

RESPONSES TO CRITICS

*Leo Zaibert**

TABLE OF CONTENTS

I. CAHILL AND THE VARIETIES OF PLURALISMS.....	1023
II. LEE, THE MEANING OF LIFE, AND THE MANAGEMENT OF OUR EMOTIONS.....	1031
III. BERGELSON, THE UNEASY RELATIONSHIP BETWEEN PUNISHMENT AND FORGIVENESS, AND BETWEEN THE AXIOLOGICAL AND THE DEONTIC.....	1040
IV. CHIESA, FREE WILL, AND THE POSSIBILITY OF DESERVED SUFFERING.....	1048
V. YANKAH, LEAVING CHILDHOOD BEHIND, AND MISSING THE MORAL FOREST FOR THE POLITICAL TREE	1054
VI. CONCLUSION	1063

I am very grateful to Rutgers Law School for the Symposium it organized around my book, *Rethinking Punishment*, which took place in the Baker Courtroom at the Center for Law and Justice of the Newark Campus on September 21, 2018. My gratitude above all goes to Vera Bergelson, who first thought of this idea, and then took care of the all-too-often thankless tasks related to the organization of the event. I am, of course, also indebted to all the participants at the Symposium: Vera Bergelson, Mitchell Berman, Michael Cahill, Luis E. Chiesa, Stephen P. Garvey, Youngjae Lee, Alice Ristroph, and Ekow Yankah, and to the members of the audience—some of whom, particularly Anthony Dillof and Douglas Husak—participated quite actively. I feel simultaneously honored and humbled to have such a distinguished group of colleagues, whose work I admire so much, engage with my book. Finally, I am very grateful to the editors of the *Rutgers University Law Review*, for deciding to memorialize substantial portions of what transpired during that Symposium.

*Professor of Philosophy and Department Chair, Union College. Email: zaibertl@union.edu.

The Symposium fostered precisely the sort of robust and refreshing discussion that I had hoped *Rethinking Punishment* would generate. These pages afford me with the opportunity to try to alleviate the effects of those bothersome bouts of *sprit d'escalier* that often seize us (or at least some of us) after unrehearsed public interventions. Independently of the potential therapeutic effects these pages may have on me, however, I believe that the discussions contained herein should contribute to the advancement of the debate over the justification of punishment. One motivation I had for writing the book was my sense that this debate had reached a sort of stalemate, and that it was really very hard to find new strategies to approach this important topic. For a variety of reasons, I have come to find serious problems with the traditional approaches to the debate: as much with the “pure” justifications—retributivism and consequentialism—as with the extant “mixed” justifications. *Rethinking Punishment* seeks to offer a new perspective from which to approach the justification of punishment. Obviously, the book does not resolve all relevant problems, and, as some of my critics intimate, it perhaps does not resolve even those problems it should have resolved. But I am convinced that the new perspective that the book suggests is at least promising—and this is a precious opportunity to further explain why I think that this is so.

Below I respond to each critic more or less individually, except for some brief concluding remarks that I think apply more or less globally. As it turns out, my responses appear in the same order that the papers were presented at the Symposium, but no special significance is to be attributed to the order in which I proceed. Unavoidably, some repetition is to be found below, in the sense that some issues appear in more than one response, even if in slightly different ways. I have not attempted to respond to each and every question posed or every objection raised, or with the same degree of depth regarding those with which I do engage. While the number of issues my critics have raised may speak favorably of the book's scope, it may render the prospect of providing fully satisfactory answers to all of these issues more elusive. I have tried to focus on those questions and objections that I find particularly important, and that when gathered together make good on the book's recommendation of creating a more valuable organic whole.

I. CAHILL AND THE VARIETIES OF PLURALISMS

Cahill finds the task of responding to *Rethinking Punishment* “daunting” given that he is “in firm and fundamental agreement with its central claims.”¹ Particularly salient, I think, are our agreements about endorsing pluralism in the approach to the justification of punishment, and about rejecting purely deontic versions of retributivism. I welcome our agreement about these views, in part because, echoing Cahill, such agreement increases “confidence in the soundness of these views.”² From my perspective, this increased confidence is not just a quantitative matter—not simply the result of two being greater than one—but much more importantly a qualitative one—the result of the respect I have for Cahill’s work.

Cahill further points out that these sorts of agreements have at least one potentially unhappy result: they cast doubt on the sense in which it could be said that one has “made a unique contribution to the intellectual conversation,” since, he believes, we “are not at all alone among recent commentators” in holding the views we hold.³ Thus, I would like to begin by attempting to elucidate the sense in which *Rethinking Punishment* has made an original contribution to the debate—indeed one that is more original than it may appear on first approximation.

Very roughly, a justification of punishment is pluralistic if it recognizes the existence and importance of different types of values. As it will become clear in due course, I do not believe that simply recognizing *more* values is to thereby necessarily more pluralistic: I believe that *which* values one recognizes is also very important. At the Symposium, Cahill was not alone in wondering about the extent to which my pluralism differs (or does not differ) from other forms of pluralism in the specialized literature, his own included.⁴ If my pluralism is not too different from other pluralisms, then, in that sense at least, *Rethinking Punishment* may not be as original as I claim it is. But I think that there exist importantly different *types* of pluralism—a fact that itself has not

1. Michael T. Cahill, *Thoughts on Zaibert’s Rethinking*, 71 RUTGERS U. L. REV. 937, 937 (2020) (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

2. *Id.*

3. *Id.*

4. This line was also advanced by Berman, and it largely informs his review of *Rethinking Punishment* commissioned by a different publication. See Mitchell N. Berman, *Rethinking Punishment*, NOTRE DAME PHIL. REVS. 1 (2018), <https://ndpr.nd.edu/news/rethinking-punishment/> (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

been sufficiently recognized in the specialized literature. One of the pluralisms I endorse—indeed the most important of them—is fundamentally different from other pluralisms, Cahill’s included.

In other words, I do not believe that defending pluralism—*simpliciter*, as it were—is what is terribly original about my book.⁵ Indeed, Cahill, myself, and others are sympathetic to some general forms of pluralism, and to its implications (both theoretical and practical), above all, on state punishment, about which there scarcely is anything terribly original at all. Rather, what I believe is particularly original (and particularly promising) about the main specific form of pluralism that I defend in the book—which I have called “proper pluralism”—is that it is centrally focused on two specific values: the value of punishment and the value of forgiveness.⁶ This is not meant to exclude the existence (or the importance) of other values, but to *focus* specifically on these two.⁷

As soon as this explanation is provided, however, at least two possible obstacles appear to stand in the way of my “proper pluralism,” one relating to its alleged “pluralistic” credentials and the other relating to its alleged “propriety.”⁸ First, it may seem as insofar as I admit that my “proper pluralism” is focused on just two values, it qualifies as a pluralism only by incorporating the barest of possible pluralities—a plurality of two. And indeed Cahill takes me to task for this: “While purporting to be pluralistic, Zaibert’s project seems entirely uninterested in incorporating many obvious competing goals, values, or costs, instead choosing to concern itself exclusively with the need to make space for forgiveness alongside retribution.”⁹ For this reason, Cahill adds, my “account seems suboptimal in its narrowness.”¹⁰ Second, it may seem that my account is overly stipulative. What exactly is it that makes my account of pluralism “proper”? As Cahill asks, why not admit that a pluralism that extended beyond mine and that “recognized” other values (such as “efficiency, equality, procedural fairness, harm prevention, privacy, and so on”) is much more “properly” pluralistic than mine?¹¹

5. I do not even believe that the originality of the book is exhausted by my take on the problem of pluralism, but this is the topic at hand now. Other aspects of the book that I *also* find original shall emerge in due course below, organically.

6. LEO ZAIBERT, *RETHINKING PUNISHMENT* 30 (2018).

7. *Id.*

8. *Id.*

9. Cahill, *supra* note 1, at 939.

10. *Id.* at 939 n.8.

11. *Id.*

A quick answer to these two closely inter-related questions is that my pluralism does “recognize” those other values that Cahill mentions: I am not “exclusively” concerned with the value of inflicting deserved suffering and of its merciful remission—I am merely *focusing* on these two, without at all denying the existence of many other important values.¹² In other words, I am both a pluralist in the usual sense in which Cahill and others also are pluralists, and I am a pluralist in the sense of being a “proper pluralist,” a sense in which very few other punishment theorists are pluralists. But the quick answer is in a way evasive, in that I certainly do not *engage* with those other values whose existence I do nonetheless recognize in *Rethinking Punishment*.¹³ And yet the quick answer really is, I think, the beginning of wisdom. My focus on the values of deserved suffering and of its merciful remission—my “proper pluralism”—is not only *consistent* with the more general pluralism of which Cahill (and others) speak,¹⁴ and regarding which he (and others) have said so many insightful things, but *feliciously* so. What I mean is this: even if my proper pluralism is indeed *focused* on but two values, I do want to continue to refer to my position as pluralistic, for independently of the peculiarities of the ways in which my version of proper pluralism differs from other versions (a point to which I will turn immediately), the convergence of these different versions reveal, I think, valuable lessons for our effort to justify punishment.

In *Rethinking Punishment* I argue that the tension between these two specific values—the infliction of deserved suffering and its merciful remission—is (a) one that has been systematically ignored by contemporary punishment theorists, and (b) crucial in our efforts to advance the discussion of the justification of punishment beyond the stalemate of sorts in which I suggest it is today.¹⁵ Let us take these two claims in order. The two values on which I focus are not just any two values, chosen amongst many. Rather, they occupy a different theoretical plane altogether, as it were. Consider the following analogy, which may help explain what I have in mind.

12. ZAIBERT, *supra* note 6, at 30.

13. *See id.*

14. *See generally* Michael T. Cahill, *Punishment Pluralism*, in *RETRIBUTIVISM: ESSAYS ON THEORY AND POLICY* (Mark D. White ed., 2011).

15. ZAIBERT, *supra* note 6, at 22–24.

Imagine a time in which nutritionists were divided into “monistic” (those who believed that a healthy lifestyle is exclusively a matter of eating less sugar) and “pluralists,” (those who believed that there are a number of factors that go into a healthy lifestyle: eating more fiber, eating more fruits and vegetables, drinking more water, etc.). And imagine that at some point a nutritionist came along and suggested that the starting point of pluralism ought to be the recognition of the importance of *both* diet and exercise. Imagine that previous nutritionists had systematically ignored this focus on “exercise.” It seems to me that the nutritionist who introduces the concern with exercise is really offering a form of nutritional pluralism that is *importantly* different from earlier nutritional pluralisms. Diet and exercise—as opposed to, say, more fruits and more fiber, or more water and smaller portions, etc.—seem to me to occupy a different theoretical plane. The Kuhnian expression “paradigm shift” is often abused, but I think it is pertinent in this case: this nutritionist is offering a paradigm shift regarding nutrition. This analogy seeks to capture what I am trying to do when I claim that the fundamental (and hitherto under-investigated) axiological conflict in the justification of punishment is the tension between punishment and forgiveness.¹⁶ (The introduction of “exercise” in the thought experiment corresponds to my introduction of “forgiveness” in *Rethinking Punishment*, of course.)¹⁷

To the extent that one finds the exercise-forgiveness analogy I have just presented compelling, then one is, I think, likely to see the originality of my version of proper pluralism. The analogy is simply meant to capture something special about a given pair of values as opposed to different pairs (or larger collections) of values. But it is important to highlight two dimensions along which the analogy does not even fully capture the novelty (and, I hope, the promise) of my version of proper pluralism.

First, the two values around which my pluralism is constructed are *conceptually* opposed: at one and the same time, if you punish you are not forgiving (and vice-versa), whereas there is no conceptual opposition between diet (however multifariously understood) and exercise. Moreover, since there is *prima facie* value in both punishment and forgiveness this conceptual opposition is drenched in normative complexity. In the book I recommend paying close attention at the often-

16. See *id.* at 23, 220.

17. See *id.* at 17.

used expression “the problem of punishment,” and I suggest distinguishing between *practical* problems of punishment and a *theoretical* problem of punishment.¹⁸ My focus in *Rethinking Punishment* is above all on the latter. And most of the problems around which extant versions of pluralism are constructed concern practical matters: resource allocation, political priorities and other matters of expediency, and so on. There is absolutely nothing wrong with caring about those sorts of problems: as I put it very early in the book, these problems are “both pressing and depressing.”¹⁹ But I focus on the other sort of problem, the theoretical problem: even if we stipulated away all the problems that can arise in the context of conflicts between the important values that Cahill mentions, even if we (unrealistically) imagined a society made up of people of unlimited good faith and blessed with unlimited resources, and much more, the theoretical problem of punishment on which I focus on in *Rethinking Punishment* would remain. Punishing *means* (synchronically) failing to realize the value of forgiveness (and, of course, forgiving *means* (synchronically) failing to realize the value of punishment).

Second, and much more importantly, the imagined nutritionist I presented above, introduces exercise into the discussion of nutrition out of the blue, as it were. But the seminal idea that there is something qualitatively unique about the specific pair of values around which my proper pluralism is built is not really mine alone, even though it has not been systematically mobilized within the context of the contemporary specialized literature on punishment. In a sense, I stand on the shoulders of giants.²⁰ The originality of my focus obtains mostly within the context of contemporary punishment theory, and not so much within the context of a more generally understood history of ideas. The tension I have in mind is (a version of) the *classical* tension between justice and mercy. Ever since ancient times it has been the tension between these two specific values that has attracted the attention of thinkers of all persuasions. As I put it in the book, I find it “not ... insignificant” that even Isaiah Berlin, one of the fathers of contemporary pluralism, would choose to refer to *this* specific pair of values when describing his

18. *Id.* at 2.

19. *Id.*

20. *See id.* at 1–2. (referencing Isaiah Berlin, Franz Brentano, Jonathan Dancy, G. E. Moore, Michael Stocker, and Bernard Williams). Within the context of contemporary punishment theory, and as noted in the book, John Tasioulas comes closest to embracing the sort of proper pluralism I defend. *See id.* at 29.

pluralism.²¹ The situation within contemporary punishment theory is, however, quite different; in this specific context, forgiveness (or mercy) is hardly ever discussed.²² The standard opponent of “retributivism” is “consequentialism,”²³ and it is this particular opposition that either informs or is assumed by versions of pluralism such as Cahill’s “consequentialist retributivism” and other pluralistic positions.²⁴

Within the more general context of the history of ideas, the standard opponent of “retributivism” is not *consequentialism* but *forgiveness*. And it is this other opposition that informs my proper pluralism. With very few exceptions (which I discuss in the book), the vast majority of contemporary punishment theorists, even those who *are* (in a sense) pluralists, like Cahill, are not pluralists in this “proper” or “theoretical” sense I defend in *Rethinking Punishment*—a sense that derives from focusing on the tension between punishment and forgiveness. In the book I have tried to unmoor this classical opposition from its traditional surroundings and bring it to bear upon contemporary punishment theory, where I think it has received insufficient attention and where I think it has much to contribute.

For all I know the name I chose for my version of pluralism—“proper pluralism”—may be off-putting and thus not the best choice. Perhaps it would have been helpful to refer to my version of proper pluralism as “theoretical pluralism,” and to other, extant versions of pluralism as “practical pluralisms,” with the caveat that I do not intend “theoretical” to mean “superior”—just as I do not intend “proper” to mean “superior.” In referring to my pluralism as “proper,” I simply meant “pluralistic in this *foundational* way.” I thought that this name was both less cumbersome than its alternatives (would “global pluralism dismissed by contemporary punishment theorists” be better?), and that it somehow managed to at least gesture at the long and venerable pedigree of this type of pluralism *outside* of contemporary punishment theory.

I believe that my proper pluralism is only “narrow” if we attend exclusively to the number of values on which it *focuses*, and ignore both

21. *Id.* at 23.

22. As I explain in *Rethinking Punishment*, for current purposes mercy and forgiveness can be treated as synonyms. *See id.* at 3–4, 119–21. I continue to do so here.

23. David Dolinko, *Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment*, 16 LAW AND PHILOSOPHY 507, 507–08 (1997) (“One feature on which retributivists and their critics have generally agreed is that retributivism is very much a non-consequentialist theory.”).

24. Cahill, *supra* note 1, at 937.

(a) the qualitative significance of these two values in particular, and (b) the number of values that it nonetheless *recognizes*. If we do that, my pluralism admittedly cannot be any narrower. But why restrict ourselves in this way? If I am right about the theoretical and normative significance of the two values around which my proper pluralism is built, then the fact that these are really only two values and not, say, twenty, strikes me as somewhat off-target. For reasons I discuss in the book, I continue to believe that this pair of values *is* particularly rich and interesting, and contemporary punishment theorists *have* systematically ignored it.

Cahill poses a number of other penetrating questions, several of which are amenable to be grouped together under a heading questioning the “practical implications” of my view. Perceptively, Cahill wonders about how forgiveness is supposed to operate within the context of an actual criminal law, about who is supposed to grant it, and even about what exactly forgiveness would *be* in that context. These are important questions and I am aware that I do not have full—or fully satisfactory—answers for them. But I hope that at least some of what I have said above—and some of what I say below—can be helpful, even if only in general and indirect ways.

Cahill believes in the “conceptual priority of having to define a practice before one can attempt to justify it.”²⁵ And he poses a number of important conceptual questions about punishment: Does it have to be intentionally inflicted?²⁶ Must it be experienced or understood (by the putative punishee) in a certain way? Must it be administered by the state? So much do I agree with Cahill on this point that this was one of the central concerns in my previous work on punishment, before I wrote *Rethinking Punishment*. For example, although I wanted my book *Punishment and Retribution* to be about the justification of punishment, I felt forced to deal first with a host of conceptual points such as the ones Cahill has in mind.²⁷

In fact, in that book (and also elsewhere) I tackled *the very questions* about punishment that Cahill now poses. Something similar has happened within the context of my work on forgiveness. My main articles on forgiveness were supposed to be about its justification, but I got

25. *Id.* at 942.

26. *Id.*

27. LEO ZAIBERT, *PUNISHMENT AND RETRIBUTION* 3–5 (Ashgate Publishing Co.) (2006).

tangled up on conceptual matters.²⁸ Admittedly, I am not as confident as to my views on the conceptual contours of forgiveness as I am about my views on the conceptual contours of punishment. As I explained in the book, this is in part the result of forgiveness being somehow much more paradoxical than punishment. Degrees of confidence aside, I am by no means certain that I have said anything approximating the “the last word” on either of these topics, even if I may have said more about punishment than about forgiveness. But I could not have written *Rethinking Punishment* and tackled the justificatory problem I tackle therein without having had some views as to the conceptual contours of punishment and of forgiveness. And regarding this methodological strategy, too, I find myself in fundamental agreement with Cahill.

But this point sets the stage for, in the midst of our many and fundamental points of agreement, the main disagreement between Cahill and myself. And it strikes me as a “meta-level” disagreement: a disagreement about the state of the contemporary specialized literature on punishment. Cahill admits that he had “hoped” that in *Rethinking Punishment* I would not have devoted so much attention to the “justification” of punishment, and instead focused on the “meaning” of punishment.²⁹ For he believes that there has been “progress toward a broad consensus (if by no means universal) view regarding the justification of punishment.”³⁰ I must confess that I do not see the progress toward a consensus that Cahill sees: (pure) retributivism seems to me as (theoretically) oblivious to consequentialism as it always was—and as it *must* be, logically speaking—and vice-versa. These positions display an “inability to share the stage” with others, as Michael Moore has put it when describing retributivism.³¹ While mixed justifications keep gaining in sophistication, they still do not seem to me to be fully tenable. Given the main theses I advance in *Rethinking Punishment*,

28. See, e.g., Leo Zaibert, *The Paradox of Forgiveness*, 6 J. OF MORAL PHIL. 365 *passim* (2009) [hereinafter *The Paradox of Forgiveness*]; Leo Zaibert, *Forgiveness: An Introduction*, 92 THE MONIST 481 *passim* (2009); Leo Zaibert, *Punishment and Forgiveness*, in PUNISHMENT AND ETHICS: NEW PERSPECTIVES 92, 106 (J. Angelo Corlett & Jesper Ryberg, eds., 2010); Leo Zaibert, *On Forgiveness and the Deliberate Refusal to Punish: Reiterating the Differences*, 9 J. OF MORAL PHIL. 103 *passim* (2012).

29. Cahill, *supra* note 1, at 942–43.

30. *Id.*

31. MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF CRIMINAL LAW 28 (1997). While Moore makes the point about retributivism alone, the point applies to classical utilitarianism as well. See LEO ZAIBERT, DESERT-SENSITIVITY AND MORAL EVALUATION (forthcoming 2020).

these mixed justifications strike me as incomplete, at least in the sense of not engaging with the theoretical problem of punishment I discuss in the book.

Of course, to the extent that Cahill may be right about there being less of a stalemate in the debate over the justification of punishment than I claim there is, perhaps there is less reason to be impressed by the change of perspective that I offer in *Rethinking Punishment*. But to the extent that I am right in seeing the debate over the justification of punishment as really stalemated, then the importance of the change of perspective that I offer in the book may be admitted as potentially quite valuable, even if some of my views in the book may be too general, too abstract, or too inchoate.

I will not here offer an item-by-item enumeration of the things in *Rethinking Punishment* that I think are both original and potentially fruitful (although, again, some of that will, I hope, happen organically as I respond to the different contributions, below). But consider one last time my explanation as to why my version of proper pluralism is original. If I am right, punishment theorists have been overly focused on one type of justificatory debate—that opposing the traditional justifications of punishment: retributivism, consequentialism, and the mixed justifications—at the expense of another, and in a sense more fundamental (at least in the sense of being more general, or logically prior) debate. This neglected debate opposes the value of punishing *at all* (independently of whatever turns out to be one’s favored justification of punishment) and forgiving (in the sense of deliberately refusing to punish). I cannot help thinking that if this is on the right track, there should be important consequences both for how we theorize punishment and its justification, and, at the practical level, for which punitive schemes deserve our support.

II. LEE, THE MEANING OF LIFE, AND THE MANAGEMENT OF OUR EMOTIONS

I frequently begin my “Introduction to Philosophy” courses warning students that if they expect me to instruct them about the meaning of life, they will be disappointed. And yet Lee is right in that the meaning

of life is an important topic in *Rethinking Punishment*.³² Lee is right, too, in that rare as this topic is amongst contemporary philosophers in general, it is even rarer amongst contemporary punishment theorists in particular. And this is indeed part of why in the book I try to enrich and refresh this specific field—a field so dear to me—by putting it into closer contact with a series of other discussions, and in particular with recent developments in moral philosophy.

While Lee is in general appreciative of my efforts (or so it seems to me), several aspects of my views in the book worry him. Before addressing these specific worries, I would like to underscore that many (if not most) of Lee's worries are not really caused by the views that I actually defend. He is worried about how others may misinterpret me, despite what I actually say. Furthermore, it is not even that Lee believes that I have somehow put things in ways which lend themselves to misinterpretation, so that I could perhaps be criticized at least for that, even if not for my views themselves. Thus, it is not easy to know how exactly I am supposed to proceed in assuaging these sorts of worries.

In order to see the connection between the lofty topic of the meaning of life and punishment, it is important to underscore that my focus on the theoretical problem of punishment is a focus on a *general* problem of punishment. As we just saw in the context of my discussion of Cahill's contribution, this focus is not meant to either ignore or otherwise diminish the importance of criminal punishment by the state. I do not wish to deny that there are peculiarities about state punishment that may perhaps not obtain in other contexts (or vice-versa). Still, and as we have also seen, knowing what punishment (in general) is, or how to justify it (in general), cannot possibly hurt our efforts to understand what state punishment (in particular) is, or how to justify it (in particular): knowing these things may actually be necessary for these latter efforts to succeed. Despite these differences, it seems to me that there are at least some common conceptual elements that, as long as we are talking about punishment, obtain in all contexts. Chiefly amongst them is that a punisher seeks to make (someone she perceives to be) a wrongdoer suffer. Suffering is, as a matter of definition, inseparable from punishment.

This is why in the book I connect the problem of punishment to the problem of theodicy, that is, to the problem of why there is suffering in

32. Youngjae Lee, *State Punishment and Meaning in Life*, 71 RUTGERS U. L. REV. 945, 945 (2020) (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

the world. Theodicies are often discussed within religious or theological contexts: how can an allegedly benevolent God allow for (any) suffering to exist? But theodicies can also be secular—as mine is meant to be. If suffering is (in general) bad, how could it possibly be good to deliberately bring it about (via punishment)? Although the connection between theodicies and punishment strikes me as obvious, contemporary punishment theorists have essentially (and cavalierly) ignored it. This lack of attention to the problem is partly explained by the fact that many punishment theorists see punishment as an *instrument*, or, to use terms I borrow from Berlin and from Leibniz (respectively) as a *technology*, or a *medicine*.³³ From that very widespread perspective, punishment may cause suffering, but since its very point is to diminish suffering overall, then it is not really a matter for theodicy after all. To those who believe that mankind “is under the governance of two sovereign masters, pain [i.e., suffering] and pleasure,” as Bentham famously has it, the idea of deserved suffering possibly adding meaning to the world has to be anathema.³⁴ The only approach to the justification of punishment that does not reduce it to an instrument is retributivism; the only approach that *can* see deserved suffering as a means to “suffuse the world with meaning,” is retributivism.³⁵

Lee thus correctly focuses on one of the several arguments and thought-experiments I discuss in the book that show how distributions according to desert, in and of themselves, appear to have intrinsic value: W.D. Ross’s example of two worlds identical in every respect (including the amounts of happiness and suffering in each), which differ only on the distribution according to desert in one and not in the other.³⁶ Again, while seemingly sympathetic to my point, Lee distances himself from the “somewhat abstract” nature of this sort of argument and turns our attention to a series of “concrete” examples: the case of the repeated rapes

33. ISIAH BERLIN, *Two Concepts of Liberty*, in LIBERTY 166, 166 (Henry Hardy ed., 2002); GOTTFRIED WILHELM VON LEIBNIZ, THEODICY: ESSAYS ON THE GOODNESS OF GOD, THE FREEDOM OF MAN, AND THE ORIGIN OF EVIL 425 (Austin Farrer ed., 1951).

34. JEREMY BENTHAM, *The Principles of Morals and Legislation*, in THE WORKS OF JEREMY BENTHAM 1 (John Bowring ed., 1962) (emphasis omitted).

35. ZAIBERT, *supra* note 6, at 10; see also Leo Zaibert, *The Instruments of Abolition, or Why Retributivism is the Only Real Justification of Punishment*, 32 L. & PHIL. 33, 33–58 (2013).

36. See ZAIBERT, *supra* note 6, at 37–39.

of Amber Rose Carson and examples from the movies *While We're Young*, *The Untouchables*, and *Minority Report*.³⁷

Lee groups together these concrete examples by attending to the fact that the “meaning” that some of the characters in these examples find in punishment may be the wrong sort of meaning: sometimes people (particularly victims) tend to be overly severe in their reactions to wrongdoing, and sometimes large segments of society prefer—consciously or otherwise—false but convenient narratives. No doubt this is possible, and in contexts that extend beyond punishment. But I think that despite the way in which truth is admittedly under attack these days, there is no doubt, either, that this is only a possibility and not a certainty. Be that as it may, it is this possibility that sets the stage for what I take to be Lee’s main worry—to repeat: a worry that concerns a potential misinterpretation of my views, to which I really have contributed nothing (beyond offering the views themselves). The worry is that my emphasis on *suffering* may appear to give an imprimatur of sorts to those overly severe tendencies in people and society that both Lee and I would like to curb.

This seems to be the reason Lee favors “approaches that deemphasize suffering in accounts of punishment,”³⁸ and why he has reservations about my insistence that there is a “conceptual connection between punishment and suffering.”³⁹ Lee thinks that my granting this conceptual connection may in fact bring about a state of affairs whereby my “talk about the value of forgiveness gets ignored.”⁴⁰ Needless to say, I do not want the value of forgiveness to be ignored, since I do believe that one of the central contributions of *Rethinking Punishment* is, precisely, that it brings the discussion of the justification of punishment into much closer contact with the discussion of the justification of forgiveness than is typically done. But I do not think that Lee should have this worry, and explaining why this is so may allow me to clarify some important aspects of some of the views I defend in the book.

The fact that, as a matter of definition, attempting to punish *means* attempting to inflict suffering says very little (if anything at all) as to whether or not such infliction of suffering would be *justified*. For example, Lee and I—and many others—look askance at the

37. Lee, *supra* note 32, at 946–52.

38. *Id.* at 958.

39. ZAIBERT, *supra* note 6, at 3.

40. Lee, *supra* note 32, at 958.

dysfunctional criminal justice system in the United States, and at the myriad ways in which it generates suffering. This suffering is all too often undeserved; it is often mindless, and often downright cruel and discriminatory. Depressing as these facts are, they are independent from the conceptual problem as to what punishment *is*. To think otherwise is to place what should be the independent and neutral analytical task of defining punishment in the service of the advancement an extraneous agenda having to do with limiting the reach and scope of the criminal law. (It just so happens that Lee and I both support this agenda—but that is beside the point.)

As I explain in *Rethinking Punishment*, I use “suffering” in a very broad sense, very similar indeed to the also very broad sense of “consequences normally considered unpleasant” found in what is known as the standard definition of punishment.⁴¹ In *this* regard, there is nothing terribly unorthodox about the substance of what I am saying. While in substantive terms the actual word I chose and the word found in the most widely endorsed definition of punishment are equivalent, “suffering” is admittedly more poignant than the usual but rather bloodless “consequences normally considered unpleasant” locution.⁴² This is precisely because I want to avoid sugarcoating; I want to remind ourselves of what punishment *really* is: to punish is to deliberately make someone suffer (as a response to her wrongdoing, etc.). As noted above, the most horrific aspects of the American criminal justice system are not typically related to this inescapable sort of suffering, but instead to all the undeserved suffering it generates.

But even an imaginary, perfect criminal justice system would inescapably be, as long as it is a system of *punishment*, a system seeking to make people (wrongdoers) suffer, and it is one of the central theses of *Rethinking Punishment* that this fact does not just disappear because punishment is justified.⁴³

41. See ZAIBERT, *supra* note 6, at 5, and the references to Anthony Flew, S. I. Benn, and H.L.A. Hart—the champions of that standard definition—therein.

42. *Id.* at 5.

43. The linguistic connection between the criminal law and punishment is clearer in languages other than English: *derecho penal*, *diritto penale*, *droit penal*, *Strafrecht*, for example, all convey the intimate connection between this branch of the legal system and the very act of punishing. While we have “penal law” in English too, the connection between it and the very act of punishing is not quite as obvious. The title of an article that I wrote in Spanish included the phrase “El Derecho Penal sin Penas,” and the oddity to which I was alluding seems to me much more neatly captured in Spanish than in its English equivalent

It is precisely in light of what punishment *is*, that in the very last sentence of the book, I recommend a certain “circumspection” when we (or the state, etc.) contemplate the possibility of punishing a wrongdoer.⁴⁴ So, I do agree—wholeheartedly—with Lee about the fact that

given the enormous amount of harm that the state can bring about in people’s lives through its coercive and judgmental uses of its power to criminalize and punish, the observation that the state is merely giving people *what they want* or satisfying *the people’s thirst for revenge* by itself cannot serve as a justification for the institutional setup.⁴⁵

A justificatory strategy based on that sort of populism would be very uninspiring indeed.

And it is clear that Lee does not think that such is my strategy, or that I appeal to observations of that sort. In fact, I cannot but celebrate that Lee thinks that the sort of argument likely to allow the criminal law to overcome populism and the tyranny of the majority “could look a lot like Zaibert’s argument about desert and meaning of life,” or that “without an argument like Zaibert’s, we can end up with a theoretical justification for the institution of punishment that is hollow at its core.”⁴⁶

Despite all this, Lee worries that my intentions in *Rethinking Punishment* may get lost in the shuffle, as it were, and that my views may have the perverse effect of causing just the opposite of what I want.⁴⁷ In particular, Lee worries about the ways in which my emphasis on suffering may thwart what he deems a “core purpose” of the criminal law: “to manage the punitive and retaliatory emotions of those who have been victims of wrongdoers,” (a purpose in turn grounded on what Lee sees as the coercive, judgmental, and preemptive characteristics of the criminal law).⁴⁸

In order to avoid hollow institutions (and hollow practices more generally) it is important to remember that punishment can best add

“A Criminal Law without Punishments.” See Leo Zaibert, *Política, Ciudadanía, y Liberalismo: El Derecho Penal sin Penas*, in DERECHO PENAL DEL ENEMIGO: EL DISCURSO PENAL DE LA EXCLUSIÓN 1149–73 (Cancio-Meliá y Gómez-Jara-Díez eds. 2006).

44. ZAIBERT, *supra* note 6, at 242.

45. See Lee, *supra* note 32, at 952.

46. See *id.* at 953.

47. See *id.* at 957–58.

48. Youngjae Lee, *Why Proportionality Matters*, 160 U. PA. L. REV. 1835, 1836–38 (2012).

meaning to our lives when it is not a mere “technology,” or a mere “medicine,” in the sense of these terms explained above.⁴⁹ Again, my goal in mobilizing these admittedly unfamiliar terms in the book is to highlight the risks here—precisely the sorts of risks to which Lee is alive. The “management” of emotions of which Lee speaks better be done in a way that is responsive to *real* human nature and to *really* meaningful considerations: paradigmatically, punishment adds meaning to our life when it is *just*. And, paradigmatically, punishment is just when it is *deserved*.

Thus, in addition to those sensible core purposes Lee mentions, another absolutely core purpose of criminal law better be to impart *justice*, even if in this context—the context of responding to wrongdoing—justice is a matter of making the wrongdoer suffer (deservedly, proportionally, etc.). In this context, justice is achieved *by means of* suffering. And it is because this sort of justice—retributive justice—must be done by means of suffering that I find it so inherently dilemmatic—so *theoretically* dilemmatic—and indeed much more complicated than it is all too frequently assumed to be.

It is regarding this point that it seems to me Lee and I do have a disagreement. It is not that Lee denies that justice is important or valuable. Rather, it is that he thinks that the *suffering* on which I focus is optional: “the right level of condemnation need not be expressed in terms of inflictions of suffering, as, say, a symbolic response can suffice.”⁵⁰ What matters, in Lee’s opinion, is that the state response to wrongdoing be *fitting* and fittingness need not require the infliction of suffering.⁵¹

I will shortly elaborate on a sense in which I *very much* believe that things other than giving people the suffering that they deserve can be fitting (and sometimes more fitting than inflicting deserved suffering). But I want to register my hesitation regarding Lee’s position. It strikes me that a criminal law that systematically failed to mete justice on a regular basis would hardly deserve its name, and would be very difficult to defend. Perhaps there are cases in which Lee’s “symbolic responses” may indeed be better, all-things-considered, than inflicting deserved suffering (even though, *ex hypothesi*, the latter are *also* fitting). But it

49. ZAIBERT, *supra* note 6, at 10.

50. Lee, *supra* note 32, at 957.

51. *Id.*

seems to me undeniable that there are cases in which mere “symbolic responses” would not be fitting enough—or even fitting at all: some situations call for *real* and *robust* justice, and not just for *symbolic* justice.

In any case, what is crucial to keep in mind is that these other responses—fitting as they may be—are *not* punishments. This is part of the reason for my insistence that *punishment* better not be seen as a mere technology or medicine, seeking to alleviate this or that societal problem. There is of course nothing wrong with alleviating societal problems, but this should not lead us to reduce punishment to a bare mechanism for such alleviations. And this is why I quoted Strawson to the effect that “to speak in terms of social utility alone is to leave out something vital in our conceptions of these practices.”⁵² And why I have echoed his view that it is a mistake

to forget that these practices [which include inflicting deserved suffering, and believing that inflicting deserved suffering can add value to organic wholes], and their reception, the reactions to them, really are expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes. Our practices do not merely exploit our natures, they express them.⁵³

My conceding that on this or that occasion a token non-punitive response can be more fitting—and all-around better—than a token punitive response in no way affects the *definitional* point on which I have insisted above: *punishment*, by definition, is a matter of attempt to inflict (deserved) suffering on a (perceived) wrongdoer. So, I can indeed agree with Lee that “to the extent that the state can find a way to devise a fitting and appropriate response to criminal wrongdoing without inflicting suffering, that is a goal worthy of our attention.”⁵⁴ But I would insist that, to the extent that these imagined responses do not seek to inflict suffering, they are not punitive. And while there are many very good reasons recommending that the state be less punitive, I have trouble imagining a state that *never* punished.

52. ZAIBERT, *supra* note 6, at 69 (quoting P.F. STRAWSON, *FREEDOM AND RESENTMENT AND OTHER ESSAYS* (Routledge 2008)).

53. *Id.* at 69 (emphasis omitted). I will come back to these passages from Strawson in the context of my discussion of Chiesa, below.

54. Lee, *supra* note 32 at 957.

Lee is perfectly aware of the importance that in *Rethinking Punishment* I give to forgiveness.⁵⁵ As we saw in the context of my replies to Cahill, I consider this emphasis on the tension between punishment and forgiveness a truly unique contribution of the book (and an important differentiator in terms of versions of pluralism). But more substantively, one of the most difficult aspects of the discussion of forgiveness is to determine what exactly is the type of reason that renders a certain refusal to punish a genuine instance of forgiveness. As I will explain in more detail below, in the context of my discussion of Bergelson's contribution, it is a relatively common view amongst forgiveness theorists to insist that forgiveness must be done "for a moral reason."⁵⁶ But it is not clear what exactly that reason is.

While admittedly incomplete, in the book I suggest that the type of reason that may do the work we need done here better be non-instrumental. Perhaps surprisingly, this is the same sort of reason that also underwrites *axiological* retributivism.⁵⁷ It is a reason that is, in any event, robustly connected to truth, and not to the production of this or that particular goal, however attractive such a goal may be. And I think that this sort of reason is also what Lee is after when he worries about the ways in which we can avoid finding meaning in the wrong ways. If I am right, society needs both *true* justice, which sometimes requires real, robust punishment linked to the value of justice as such, and *true* forgiveness, which requires an elusive but at the very least non-instrumental reason for not punishing. (It may need, also, non-punitive but fitting responses to wrongdoing that are not instances of forgiveness.)

Finally, Lee is also aware that I share his preoccupation regarding the "real human costs" that even a "perfect" criminal justice system would generate.⁵⁸ And he realizes that the sort of vulgar majoritarianism that would underwrite simply giving people what they want without regards to truth or justice is not at all my position.⁵⁹ Still, it is possible, I guess, that my insisting on "a 'conceptual connection between

55. *Id.*

56. Jeffrie G. Murphy, *Forgiveness and Resentment*, in JEFFRIE G. MURPHY & JEAN HAMPTON, *FORGIVENESS AND MERCY* 503, 508 (1988).

57. Early on, when I thought that I was going to be able to discuss the justification of forgiveness (as opposed to its conceptual contours, as explained in my discussion of Cahill's contribution), the working title of what eventually became my "The Paradox of Forgiveness" (*op. cit.*), was "A Retributive Theory of Forgiveness."

58. See Lee, *supra* note 32, at 958.

59. See *id.* at 953–54.

punishment and suffering’ would [unjustifiably] still tend to make us inured to the existence of suffering in our criminal justice system while [my] talk about the value of forgiveness [unjustifiably] gets ignored.”⁶⁰ I just do not know what I am supposed to do regarding this extreme misunderstanding of my views, particularly when my motivation for speaking bluntly about “suffering” rather than about “consequences normally considered unpleasant” is precisely to try to keep the real human costs of punishment—even when justified—front and center.⁶¹

III. BERGELSON, THE UNEASY RELATIONSHIP BETWEEN PUNISHMENT AND FORGIVENESS, AND BETWEEN THE AXIOLOGICAL AND THE DEONTIC

Bergelson begins her contribution by describing the “central theme” of *Rethinking Punishment*.⁶² She does this accurately, focusing on what I have just been calling the tension “between justice and mercy, punishment and forgiveness.”⁶³ Interestingly, rather than calling it a tension, she calls this an “uneasy relationship.”⁶⁴ I do not disagree with Bergelson’s characterization, and in fact in the book I go to great lengths in trying to embrace the really complicated nature of the relationship between these phenomena.⁶⁵ But Bergelson’s characterization is meant, in part at least, to register a disagreement with my views on the matter—or at least with what she takes my views on the matter to be. In other words, Bergelson accepts my (implicit) challenge of focusing on exactly what I think we should focus—justice and mercy—and still she finds the results of my suggested approach unacceptable.

It should not be terribly surprising that I agree with Bergelson (and others) on the importance of desert in general, and of deserved suffering in particular. After all, despite a number of important qualifications,⁶⁶ I continue to consider myself a retributivist, at least in the eminently *axiological* sense I explain in *Rethinking Punishment*: suffering, when it

60. *Id.* at 958 (quoting LEO ZAIBERT, *RETHINKING PUNISHMENT* 3 (2018)).

61. *See* ZAIBERT, *supra* note 6, at 5.

62. Vera Bergelson, *Justice or Mercy?*, 71 RUTGERS UNIV. L. REV. 959, 959 (2020) (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

63. *Id.*

64. *Id.*

65. ZAIBERT, *supra* note 6, at 21–26.

66. Some of these qualifications have already been presented above, and others will be presented below.

is deserved, can add, in and of itself, value to the organic whole in which it appears.⁶⁷ I am convinced that the moral valence of suffering is *not* invariant: the fact that suffering is deserved *can* change its moral valence, and it frequently *does* change it. Furthermore, these retributivist inclinations at the axiological level better translate into *something* at the level of action. In other words, it is not surprising that Bergelson and I—and others who share some of these highly plausible retributive intuitions—would in fact treat a number of cases quite alike. There *are* many cases wherein Bergelson’s bottom line as to what is to be done will be very similar to my bottom line as to what is to be done.

But two qualifications are in order. First, I suspect that Bergelson thinks that there is more disagreement between us than there really is, because I think that Bergelson misattributes to me the view that “it is incorrect to think that, even in paradigmatic cases, the value of inflicting the deserved punishment is greater than the value of its merciful remission.”⁶⁸ My view, however, is not quite that, but rather that contemporary punishment theorists tend to “operate as if all cases, or at least the paradigmatic cases, of punishment were such that the difference between the value of inflicting punishment is much greater than the value of its merciful remission.”⁶⁹ The word “much” from the quoted passage disappeared from the summary of the views as Bergelson attributes them to me, and so is the fact that I am talking about what others think.⁷⁰ Moreover, I openly admit that there exist “cases in which the contrast between the specific wrong being punished and its merciful remission” is indeed “extreme.”⁷¹ In other words, while I criticize contemporary punishment theorists for too facilely assuming that the value of deserved punishment is much greater than the value of its merciful remission as if by default, I can and do believe that such cases are perfectly possible. Sometimes the value of deserved punishment is greater—indeed much greater—than the value of forgiveness; sometimes the other way around.

Second, our path toward that deontic conclusion may perhaps be different, and this difference is itself important. As (following Bernard

67. ZAIBERT, *supra* note 6, at 33.

68. Bergelson, *supra* note 62, at 960.

69. *See id.* at 960 n.9 (quoting LEO ZAIBERT, RETHINKING PUNISHMENT 235 (2018)).

70. *See id.*

71. ZAIBERT, *supra* note 6, at 235.

Williams)⁷² I note in the book, the ultimate actions of the punisher with integrity need not be too outwardly different than those of the punisher without integrity—and yet the differences between these two punishers are not altogether unimportant.⁷³ One hope I had in writing *Rethinking Punishment* is that the sort of generally retributive worldview to which Bergelson and I subscribe would be *tempered* by our realizing that the merciful remission of deserved suffering is *also* valuable. So, people can agree about their deontic views—about what is to be done in this or that case—while disagreeing at the axiological level, and about the path that takes us from the axiological to the deontic. Again, these tempered dispositions, arising from taking the axiological turn seriously (a turn that recognizes *both* the value of deserved suffering and of its merciful remission) need not entail that every single case will be treated differently.

Furthermore, it is not only regarding retributive judgments (or intuitions) that I agree with Bergelson's bottom line. I also largely agree with her conceptual analysis seeking to show that seemingly good candidates for instances of mercy are in fact not genuine instances of mercy⁷⁴—a point that may not be surprising given my remarks concerning the importance of avoiding the *wrong kind of reasons* at the end of my discussion of Lee's contribution (to which I will return shortly). Bergelson's comprehensive list of enticing but ultimately spurious instances of "mercy" includes executive amnesty, prosecutorial discretion, grand juries' decisions not to prosecute, power of nullification, judges' decisions to reduce or eliminate punishment, and executive clemency.⁷⁵

I want to make two points about the items Bergelson groups together. First, the items on her list are all related to *state* action, that is, to action taking place within a very rich and very complex institutional setting.⁷⁶ This commonality is one of the reasons why both in the context of my work on punishment and in the context of my work on forgiveness, I have tried to steer clear of these sorts of rich contexts and have instead focused

72. See Bernard Williams, *A Critique of Utilitarianism*, in VICE AND VIRTUE IN EVERYDAY LIFE: INTRODUCTORY READINGS IN ETHICS 124, 128–132 (5th ed. 2000).

73. ZAIBERT, *supra* note 6, at 108.

74. See Bergelson, *supra* note 62, at 963–65.

75. See *id.*

76. See *id.*

on “punishment outside the [s]tate”⁷⁷ and on “pure forgiveness.”⁷⁸ I have never denied that the state punishes (in fact, I have criticized how much and how severely some states punish), or that state punishment is a particularly important (politically and morally) instance of punishment. Neither have I denied that forgiveness can be an overt, communicative act. But I have insisted that if we wish to understand the *conceptual* contours of punishment and forgiveness, we can benefit from carrying out our analyses, at least at the outset, in ways that avoid the complexities—and indeed the *distractions*—that rich institutional settings generate.⁷⁹ I have thus recommended a version of the phenomenological reduction, or bracketing, which tries to isolate the essential conceptual bits that constitute punishment or forgiveness.

So, a presidential pardon, for example, is not quite an instance of forgiveness—just as Bergelson points out,⁸⁰ even if perhaps we arrive at this conclusion through different paths.⁸¹ And some instances of state “punishment” are not really instances of punishments either (say, those cases in which the “punisher” does not believe that the “punishee” did anything wrong, etc.).

The second point concerns my adhering to the view that forgiveness (or mercy) needs to be granted “for a moral reason.”⁸² Bergelson agrees that this view, which in the book I analyze in the context of discussing the seminal way in which Jeffrie Murphy developed it, a view that I find fundamentally correct, although (as explained in the book) I do not think it goes far enough.⁸³ But Bergelson parts company with me when I insist on the fact that much more work needs to be done in order to further specify the kind of moral reason that is been envisaged here. In *Rethinking Punishment*, I recall Claudia Card’s example of releasing a mafia don in order to prevent his associates from wreaking havoc on our city and its innocent inhabitants.⁸⁴ I follow Card in thinking that this is not a real instance of forgiveness. Given her treatment of the items on

77. See ZAIBERT, *supra* note 6, at 22.

78. See *The Paradox of Forgiveness*, *supra* note 28, at 381.

79. See, e.g., ZAIBERT, *supra* note 6; *The Paradox of Forgiveness*, *supra* note 28, at 365.

80. See Bergelson, *supra* note 62, at 966 n.52.

81. While the linguistic proximity between pardoning and forgiving obtains in English, in other languages, such as Spanish (and other Romance languages), the connection is even closer, since the very same term “perdón” can be used in both cases.

82. ZAIBERT, *supra* note 6, at 191.

83. See Murphy, *supra* note 56, at 24–34; ZAIBERT, *supra* note 6, at 191.

84. ZAIBERT, *supra* note 6, at 191.

her list of spurious instances of mercy/forgiveness, I am confident that Bergelson agrees with this too.⁸⁵ But I have argued that part of the reason the expression “for a moral reason”—useful as it unquestionably is—requires much more work is that preventing the havoc that the mafia don’s associates may wreak *is* a perfectly recognizable moral reason: we want to prevent harm to innocent people, etc.⁸⁶ Thus, the bald demand that forgiveness/mercy be done for a moral reason is evidently insufficient: the release of the mafia don in Card’s example *is* done for a moral reason, and yet it *is not* an instance of forgiveness. And, similarly, the actions in Bergelson’s list may perhaps be all carried out for moral reasons, and yet they may fail to constitute genuine instances of forgiveness.

Thus, I fail to understand why Bergelson disagrees with my invitation to further explore the kind of moral reason that is necessary to render a certain apparent instance of forgiveness a genuine instance of forgiveness.⁸⁷ In fact, I find Bergelson’s own forays into this territory quite insightful and helpful. Bergelson suggests two characteristics that would exclude a moral reason from being the right kind of moral reason (in the sense of being the reason that generates genuine forgiveness).⁸⁸ First, the right kind of reason better not be “instrumental” (just as I argued in the context of my discussion of Lee’s contribution), and, second, that it must be motivated by a desire to “benefit *the punishee*.”⁸⁹ Both of these are considerations that in *Rethinking Punishment* I suggest must be part of the story as to what renders a reason the right kind of reason in the context of mercy.⁹⁰ Of course, neither Bergelson nor I believe that these two characteristics constitute a complete analysis of what exactly is the right kind of reason in this context. But they are a good starting point, and a starting point that, again, I consider promising, and in any event that should help in the further specification of the kind of moral reason required in these cases.

Bergelson takes me to require that the right kind of reason, the reason that would render seeming instances of forgiveness and seeming instances of punishment genuine instances of forgiveness and genuine

85. See Bergelson, *supra* note 62, at 966–69.

86. ZAIBERT, *supra* note 6, at 192.

87. Bergelson, *supra* note 62, at 965–66.

88. *Id.*

89. *Id.*

90. See ZAIBERT, *supra* note 6, at 192–93.

instances of punishment be *objectively* “good and correct.”⁹¹ And she disagrees with this view because she believes that the right kind of reason could be “misguided and mistaken,” and yet be enough, if the person genuinely holds the reason “subjectively.”⁹² Again, I do not disagree with Bergelson. As evidence of my adherence to an “objective” standard for determining the right kind of reason, Bergelson cites this passage from *Rethinking Punishment*: “[G]enuine cases of forgiveness... [are those] in which although it would be just—and *valuable*—to punish someone who deserves punishment (and who has no excuses, mitigations, or justifications), we nonetheless conclude that it is all-things-considered better—i.e., *more valuable*—if we remit this punishment.”⁹³

I am not sure I see where the “objectivity” that Bergelson detects here is: I only claim that we need to conclude something (“subjectively,” in Bergelson’s terms), not that we should *correctly* conclude it (“objectively” in Bergelson’s terms). Perhaps it is the mere fact that I am against allowing the wrong kind of reason to do the work we need here. But I think that this is different from the subjective/objective distinction. It is enough for my purposes that the punisher or the forgiver acts for the right kind of reason, even if she is mistaken. For example, if a person refused to inflict deserved suffering on a wrongdoer because she had a plane to catch, this would be the wrong kind of reason (and would not be a genuine instance of forgiveness), even if she is subjectively right about her itinerary. Similarly, if a person refused to inflict deserved suffering on a wrongdoer because she thinks that so doing so would generate the most valuable organic whole possible, this would be the right kind of reason (and it would be a genuine instance of forgiveness), even if she is subjectively mistaken about the value of the organic whole she has in mind. Thus, we can indeed be subjectivists (in Bergelson’s sense) and still insist that the action be motivated by the right kind of reason.

So, I think that in some important cases in which Bergelson believes that there are disagreements between us, there really are no significant disagreements. But I want to conclude with one instance of the opposite situation: a case in which Bergelson claims to agree with me, but regarding which I think that we are in a significant disagreement. I am surprised by the ease with which Bergelson appears to grant the

91. Bergelson, *supra* note 62, at 965.

92. *Id.*

93. *Id.* at 965 n.47 (quoting LEO ZAIBERT, *RETHINKING PUNISHMENT* 178 (2018)).

remainder thesis with which I operate in the book, and which I think has been largely absent in contemporary criminal law theory. Not only does Bergelson agree that “punishment . . . generates remainders,” but she believes that “everything else” generates remainders too.⁹⁴

Bergelson’s point is well taken (after all, an important part of my strategy in the book is to underscore the complexity of moral life *in general*, and punishment is evidently not the only complex thing in our complex moral life). But I do not share Bergelson’s confidence in the view that quite *everything* generates remainders. Imagine giving the award for the best poem in a poetry competition to the person who wrote the unquestionably best poem (and who, additionally, happened to be the nicest person, who needed the award the most, and who would put it to the best use, etc.). I am not sure that I can discern what a remainder may turn out to be in such a case. To suggest that, perhaps, the remainder would be our inability to give the award to all competitors (or to all who need it, or to all nice people, etc.) risks utterly undermining the very cogency of giving an award in a competition, or flattening the notion of a remainder at play here.

The remainders on which I focus obtain when we find ourselves in situations in which we cannot fail to do wrong, no matter what we choose to do, and even though some of what we choose to do is indeed the right thing to do: our predicament in these situations is that we cannot do right without simultaneously doing wrong. (This situation does not seem to me to obtain in the poetry competition just described.) And it seems to me that the typical retributivist—and at any rate Bergelson—cannot easily join me in embracing the remainder thesis as I have just described it. After all, Bergelson is committed to the view that the value of deserved suffering is “fundamentally superior to the value of its merciful remission,” and that therefore punishers do not need to necessarily dirty their hands, certainly not when they are punishing justly.⁹⁵ This does not seem to me to be compatible with the idea that a punisher who does the right thing (i.e., she punishes justly, because she is justified in punishing, etc.) is also doing something somehow wrong (inflicting suffering), and *is* dirtying her hands. This is why, while still considering myself a retributivist, I distance myself from typical retributivism. I want to

94. *Id.* at 961.

95. *Id.*

distance myself from any justification of punishment that does not take seriously the intrinsic complexity of punishment.

In fact, I would consider Bergelson a mercy skeptic, at least about mercy's role within the context of the criminal law. She is, after all, perfectly emphatic in that "what we routinely call 'mercy' in the criminal law is usually not mercy," and that a well-functioning system of criminal law does not really need mercy at all.⁹⁶ And it is because of this skepticism that she concludes that within the context of this well-functioning criminal law, punishers need not dirty their hands, or, in other words, they do not face the sort of conflict I claim they do face.⁹⁷ Bergelson is in good company regarding this sort of mercy skepticism; in the book, I discuss versions of it I find in Heidi Hurd, in R.A. Duff, and even in the whom I consider the greatest contemporary defender of the role of mercy *within* criminal law, John Tasioulas.⁹⁸

But about this issue, Bergelson and I, do, at long last, have a genuine disagreement. Since Bergelson is, in the final analysis, a mercy skeptic, she disagrees with my view whereby we cannot punish without dirtying our hands.⁹⁹ En route to this disagreement Bergelson (understandably) disagrees with my mobilization of *Billy Budd*.¹⁰⁰ Above all, Bergelson suggests that this fictional case is quite unlike real cases we face in criminal law.¹⁰¹ More generally, Bergelson thinks that "to test our intuitions on whether the punisher's remorse is an inherent part of punishment, we should not get distracted by the most egregious, painful, or deadly forms of punishment. Instead, we should mentally replace them with lesser forms of punishment, such as fines or community service."¹⁰²

Again, Bergelson's point is well taken. But I have not suggested that *Billy Budd* is a representative example of a criminal case in our courts. Instead, I have argued that "there really exists a certain opposition between certain *moral* emotions and the type of moral simplemindedness that I have criticized in this book and that *Billy Budd* exposes in such a

96. *Id.*

97. *Id.*

98. ZAIBERT, *supra* note 6, at 118–46, 189.

99. *See* Bergelson, *supra* note 62, at 976.

100. *See id.* at 973–74.

101. *See id.*

102. *Id.* at 975.

masterful fashion.”¹⁰³ In other words, *Billy Budd* is valuable for my purposes because it portrays an *admittedly* extreme case of the conflict between justice and mercy. But—and here’s the crucial point—I *do* want to claim that punishment necessarily generates this *type* of conflict, even if not quite in such intense ways: this is, after all, the theoretical problem of punishment. To the extent that Bergelson is a mercy skeptic, and to the extent that she downplays the importance of the sort of remainders I am talking about, she really is not supposed to agree with me on this.

There is a corollary to Bergelson’s disagreement with my take on the theoretical problem of punishment. She believes that my position is in fact “weak[]” and “artificial[]” in the sense that there are cases in which we can punish without generating remainders, without getting our hands dirty, and without feeling much tragic-remorse or agent-regret.¹⁰⁴ I do recognize that there are cases in which the difference between the necessarily opposed values—the value of inflicting deserved suffering and the value of its merciful remission—is so enormous that the fact that one of them goes unrealized does not give rise to particularly intense forms of remorse or regret. My view is that at the axiological level, however, the conflict between these two values necessarily obtains, although in those cases in which one of the values really is overwhelmingly greater (or more important, etc.) than the other value the conflict may perhaps not be felt too intensely. In no way does this militate against my view that contemporary punishment theorists have too often and too uncritically assumed that the typical case in our courtroom is one in which the value of inflicting deserved suffering is indeed overwhelmingly greater than the value of its merciful remission (that is if they find *any* value in the latter at all).

IV. CHIESA, FREE WILL, AND THE POSSIBILITY OF DESERVED SUFFERING

Part of the reason for the peculiar situation I described in my response to Bergelson’s contribution (whereby what appears as disagreements are not really so, and what appears as an agreement is not really so) is, to an extent, the result of paying insufficient attention to the distinction between the axiological and the deontic. I think something similar happens here in the context of my response to Chiesa’s

103. ZAIBERT, *supra* note 6, at 222.

104. See Bergelson, *supra* note 62, at 976.

contribution. The difference is that, unlike Bergelson, Chiesa is not *directly* objecting against me, which renders my task of responding to him more complicated.

Rethinking Punishment contains an “Appendix” in which I discuss the late Derek Parfit’s view that suffering cannot be deserved.¹⁰⁵ Were Parfit correct, the thesis that there is any *tension* (or uneasy relationship) between deserved suffering and its merciful remission would be a non-starter, for there simply would not be such a thing as deserved suffering. I find Parfit’s position odd mainly because he is not a skeptic about desert *in general*: it is only suffering that he thinks cannot be deserved.¹⁰⁶ Unable to find any arguments in Parfit’s work that would explain why suffering is so exceptionally unique, I suggested that Parfit was guilty of cherry-picking what can and cannot be deserved—and I rejected his position.¹⁰⁷ Chiesa agrees with me “that Parfit offers little support for his view” and that I am essentially correct in rejecting what I called Parfit’s “compatibilism à la carte.”¹⁰⁸ Courageously, Chiesa nonetheless undertakes to defend a “kind of selective approach to compatibilism.”¹⁰⁹ While Chiesa does offer much more support for his view than Parfit does for his own, I remain unconvinced by what I see as efforts to (arbitrarily) turn compatibilism on and off.

Chiesa laments the fact that retributivists pay so little attention to the problem of free will, given that, in his view, “the challenges posed to retributive theories of punishment by the free will problem are quite formidable.”¹¹⁰ Although Chiesa believes that the lack of attention is manifested by criminal law theorists in general, he believes that it is *retributivist* punishment theorists who are more properly at fault for being either wholly disengaged or, at best, either engaged “superficial[ly]” or “uncritically adopting some philosopher’s view on the subject”¹¹¹ Since Chiesa kindly notes the “breadth and depth” of

105. See ZAIBERT, *supra* note 6, at 243.

106. *Id.*

107. *Id.* at 246–47.

108. Luis E. Chiesa, *Selective Incompatibilism, Free Will, and the (Limited) Role of Retribution in Punishment Theory*, 71 RUTGERS U. L. REV. 977, 979–80 (2020) (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

109. *Id.* at 980.

110. *Id.* at 978.

111. *Id.* at 978 n.3.

Rethinking Punishment, he had hoped to find in it *some* engagement with this admittedly venerable philosophical problem.¹¹²

Alas, such engagement is indeed nowhere to be found in the book. Very early in *Rethinking Punishment* I announce that “I will not engage in this book with those thoroughgoing forms of skepticism, fatalism, incompatibilism, or determinism that call into question moral responsibility in general.”¹¹³

This is not, I hope, to be overly cavalier. And it is not that I doubt that the problem of free will is, in other contexts, important. Rather, it is that I think that systematically engaging with that problem would have entailed an entirely different book. Furthermore, I think that some version of compatibilism—that is, some recognition of the fact that this deeply intractable metaphysical problem *can* be bracketed—is downright necessary if we are to attempt to do *punishment* theory. For if moral responsibility turned out to be a fantasy, moral theory (or at least important aspects of it) would face the same fate. And to the extent that punishment theory presupposes *some* moral theory, then it *is* inconsistent with those deterministic positions that deny compatibilism (in the sense of “compatibilism” just sketched).

I am by no means alone in bypassing the metaphysical debate between free will and determinism. Chiesa is perfectly aware that “most philosophers and legal scholars who tackle the free will problem advocate for some kind of compatibilist account,” and, again, compatibilism *is* a way of bypassing the metaphysical debate.¹¹⁴ I would simply add that this is true of most philosophers, legal scholars, and ordinary people—even those who do not officially advocate for (or indeed have even heard of) compatibilism. Some form of compatibilism seems to be—inside and outside academia—a precondition for engaging in a *moral* discussion in good faith.

Just as I did not undertake a systematic takedown of determinism in *Rethinking Punishment*, I do not do that here either. But it may be worth underscoring a relatively small but extremely important point. A certain *robust* and non-mechanistic understanding of ethics runs through *Rethinking Punishment*: I detect strands of this sort of robust approach in a number of authors, from Leibniz forward, going through Weber, and

112. *Id.* at 978.

113. ZAIBERT, *supra* note 6, at 3 n.8.

114. Chiesa, *supra* note 108, at 987.

culminating with twentieth century luminaries such as Isaiah Berlin, Bertrand Russell, P. F. Strawson, and Bernard Williams. I find such understanding at odds with determinism in general. Within the context of responding to Chiesa, a word on Strawson, in particular, seems in order. For I cannot avoid concluding that Chiesa does not sufficiently engage with Strawson, or worse: that Chiesa misrepresents Strawson's views. (What is beyond dispute is that Chiesa and I read Strawson quite differently.) Chiesa sees Strawson as a rather unsophisticated compatibilist who insisted on the importance of "reactive attitudes" simply in order to nurture "the creation and maintenance of meaningful personal relations."¹¹⁵ But it is clear that while Strawson sought to salvage the "optimists" (a type of compatibilists), he thought this was only possible by way of introducing what he confessed was a "radical modification" of their views.¹¹⁶ After all, Strawson thought that without this radical modification the optimist's way salvaging of moral responsibility was attained too cheaply.¹¹⁷ So, this alone shows that Strawson was no garden-variety compatibilist.

Furthermore, and more substantively, Chiesa takes Strawson to endorse a rather vulgar form of consequentialism: Chiesa claims that Strawson believes that we hold people responsible "not because they are truly morally responsible," but rather because this is necessary "for constituting meaningful relationships with others."¹¹⁸ Chiesa's interpretation transforms Strawson into precisely the "one-eyed utilitarianism" which Strawson so forcefully opposed, and which I, too, sought to systematically criticize in the book.¹¹⁹ I do not see how to reconcile Chiesa's interpretations with Strawson's own words, as I quoted them in the book (and in my discussion of Lee's contribution, above):

[T]o speak in terms of social utility alone [or of nurturing meaningful relationships alone] is to leave out something vital in our conceptions of these practices [...] to forget that these practices [which include inflicting deserved suffering, and believing that inflicting deserved suffering can add value to organic wholes], and their reception, the reactions to them, really are expressions of our moral attitudes and not merely devices we

115. *Id.* at 988.

116. ZAIBERT, *supra* note 6, at 70.

117. *Id.* at 71.

118. Chiesa, *supra* note 108, at 988.

119. ZAIBERT, *supra* note 6, at 69.

calculatingly employ for regulative purposes. Our practices do not merely exploit our natures, they express them.¹²⁰

Independently of our disagreement regarding how to interpret Strawson, I think that Chiesa employs a specific type of argument I have discussed (and criticized) elsewhere. In the context of discussing Hyman Gross's misgivings about the criminal justice system, I called this type of argument "provocative, and, to a point, persuasive"—and this is how it continues to strike me now that Chiesa deploys it.¹²¹ This is not surprising, since Chiesa's argument and Gross's arguments are indeed very similar. Gross quotes Justice Brennan's view that:

[T]he accused during a criminal prosecution has at stake interests of immense importance, both because of the possibility he may lose his liberty upon conviction and because of the certainty that he would be stigmatized by conviction. Accordingly, a society that values the good name and freedom of every individual should not condemn a man for commission of a crime when there is reasonable doubt about his guilt.¹²²

And then Gross attempts to export this view about convictions in a court of law into the way in which we discuss (and come to accept or reject) philosophical theories: "It seems plain enough that for the very same reason it would be wrong to condemn anyone to criminal punishment unless criminal punishment itself were *justifiable beyond a reasonable doubt*."¹²³

Appealing to the same sort of "Supreme Court's reasonable doubt jurisprudence" as did Gross, and exploring the procedural asymmetry between inculcation and exculpation, Chiesa seeks to launch his "selective compatibilism."¹²⁴ But as I noted in the context of my discussion of Gross: there are "important differences between adjudicating guilt in a court of law and ascertaining the merits of a theory

120. *Id.*

121. Leo Zaibert, *Of Normal Human Sympathies and Clear Consciences: Comments on Hyman Gross's Crime and Punishment: A Concise Moral Critique*, 10 CRIM. L. & PHIL. 91, 93 (2016).

122. *Id.*

123. *Id.*

124. Chiesa, *supra* note 108, at 994, 1000.

(philosophical or otherwise).¹²⁵ I sought to illustrate the disanalogy by considering the fact that:

[T]here is plenty of discussion about the best procedure for treating inguinal hernias, or about how best to combat inflation. Yet no one doubts that inguinal hernias ought sometimes to be treated, or that they are in fact sometimes treated correctly—and the same holds for economic measures combating inflation.¹²⁶

I think that these points apply to Chiesa in just the same way: although provocative and to an extent persuasive, Chiesa's argument just does not hold.

Finally, I wish to discuss another of Chiesa's views: that his "selective compatibilism" leads to what punishment scholars call a "mixed theory of punishment."¹²⁷ This is odd first in that proponents of these mixed theories rarely invoke anything resembling selective compatibilism when championing their mixed justifications. But, secondly and more importantly, there is no obvious connection between these two theories. Mixed theories are often pluralistic approaches to punishment in which a variety of factors are allowed to play roles regarding the justification of punishment. As I conceded above, in the context of my discussion of Cahill's contribution, there *are* good reasons for embracing this sort of pluralism. But they need not derive from any grand metaphysical theory about the problem of free will. I, too, am a pluralist (though in at least two different senses, as explained above)—and, again, I explicitly avoid this metaphysical debate. The difference between classical compatibilism and "selective" compatibilism simply is that while the former bypasses the metaphysical discussion altogether, the latter bypasses it only sometimes, when it is convenient to bypass it. Neither form of compatibilism is specially linked to mixed justifications of punishment.

Chiesa correctly suggests that it would be fallacious to assume that the fact that believing something may be useful, or soothing, or strategically wise, is not enough to claim that it is true. And he is of course right: belief, at least in the standard case, does aim at truth. But I am afraid that it may be that Chiesa himself, via his selective compatibilism, is guilty of making this fallacious move. Either

125. Zaibert, *supra* note 121, at 93.

126. *Id.*

127. Chiesa, *supra* note 108, at 1000–01.

compatibilism—that is, the view that determinism is compatible with moral responsibility—is true or it is not. Period. Whatever good reasons we may have to be pluralists about punishment, or to be particularly circumspect when we punish—and these are weighty and many, as I endeavor to show in *Rethinking Punishment*—they have nothing to do with the truth of compatibilism. And it looks to me as if the reason why Chiesa would endorse or not endorse compatibilism (“selectively”) is whether or not it is convenient (theoretically speaking) for him to do so.

From the *normative* perspective I am quite sympathetic to the spirit that motivates some of Chiesa’s goals. For example, I share his intuition that it is worse to punish innocent defendants than to acquit guilty defendants.

Furthermore, the idea that we should be particularly circumspect when it comes to punishing is one of the main conclusions of *Rethinking Punishment*.¹²⁸ But this in no way commits us to endorsing what I have dubbed Parfit’s “compatibilism à la carte” or to what Chiesa calls “selective compatibilism.”¹²⁹ If my criticisms of Parfit and of Chiesa are on the right track, we are better off not endorsing those positions.

V. YANKAH, LEAVING CHILDHOOD BEHIND,
AND MISSING THE MORAL FOREST FOR THE POLITICAL TREE

In his far-reaching and insightful response, Yankah identifies “two” things about my views that “nag[]” him.¹³⁰ But it would be a mistake to read these as corresponding to only two criticisms, or to two small criticisms. Rather, these are best seen as mere headings, under which a number of concerns can be detected. I can only discuss what I take to be Yankah’s main worries (though, *en passant*, I may address *some* of Yankah’s other, sundry worries). First, Yankah is worried about the “robustness” of my intuitions, in particular about my alleged intuition that we can see “[deserved] suffering as a good.”¹³¹ Second, Yankah is worried about my approaching punishment morally, rather than politically.¹³²

128. See generally ZAIBERT, *supra* note 6.

129. Chiesa, *supra* note 108, at 979, 1000.

130. Ekow N. Yankah, *The Place of Retributivism in Punishment*, 71 RUTGERS U. L. REV. 1003, 1006–07 (2020) (discussing LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018)).

131. *Id.*

132. *Id.*

Regarding the first of Yankah's worries I must begin by noting that my position is not that deserved suffering is good, but that suffering's *default* value (its default badness, in fact) *can* change, depending on the context. For my purposes, desert is among the most important things that can alter the default badness of suffering: when suffering is deserved, the organic whole of which it is a part typically (though not necessarily) gains in goodness.¹³³ This aspect of my discussion draws considerably from G. E. Moore (from whom I borrow the framework of organic wholes)¹³⁴ and from Jonathan Dancy (from whom I borrow the framework of moral particularism, and the very notion of "default" value).¹³⁵

I doubt that this correct rendering of my position is likely to convince Yankah, and I suspect that he will continue to find my intuition problematically "fragil[e]":¹³⁶ he will simply reformulate his objection, attacking not the intuition that "suffering is good," but the intuition that "deserved suffering can add goodness to an organic whole."¹³⁷ The substantive response to the reformulated objection is that, even if this *view* of mine is properly called an "intuition," it is hard to avoid appeals to intuitions altogether. This is why in the book I admitted that "I, *just like my opponents*, rel[y], at some point or another, on some (ineliminable) appeal to intuition."¹³⁸ To assume that we can dispense with *all* intuitions is unrealistic, or to borrow Yankah's frequent term in his contribution (and to which I will soon return), "childish." Precisely because I wish to limit my reliance on intuitions, in the book I have tried to avoid the piecemeal approach of intuition pumps based on relatively isolated and remarkably contrived hypotheticals, and have instead sought to systematically ground intuitions on more comprehensive moral doctrines. In that vein, I have suggested that to disagree with my alleged "intuitions" is to disagree quite systemically with quite a lot, including the very importance of *context* for moral deliberation.¹³⁹

133. I leave aside the discussion of suffering *qua* part (as opposed to the whole to which it belongs), which itself changes value.

134. ZAIBERT, *supra* note 6, at 33.

135. *See id.* at 33–35.

136. Yankah, *supra* note 130, at 1008.

137. ZAIBERT, *supra* note 6, at 69.

138. *See id.* at 76.

139. This is why I not only presented my own arguments, but I also mobilized arguments by Isaiah Berlin, Jon Elster, John Rawls, Bertrand Russell, and P. S. Strawson to this effect. *See generally id.* at 71–86.

None of this, of course, *proves* that my intuitions are correct and that Yankah's are not—whatever “proving” the correctness of intuitions may turn out to be. Yankah appears to have great confidence in the role that empirical research may play in the current discussion. I will mention two reasons I do not share his confidence. First, and for whatever it may be worth, it is an easily verifiable empirical matter that *my* intuition, whereby “sometimes people should get the punishment they deserve—and sometimes simply because they deserve it,”¹⁴⁰ really is widely shared. Second, the empirical evidence that Yankah brings to bear on this discussion is orthogonal to my aims. Yankah's empirical evidence appears to suggest that many subjects derive *more* satisfaction from believing that the punishment they have inflicted served some communicative goal than they do merely from inflicting punishment as such. But since I explicitly grant that to get the wrongdoer to not only suffer (deservedly, as a response and in proportion to her wrongdoing, etc.), but to also understand the nature of her wrongdoing is often *more* valuable than to simply make her suffer, I am neither surprised nor worried by such findings. I am not at all inimical to the communicative goals of punishment, and in the book I refer to the communicative approach as “the most promising contemporary approach to the justification of punishment.”¹⁴¹

Thus, I fail to see why Yankah concludes that perhaps I may be “moving too quickly” when I claim that there is some value in wrongdoers getting the punishment that they deserve.¹⁴² I have defended the view that inflicting deserved suffering may create a more valuable organic whole than one that contains impunity. And holding the view that there may be value in deserved suffering as such in no way precludes me from holding the *utterly independent view* that it would be even more valuable if the deserved suffering were to also serve communicative (or other worthy) goals.

In any event, Yankah's skepticism about *some* intuitions influences his emotional responses to the possible distributions of good and bad things among good and bad people. Yankah claims that what he experiences when he learns of bad things happening to good people (or when good things happening to bad people) is “a sense of not only regret

140. *Id.* at 76.

141. *Id.* at 28.

142. Yankah, *supra* note 130, at 1008.

but some nagging sense of unfairness.”¹⁴³ Even more tellingly, “when something serendipitously good happens to the virtuous,” Yankah claims to feel “touched with child-like wonder.”¹⁴⁴ While he agrees that “a natural part of each of us is outraged when the vicious prosper,” and that a world in which suffering is distributed randomly is “disappointing” or even “destabilizing” or “terrifying,” he nonetheless thinks that “one sobering part of maturing is the realization that many things in life are difficult to gather into a coherent narrative that responds easily to vice and virtue.”¹⁴⁵ In other words, another part of us—the “mature” part, presumably—would learn to either not be outraged (or disappointed, destabilized, terrified, etc.) or to somehow manage this outrage.¹⁴⁶ When addressing the fact that sometimes “some cheaters really do win” and the fact that “feckless conmen can become presidents” Yankah frames his discussion in terms of how to explain this “to a child.”¹⁴⁷ The appeal to children is a conspicuous leitmotif in Yankah’s piece.

I find it odd that Yankah would so cavalierly dismiss this discussion—the discussion of the meaning of life, no less—as childish. After all, he explicitly recognizes that religions all around the world try to cope with ways in which injustice makes our search of meaning in the world more complicated.¹⁴⁸ Furthermore, Yankah rightly claims that this impulse is “not confined to the philosophically unsophisticated”: Yankah claims that Kant appeals to God “as the ultimate backstop to make good his account of desert.”¹⁴⁹ And utterly secular retributivists, say like Michael Moore, have explicitly grounded some retributive urges on a secular (and adult) desire to see justice served for its own sake, and in *this* world.¹⁵⁰ Either these admissions concerning the widespread appeal of the search for meaning suggest that Yankah sees both religious positions and the argument that he attributes to the philosophically sophisticated like Kant (or, in my example, like Moore), as somehow childish, or the admissions in fact militate against Yankah’s overall line that this search for meaning in life is childish.

143. *Id.*

144. *Id.*

145. *Id.* at 1005.

146. *See id.*

147. *Id.*

148. *See id.*

149. *Id.*

150. *See* MICHAEL S. MOORE, *PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW* 152. (Oxford: Oxford University Press) (1997).

Unlike Yankah, I find nothing childish about being outraged by the outrageous, or in striving to make the world a more just place. One need not, of course, assume that this striving can ever be “finished,” or that we should hope to one day be able to gather our moral experiences together in a completely just Grand Narrative—that assumption would be childish indeed. But that assumption is evidently at odds with some of the central tenets in *Rethinking Punishment*, since the book is committed to underscoring the extraordinary complexity of moral life.¹⁵¹ My bringing punishment theory into contact with theodicies, for example, seeks to highlight precisely how these concerns have preoccupied many of our ablest intellects throughout history. Any list of authors would be capricious and incomplete: choose any thinker who has ever thought about values, and you are likely to find someone also concerned with the prevalence of injustice in the world, and with ways of responding to it. There is nothing childish about their concerns; and my view is that contemporary punishment theory would benefit from thinking more systematically, or at least more deeply, about these foundational issues.

When Yankah deals with “intuitive” emotional responses in cases in which bad things happen to bad people—that is, in cases that can be deemed *retributive*—he does not merely dismiss them as atavistic remnants of our immature selves: he is unsure he even experiences them.¹⁵² By way of illustrating the fact that he does not hold these intuitions “robustly or coherently,” Yankah discusses what he experiences when he learns of a vicious person suffering deservedly, and he confesses that he is “unsure” that he feels “a strong sense that something good or valuable has occurred.”¹⁵³ So, for example, “learning the historical fact that Stalin suffered a major stroke or Kim Il Sung had a heart attack” leave Yankah “cold.”¹⁵⁴

These examples are a bit tendentious: one would be less cold, I am sure, if Stalin or Sung had in fact been *punished* for their wrongdoing, rather than merely having suffered unrelated ills. Nowhere do I deny this point; my point is that even an unrelated ill, may, in and of itself, add value to the organic whole of which it is a part. But the examples, together with some of Yankah’s other views, paint a fuller picture: Yankah is not at all clear about the status of “desert as a moral

151. ZAIBERT, *supra* note 6, at i.

152. Yankah, *supra* note 130, at 1004.

153. *Id.* at 1008.

154. *Id.*

matter,”¹⁵⁵ and he believes that “suffering that simply befalls others is an unlikely bearer of moral value.”¹⁵⁶ In other words, Yankah appears to doubt that there may be *any* value in *any* suffering unless it is accompanied by (or conducive to) some good consequence, and this strikes me as an undeniable endorsement of the axiology of (classical) utilitarianism. Despite generously praising aspects of *Rethinking Punishment*, and in particular the unusual conversations that it invites, Yankah, it seems to me, refuses to have a conversation about axiology.

Yankah is left cold when bad things happen to those who deserve them because “*suffering*, as opposed to *punishment*, is not obviously desert apt.”¹⁵⁷ While it is unfair to expect Yankah to have provided a full-fledged account of desert-aptness in the context of a brief commentary on my book, he does say so little about the idea that it is very hard to discern what he may have in mind. Yankah deserves to be recognized as a better chess player than I am. Why? Because he beats me much more often than I beat him, etc. Is chess-playing not “desert-apt”? My arguments in the book suggest that it is, and that there is some value in Yankah receiving the recognition that he deserves, even absent any further consequences of this recognition.¹⁵⁸

Consider, however, a seeming parallel: Lake Garda deserves to be recognized as one of the most beautiful places on earth. Why? Because it *is* extraordinarily beautiful, etc.¹⁵⁹ There is something unconvincing about this case. Places—and inanimate objects in general—may not *deserve* anything at all: *they* may not be desert-apt. But surely it is at least *fitting* that Lake Garda be so recognized. And I have argued that fittingness and desert are very close relatives (desert being a robustly *moral* form of fittingness).¹⁶⁰

Still, while it may be true that although recognizing the beauty of Lake Garda may be fitting, the lake does not deserve anything at all because the lake is not desert-apt. Only rational beings, and typically in light of their *actions*, strike me as desert-apt: unlike the beauty of Lake

155. *Id.* at 1005.

156. *Id.* at 1009.

157. *Id.* at 1008.

158. ZAIBERT, *supra* note 6, at 73.

159. See Mariam Thalos, *Truth Deserves to be Believed*, 88 PHIL. 179, 195 (2013); see also BERNARD WILLIAMS, *TRUTH AND TRUTHFULNESS: AN ESSAY IN GENEALOGY* (2004) (review 110 ff, in particular).

160. See Leo Zaibert, *The Fitting, the Deserving, and the Beautiful*, 3 J. MORAL PHIL. 331 (2006).

Garda, the atrocities committed by tyrants, and Yankah's chess playing, strike me as unproblematically desert-apt. But these are *my* views on an admittedly difficult topic; *Yankah's* invocation of desert aptness strikes me as mysterious. And I cannot avoid suspecting that there may be some cherry-picking involved in what he finds desert-apt and what he does not, of the sort I criticize in the "Appendix," and in my discussion of Chiesa's contribution above—*desert-aptness à la carte*, if you will.¹⁶¹

Despite the halo of mystery surrounding Yankah's talk of desert-aptness, it is clear that Yankah believes that to talk about desert in cases that are not "desert apt" is, yet again, childish.¹⁶² He claims that "[o]nce we leave childhood behind, we recognize[] that whether a brutal dictator versus a kind humanitarian is stricken with cancer is only metaphorically described as something they deserve."¹⁶³ But, again, I fail to see anything necessarily childish about, echoing Weber, being affected by the "incongruity between destiny and merit."¹⁶⁴ There is nothing necessarily childish in being affected by the contemplation of feckless conmen who become president, or by dishonesty, cruelty and vice triumphing over truth, kindness, and virtue. Being so affected is nothing short of being alive to the human predicament.

Yankah claims that what is not desert-apt is "suffering," as opposed to "punishment" which (presumably) is desert-apt.¹⁶⁵ The very opposition that Yankah has in mind here escapes me, since by "suffering" I simply refer to the hard treatment that is *conceptually* inseparable from punishment.¹⁶⁶ This brings me to the second of Yankah's worries. He believes that "Zaibert's passion for informing punishment theory with moral theory leaves him insufficiently sensitive to the political nature of legal punishment."¹⁶⁷ Why should my invitation to think more deeply about the morality of punishment desensitize me to anything? Nowhere do I deny the almost tautological claim that legal punishment is a political phenomenon ("legal punishment" is in this context almost a synonym of "political punishment"). Neither do I deny that the

161. ZAIBERT, *supra* note 6, at 143–47; *see supra* part IV.

162. Yankah, *supra* note 130, at 1008–09.

163. *Id.*

164. H. H. GERTH & C. WRIGHT MILLS, FROM MAX WEBER: ESSAYS ON SOCIOLOGY 275 (1946).

165. Yankah, *supra* note 130, at 1009–10.

166. *See* discussion *supra* part II.

167. Yankah, *supra* note 130, at 1007.

justification of political phenomena may be different from the justification of the non-political versions of those phenomena. My view, rather, is that even our political discussion of legal punishment can *benefit* from taking a broader and deeper look at punishment, including its morality. This strategy did not begin with *Rethinking Punishment*, although it is here that I focus most centrally on the justification of punishment. Already in my 2006 *Punishment and Retribution* I remarked on the astounding tendency of contemporary punishment theorists who even deny that, purely as a conceptual matter, punishment can *occur* outside the context of the state.¹⁶⁸

To an extent, Yankah is pushing against an open door here: he recognizes that my views are perfectly compatible with a number of options when it comes to the (political) justification of legal punishment.¹⁶⁹ But he believes that my views are somehow “simply compatible” with those options, and he finds this “simpl[e] compatib[ility]” insufficient.¹⁷⁰ The reason for this insufficiency seems to be that Yankah believes that when it comes to the justification of legal punishment, “[P]unishment as a legal practice respond[s] to a set of political justifications and set[s] aside the purely moral value of retributivism.”¹⁷¹ In other words, the problem Yankah sees in my view is not quite one of insufficiency, but actually one of misguidedness: it is misguided to investigate the morality of punishment.

Yankah’s position is thus revealed in its remarkable radicalness: he appears to model the relation between moral and political philosophy on the relation between oil and water—they do not mix. This is particularly odd if one realizes that political philosophy is but a branch of moral philosophy. Why should the “purely moral” value of desert be fully “set aside” (and not enriched by, or brought to bear in conjunction with, a number of “political” values)?¹⁷² This radicalness explains a number of peculiar aspects of Yankah’s position. For example, Yankah claims that “when Kant turned his attention to legal punishment, retributivist urges faded from the picture.”¹⁷³

168. See ZAIBERT, *supra* note 27 (objecting to Guyora Binder’s view, a view that Yankah cites approvingly).

169. Yankah, *supra* note 130, at 1006–07.

170. *Id.* at 1015.

171. *Id.* at 1015.

172. *Id.* at 1015–16.

173. *Id.* at 1014.

But to summarily dispose of Kant's retributivism as Yankah does, seems unwarranted. Kant's oeuvre contains plenty of evidence that directly challenges Yankah's take: from Kant's systematic defense of intrinsic value independently of any consequence (say, in the *Grundlegung*) to his famous stern warning in the *Rechtslehre* "woe to him who creeps through the serpent-windings of Utilitarianism to discover some advantage that may discharge him from the Justice of Punishment [i.e., from retributivism]."¹⁷⁴ It is hard to accept that these essential aspects of Kant's philosophy "fade[] from the picture" when we are in the political sphere.¹⁷⁵

Similarly, Yankah too casually assumes that to recognize value without thereby expecting an obligation to promote it is "the most natural thing in the world."¹⁷⁶ Yankah's naturalness thesis is not true of most moral philosophers I have read, or of most non-specialists with whom I have interacted. On the contrary, the overwhelmingly typical view is that, at least in principle, if something is of value then there is, at least *prima facie*, an obligation to bring it about. Curiously, Yankah himself seems to forget about his own naturalness thesis, for elsewhere in his contribution he endorses its opposite: "In the normal case, that something is of value implies that the thing (value) should be promoted."¹⁷⁷ And to elucidate—or to begin to elucidate—how to travel from the axiological to the deontic is (as we have seen above) one of the central goals of *Rethinking Punishment*.

Even Yankah's peculiar take on some of my views on forgiveness can perhaps be connected to his exaggerated distinction between moral and political philosophy. Yankah reminds me that it would be a mistake to see "forgiveness as an unmitigated good," to think of it as "always . . . praiseworthy"; and he warns against "forgiv[ing] too readily."¹⁷⁸ But these are views that I explicitly oppose in the book. In fact, a central theme in *Rethinking Punishment* is that forgiveness qua response to wrongdoing stands in need of a justification (thus not always

174. IMMANUEL KANT, *THE PHILOSOPHY OF LAW* 195 (W. Hastie, trans., Edinburgh: T. & T. Clark 1887); *see also* IMMANUEL KANT, *GROUNDWORK OF THE METAPHYSICS OF MORALS* 7 *passim* (Mary Gregor trans., Cambridge: Cambridge Univ. Press 2006).

175. Yankah, *supra* note 130, at 1014 (citing IMMANUEL KANT, *METAPHYSICAL ELEMENTS OF JUSTICE* 140, 141 (John Ladd trans., Hackett Publishing Company 2d ed. 1965)).

176. Yankah, *supra* note 130, at 1004.

177. *Id.* at 1013.

178. *Id.* at 1006.

praiseworthy or good), just like punishment. Paying more attention to matters of *justification* is one of the main lessons that forgiveness theorists can learn from punishment theorists. For their part, punishment theorists stand to learn from forgiveness theorists a greater appreciation of the intricacies of both responses to wrongdoing. Punishment (legal and otherwise), like forgiveness, is *dilemmatic*—a point neither frequently nor systematically discussed by punishment theorists—and which, finding it full of promise, I discuss at length in *Rethinking Punishment*.

VI. CONCLUSION

One of my main goals in *Rethinking Punishment* is to show that paying attention to moral philosophy may teach important lessons about the justification of punishment (including *legal* punishment). But to a large extent my emphasis is not on moral philosophy in general, but on a relatively recent and very *particular* approach to moral philosophy.¹⁷⁹ Admittedly, the approach is not widely accepted amongst moral philosophers, which is not surprising, since it is very critical of the typical ways in which moral philosophy is carried out. I detect elements of this approach in Berlin's work, in Brentano's, in Dancy's, in DeWijze's, and in Stocker's (amongst others), in ways often connected precisely to the dilemmatic nature of punishment (in *any* context), and to the way in which punishment leaves remainders *even when justified*. Above all, however, my approach is inspired by Williams's work; as I said at the outset of *Rethinking Punishment*, "I seek to deploy, within the specific context of punishment, criticisms similar to those more general criticisms Williams deployed against both utilitarianism and what he called "the morality system."¹⁸⁰ I am afraid that the significance of the *particular* approach to moral philosophy that I champion in the book is greater than my critics seem to realize.

179. In the book I emphasize how recent these sorts of developments in moral philosophy really are, and I quote Alastair Macintyre's felicitous way of putting it: "[I]f one were to publish two volumes, the first containing the entire preceding philosophical literature dealing with this topic, broadly construed, from Plato to W. D. Ross through Gregory, Aquinas, Kant, Hegel, Mill, Sidgwick, and Bradley while the second was devoted to the publications of the last thirty years, the second volume would be by far the larger." ZAIBERT, *supra* note 6, at 24 (quoting Alasdair Macintyre, *Ethics, in 2 ETHICS AND POLITICS: SELECTED ESSAYS* 85 (Cambridge: Cambridge University Press 2006)).

180. ZAIBERT, *supra* note 6, at 2.

It is not my intention to poach Williams's larger-than-life reputation. First, because despite my viewing him as the most insightful moral philosopher of at least the last hundred years, I have some noteworthy disagreements with him. (Particularly relevant is my rather more positive assessment of G. E. Moore's meta-ethical positions, for example.¹⁸¹) Second, because some of my interpretations of Williams may be rather idiosyncratic, thus casting doubt on the extent to which regarding those aspects my book can really be seen as based on (the standard interpretation of) Williams's work. (My Williams is much less relativistic than the Williams of others, and is less reliant on Greek tragedy and ethics, or on Nietzsche, for example.)

But I can admit to having inherited a certain problem from Williams. Williams's approach is often criticized for being somehow too "anti-theory," or even "romantic" (in the senses of anti-Enlightenment, or irrationalist).¹⁸² Notwithstanding the perspicuity of Williams's criticisms of others, some of Williams's critics have asked, almost in exasperation: But what, *concretely*, does he have to offer instead? Some of my critics (here and elsewhere)¹⁸³ ask a similar question of *Rethinking Punishment*: What, *concretely*, does it have to offer?¹⁸⁴

While it may not be concrete enough for some, inviting punishment theorists to think more deeply about the complexity of moral life, about the inescapably dilemmatic nature of punishment, about the theoretical problem of punishment (and the peculiar type of pluralism it engenders), and about the types of remainders that justified punishment generates, strikes me as likely to advance the field. Truth matters, and I believe that these points deriving from the often-overlooked complexity of moral life *are* true. But, if nothing else, these considerations may alert punishment theorists about mistakes to avoid.

181. See *id.* at 58 n.84.

182. See *id.* at 166–68 (referencing several works discussing Williams's work).

183. See generally *Leo Zaibert: Rethinking Punishment*, C4EJOURNAL: PERSP. ON ETHICS, <https://c4ejournal.net/category/leo-zaibert-rethinking-punishment/> (last visited Nov. 2, 2019) (showing a book forum of reviews of *Rethinking Punishment*); *16 Eunomia No. 16 (Abril 2019–Septiembre 2019)*, UNIVERSIDAD CARLOS III DE MADRID, <https://e-revistas.uc3m.es/index.php/EUNOM/issue/view/573> (last visited Nov. 2, 2019) (referencing a Spanish magazine which includes reviews of *Rethinking Punishment*).

184. See ZAIBERT, *supra* note 6, at 165–71 (discussing whether this sort of question may not be even more pertinent within the specific context of punishment theory than in the context of moral philosophy in general).

Consider, for example, Yankah's favored "republican justification of punishment": punishment is justified when it effectively "rebind[s] the offender to civic society."¹⁸⁵ I am too much of a liberal to see the value of communal bonds, and much less of *enforced* communal bonds, as sharply as Yankah sees it. Furthermore, *if* communal bonds are indeed valuable, it is not clear to me what, if anything, the connection between such value and our responses to *wrongdoing* as such is supposed to be. (Should we not make sure that *everyone*—and not just wrongdoers—is bound or rebound to civic society?) But, even putting these matters aside, to the extent that this justification of punishment ignores the dilemmatic nature of punishment and the remainders it generates, it strikes me as risking making some of the mistakes I highlight in *Rethinking Punishment*. Yankah may have simply substituted "utility" or "suffering reduction" or "desert" with "rebinding offenders to their community," without overcoming the inadequate *structure* of traditional justifications of punishment.¹⁸⁶ After all, it seems as if, by Yankah's lights, provided wrongdoers are reconnected to the values of their communities, there would be nothing to lament—no agent regret, no tragic remorse, not even garden-variety sadness—about the *suffering* that punishment, as a matter of definition, seeks to inflict upon them in the process of rebinding them to their community.¹⁸⁷

So, when critics express a certain dissatisfaction as to the humbleness (or lack of concreteness, etc.) of my view—for example, when Yankah claims that "a pluralistic theory that insists on the value of retributivism only to note that this value always collapses when in competition with distinctly political concerns risks rendering retributivism trivial"—they are missing something crucial about what my championing of this very *particular* view of moral philosophy entails.¹⁸⁸ It should by now be clear that this talk of "collapsing values" is *precisely* what my view does not claim.¹⁸⁹ The value of deserved suffering *never* collapses (either to political or to non-political concerns):

185. See Yankah, *supra* note 130, at 1019.

186. See *id.*

187. As I argued above, mainly in the context of my discussion of Lee's contribution, fitting responses to wrongdoing need not be punitive. See discussion *supra* part II. I suspect that this is true of Yankah's "rebinding," but he himself advertises it as justification of *punishment*. See Yankah, *supra* note 130 at 1020. And as long as it is supposed to be of *punishment* then it must contend with its hard treatment aspect, i.e., with suffering. *Id.*

188. Yankah, *supra* note 130 at 1016.

189. See *id.*

this value always *remains*, even when it is “defeated,” just as the value of its merciful remission *never* collapses, even when “defeated.” *This* is the theoretical problem of punishment, from which a special form of pluralism is generated, and from which admittedly complex axiological and emotional situations are engendered. Taking the theoretical problem of punishment seriously entails not eschewing the intimate interaction between moral and political philosophy, and recognizing the complexity of moral life.

This is why, when discussing Cahill I insisted on the uniqueness of the pluralism that flows from *this* complexity,¹⁹⁰ when discussing Lee I insisted on the importance of *real* justice and *real* punitive responses even when they involve suffering,¹⁹¹ when discussing Bergelson I insisted on the depth and density of the remainders of which I speak,¹⁹² and when discussing Chiesa I insisted on the ways in which determinism flattens moral reality.¹⁹³ None of this is meant to deny a certain weakness that I concede I may have inherited from Williams: I *do* owe a story as to how exactly this theoretical complexity is to play out in concrete cases. But it is meant to underscore that arguing for this theoretical complexity is already a significant step, particularly in light of the frequency and ease with which contemporary criminal law theorists overlook or dismiss it.

190. See discussion *supra* part I.

191. See discussion *supra* part II.

192. See discussion *supra* part III.

193. See discussion *supra* part IV.