior (e.g., revenge, stress, and expediency). See, e.g., James D. Morrow, *The Institutional Features of the Prisoners of War Treaties*, 55 INT’L ORG. 971, 976, 984 (2001). For a detailed account on the reasons soldiers commit

Even this cursory analysis suggests that EDT methods and insights can aid international efforts to increase IHL compliance. Assessing combat soldiers' intuitions of justice and *predicting* their compliance with IHL rules can be a valuable first step to decrease violations in compliance-inhibiting situations. The means may range from tightening personnel supervision, to improving IHL training, to increasing reliance on technology.³⁴⁸ Additionally, an unorthodox and intriguing option is more available now than ever before: instead of modifying combat soldiers' behavior to comply with IHL rules misaligned with their intuitions of justice, it may be advisable to modify these IHL rules (or their interpretations).

Until now, this option—aligning the law (IHL) with community (combat soldiers') intuitions of justice, as the strictly utilitarian standard EDT could recommend—may have appeared unappealing due to standard EDT's ostensible vulnerability to immorality objections and additional criticisms. Yet, because reconceptualized EDT undermines these concerns, this option's appeal increases. Moreover, reconceptualized EDT illuminates the unsettling prospect that efforts to increase compliance by fitting soldiers' behavior into the Procrustean bed of IHL rules that are misaligned with their intuitions of justice may not only *fail* but be normatively *inappropriate*. Modifying IHL to fit combat soldiers' intuitions of justice (as ascertained through the normative lens of reconceptualized EDT) may be the normatively *right* thing to do. Relatedly, reconceptualized EDT's ability to engage not only in *predictive* but also in *justificatory* discourse, may both facilitate aligning IHL with combat soldiers' intuitions of justice in a *normatively responsible manner* and help *persuade* combat soldiers to comply with IHL.

Naturally, these are merely preliminary bearings for future explorations. Extensive research will be required to ascertain whether they are merited. Yet reflecting on these matters may be valuable in itself and open additional investigative avenues. For instance, whether using combat soldiers' (or other community) intuitions of justice in a normatively responsible manner is feasible may hinge on larger questions pertaining to the potential of a constructive relationship between EDT and *philosophy*.

The topic of relationship between EDT and philosophy has been present throughout the Article, including in discussions on immorality

atrocities in war see Mark J. Osiel, *Obedying Orders: Atrocity, Military Discipline, and the Law of War*, 86 CALIF. L. REV. 939, 1028–44 (1998).

348. Curiously, one such possibility may be deployment of non-human, autonomous weapon systems. For suggestions that these systems may improve upon human performance in mitigating dangers of armed conflict, see, e.g., Kenneth Anderson & Matthew C. Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*, 15 (Stanford Univ. Hoover Inst., American Univ. Wash. College of Law, Research Paper 2013-11, 2013, Columbia Pub. Law Research Paper 13-351, 2013).

objections and (most recently) on “true” retributivists’ criticisms. Yet, beneath separate EDT critiques, a more general issue has been lurking. As one commentator explicitly and forcefully put it: standard EDT (with its recommendations based on experimental assessments of intuitions of justice) attempts to substitute what is ultimately a *moral* inquiry for an *empirical* one; yet EDT’s empirical inquiry can only sketch a shallow “two-dimensional” picture of the considerations relevant for desert judgments, while moral inquiry can represent them comprehensively.³⁴⁹

We encountered such claims for philosophical superiority while discussing “true” retributivists. I stressed that its crux is the ostensible superiority of philosophical *methodology* (e.g., the method of reflective equilibrium). I showed, moreover, that reconceptualized EDT’s methodology’s *normative* aspects bear some similarity to and enable comparison between it and philosophical methods, thereby potentially challenging the latter’s superiority.

Curiously, the ostensibly superior philosophical methods had come under pressure from a different source. Since the beginning of this century, a vibrant discussion was spurred by the Experimental Philosophy movement, which brought social-scientific thought to bear on philosophy.³⁵⁰ Experimental philosophers challenged traditional “armchair” philosophical methods, especially methods employing intuitions (e.g., the method of reflective equilibrium), and demonstrated that philosophers are susceptible to biasing influences that are extremely difficult to eliminate.³⁵¹ Notably, some experimental philosophers called for assessing intuitions via the less bias-prone *experimental* methods in order to make progress on philosophical questions.³⁵²

The parallels between EDT and experimental philosophy are striking: both apply social scientific insights on law and philosophy, respectively, and are mutually skeptical towards the same philosophical methods. The similarities increase further for reconceptualized EDT,

349. For claims along this line see Sigler, *supra* note 4, at 1179–81, 1186.

350. For a general overview of this movement, see Joshua Knobe & Shaun Nichols, *Experimental Philosophy*, STAN. ENCYCLOPEDIA OF PHIL. (Dec. 19, 2017), <https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy> [https://perma.cc/63H5-AUSF]. For claims about earlier origins of the movement and its name, see Knobe & Nichols, *supra* note 107, at 3, 7, 13; Justin Sytsma & Jonathan Livengood, *Experimental Philosophy and Philosophical Disputes*, 13 *ESSAYS PHILOSOPHY* 145, 159 n.1 (2012).

351. For a claim that such methods cannot “sufficiently detect its own susceptibilities to *bias and error*” see Jonathan M. Weinberg, *The Methodological Necessity of Experimental Philosophy*, 25 *DISCIPLINE FILOSOFICHE* 23, 27 (2015). See also *supra* note 234.

352. Knobe & Nichols, *supra* note 350 (section 1); Mark Alfano & Don Loeb, *Experimental Moral Philosophy*, STAN. ENCYCLOPEDIA OF PHIL. (Mar. 19, 2014), <https://plato.stanford.edu/archives/fall2017/entries/experimental-moral> [https://perma.cc/8N7D-7VW8] (section 2.2).

which—like some experimental philosophers—sees value in using intuitions derived via experimental methods for *normative* discourse.

The nexus between EDT and experimental philosophy is well worth exploring. First, the methodological discourse driven by the latter in *philosophical* literature further challenges *legal* literature on EDT's assertions on philosophical methods' superiority. Yet, it strengthens *reconceptualized* EDT's appeal further. As we have seen, reconceptualized EDT can address superiority claims directly on the *normative* plane, as it engages in a *normative* inquiry. Additionally, experimental philosophy discourse suggests that reconceptualized EDT methodology might be *preferable* to traditional philosophical methods. It is not vulnerable to the biasing influences highlighted by experimental philosophy and comparably relies on the less bias-prone *experimental* methods. Accordingly, reconceptualized EDT's amalgamation of normative considerations and empirical methods arguably offer uniquely valuable opportunities for progress on philosophical questions pertaining to desert.

Again, this analysis only hints at a possible path for future investigation, and it is far from suggesting that reconceptualized EDT provides a panacea for all desert-related philosophical conundrums. Yet, it suggests that *reconceptualized* EDT presents opportunities for constructive partnership between EDT and philosophy, a significant accomplishment in and of itself. Recall, one of *standard* EDT's most valuable achievements was enabling direct engagement between retributivism and utilitarianism—talking about *desert* while using utilitarian vocabulary.³⁵³ Yet, this achievement also exposed standard EDT to critiques ranging from immorality objections to the superiority claim raised by deontological desert theorists and philosophers. Now, mirroring the move made by standard EDT several decades earlier, reconceptualized EDT comes full circle. It enables direct engagement between empirical desert researchers and deontological desert theorists and philosophers—talking about experimentally derived intuitions while using philosophical vocabulary.

Reconceptualized EDT, as we have seen throughout this Article, safeguards standard EDT from immorality objections, strengthens and augments its ability to overcome individual criticisms, and instills trust that its findings and insights can be safely relied on in criminal law—and beyond. In addition, reconceptualized EDT's ability to speak the scientific language of utilitarianism and the normative language of philosophy holds another promise. It may foster dialogue that transcends familiar divides and carry on standard EDT's project, by further bridging the gap between retributivism and utilitarianism.

353. See *supra* Section I.

