


RESEARCH ARTICLE

The Dogma of Opposing Welfare and Retribution

Leora Dahan Katz 

Faculty of Law, Hebrew University of Jerusalem, Jerusalem, Israel
Email: leora.dahankatz@mail.huji.ac.il

Abstract

There is a common refrain in the literature on punishment that presumes the mutual exclusivity of defending retribution and adopting a humanistic or welfare-oriented outlook. The refrain, that if we want to be humane, or care about human welfare, we *must* abandon retributive punishment, anger, and resentment is readily repeated, endorsed, and relied upon. This article suggests that this opposition is false: retribution and welfare-orientation can not only be endorsed concomitantly, but are complimentary projects, and may even be grounded in the same normative basis, such that if we endorse one we are already committed to ideas that ground reason to care about the other.

My primary target will be claims that aim to undermine retributivism by demonstrating the desirability of welfare-orientation. If both can live together, demonstrating the attractiveness of one goes nowhere toward displacing the other. Further, establishing this claim invites further inquiry into classic questions about the “barbaric,” or “morally repugnant” credentials of retributivism. Confronting these claims will elucidate the consistency of adopting both retributive and welfare-oriented views, which, I suggest, can be jointly adopted and pursued.

Keywords: Retributivism; Punishment; Welfare; Consequentialism; Justification; Mass incarceration; Suffering

1. Introduction

Retributivism is commonly regarded as a position that glorifies human suffering, even if restricted to the suffering of the deserving. Welfare-oriented approaches, on the other hand, aim to preserve, enhance, promote, maximize, etc. human welfare or well-being. Taken in this oversimplified form, the two appear to be directly at odds with one another, where retributivism aims at human suffering while welfare-oriented approaches aim at its alleviation. A less caricatured look at the positions, however, demonstrates that this opposition presents a false dichotomy. Yet, in the literature on punishment, this opposition is often relied on not only to advance false claims about retributive theory, but to argue for its displacement as a justification of punishment given the conflict.

Recent work on punishment that illustrates this common refrain includes monographs by Martha Nussbaum, Erin Kelly, and Vincent Chiao, which all, in defending

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (<http://creativecommons.org/licenses/by-nc/4.0>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

a consequentialist outlook on punishment, assume a basic dichotomy between the aims of retribution and welfare.¹ This dichotomy then serves as a basis to argue against retributivism as a viable justification of punishment. In the following, I aim to challenge this received wisdom by looking more carefully at these antiretributive arguments and to demonstrate that the proposed welfare-oriented aims may be achieved without resorting to the abandonment of retribution. I present a picture of complementarity that invites welcoming the welfare-based criticism offered by critics, but denies that these provide grounds for rejecting retributivism. Further, I argue that the normative core that underlies interest in human welfare can itself be understood as underlying the very grounds of retributivism, allowing not only for potential complementarity, but for more robust congruence.

Section II sets forth the relevant line of objection against retributivism, which has recently received much support. Disentangling various arguments made within this line of objection, Section III identifies and addresses two central arguments advanced in the literature, which prove false: (1) that retributivism entails the rejection of welfare-promotion, and (2) that retributivists cannot object to harsh and cruel punishment nor to blatant injustices such as the injustices of mass incarceration. The section concludes by offering reasons to conceive of the projects of welfare-promotion and retributive punishment not only as potentially complementary projects, but as projects that emerge from the same set of normative commitments, though it does not aim to offer conclusive arguments to this effect.

Section IV then attempts to excavate the underlying concerns that may rest beneath the surface of these arguments, offering alternative welfare-oriented arguments that may be and have been raised against retributivism: primarily, the traditional concern with retributivism's endorsement of the goodness of suffering. It elucidates the myriad ways in which retributive theories need not be (and often are not) committed to the intrinsic goodness claim, and proposes that in fact, retributive views generally do not attribute intrinsic goodness to suffering. Section V concludes that while tensions remain, the proposed line of argument against retributivism fails.²

II. The Line of Objection

Contemporary retributive thought has flourished since the 1970s. The so-called retributivist revival has seen an outpouring of in-depth engagement with different strains of retributive thought, which attempt to carefully address myriad aspects of punitivity and defend an ultimately desert-based justification of punishment.³

¹MARTHA C. NUSSBAUM, *ANGER AND FORGIVENESS* (2016); ERIN I. KELLY, *THE LIMITS OF BLAME* (2018); VINCENT CHIAO, *CRIMINAL LAW IN THE AGE OF THE ADMINISTRATIVE STATE* (2018).

²Note that much of what follows does not directly address the category of views known as "legal retributivism," but is rather focused on a defense of moral retributivism.

³See, e.g., Herbert Morris, *Persons and Punishment*, 52 *MONIST* 475 (1968); Jeffrie Murphy, *Marxism and Retribution*, 2 *PHIL. & PUB. AFFS.* 217 (1973); ANDREW VON HIRSCH, *DOING JUSTICE* (1976); MICHAEL S. MOORE, *PLACING BLAME* (1997); Jean Hampton, *Correcting Harms versus Righting Wrongs*, 39 *UCLA L. REV.* 1659 (1992); Richard Dagger, *Playing Fair with Punishment*, 103 *ETHICS* 473 (1993); Dan Markel, *What Might Retributive Justice Be?*, in *RETRIBUTIVISM: ESSAYS ON THEORY AND POLICY* 49 (Mark D. White ed., 2011); Douglas Husak, *Retributivism In Extremis*, 32 *LAW & PHIL.* 3 (2012); Mitchell Berman, *Rehabilitating*

While retributivism encompasses a plurality of views, central to this set of views is the contention that the punishment of the deserving is right or good, and that punishment is justified in terms of the desert of the agent.⁴ Retributivism is, of course, not alone in endorsing punishment, that is, in endorsing the deliberate imposition of burdens, deprivations, or suffering on wrongdoers in response to their wrongs.⁵ Despite the welfare-inhibiting implications of such endorsement, all nonabolitionist views, including retributivism's traditional consequentialist competitors, endorse such imposition.⁶ However, to ground punishment's justification, competitors rely on the nonretributive goods that punishment has the capacity to produce, such as deterrence, incapacitation, and rehabilitation, while retributivism alone relies on the normative force of desert.

From the outset, such revival has been accompanied by a broad range of critical reactions that have produced a counter-literature to match the rise of retributive theory.⁷ Recently, however, a particular line of objection has emerged as a dominant theme in antiretributivist writings, in the center of which stands the presumption that if we care about human well-being or welfare,⁸ retributivism must be rejected.⁹

Current interest in human welfare and well-being in the context of punishment follows a much-needed resurgence of political interest in the harshness of our punitive systems, not least of which the American criminal justice system. This includes concern with solitary confinement, life without parole, prison suicide rates, and more particular to the American context, mass incarceration and the reproduction of patterns of enslavement and racial discrimination, to name just a few ills. Public interest in the welfare of persons, including that of the incarcerated, has redirected attention toward the need to cure the ills of the system.¹⁰ These are welcome cries.

Retributivism, 32 LAW & PHIL. 83 (2013); Leora Dahan Katz, *Response Retributivism: Defending the Duty to Punish*, 40 LAW & PHIL. 585 (2021).

⁴See, e.g., C.L. TEN, CRIME, GUILT, AND PUNISHMENT (1987), at 46; Husak, *supra* note 3, at 4; R.A. DUFF, PUNISHMENT, COMMUNICATION, AND COMMUNITY (2001), at 3.

⁵See further Antony Flew, *The Justification of Punishment*, 29 PHILOSOPHY 85, 85–87 (1954); S.I. Benn, *An Approach to the Problems of Punishment*, 33 PHILOSOPHY 325 (1958); H.L.A. HART, PUNISHMENT AND RESPONSIBILITY (1968), at 4–5.

⁶For an elaboration on consequentialist versus nonconsequentialist versions of retributivism (the former of which aim to promote a retributive good), see Mitchell N. Berman, *Two Kinds of Retributivism*, in PHILOSOPHICAL FOUNDATIONS OF CRIMINAL LAW 452 (R.A. Duff & Stuart Green eds., 2011). For a nonconsequentialist, nonretributive justification of punishment, see, e.g., VICTOR TADROS, THE ENDS OF HARM: THE MORAL FOUNDATIONS OF CRIMINAL LAW (2011).

⁷See, e.g., TED HONDERICH, PUNISHMENT: THE SUPPOSED JUSTIFICATIONS (1970); J.L. Mackie, *Morality and the Retributive Emotions*, 1 CRIM. JUST. ETHICS 3 (1982); David Dolinko, *Three Mistakes of Retributivism*, 39 UCLA L. REV. 1623 (1991); Russ Shafer-Landau, *The Failure of Retributivism*, 82 PHIL. STUD. 289 (1996); DEIRDRE GOLASH, THE CASE AGAINST PUNISHMENT (2005); DAVID BOONIN, THE PROBLEM OF PUNISHMENT (2008).

⁸I use these terms interchangeably here.

⁹I refrain from framing this in terms of welfarism, since on some understandings, retributivism is definitionally excluded by welfarism. See note 13 below. Indicatively, the *Stanford Encyclopedia of Philosophy* entry on well-being states that while welfarism is a varied concept, “however it is understood, it does seem that welfarism poses a problem for those who believe that morality can require actions which benefit no one, and harm some, such as, for example, punishments intended to give individuals what they deserve.” Roger Crisp, *Well-Being*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Fall 2017), <https://plato.stanford.edu/archives/fall2017/entries/well-being/>.

¹⁰Such interest is often driven by alarms raised in light of economic rather than humanist concerns, though the latter have happily gained traction following the former.

And one hopes they will motivate change and reform and produce better and more humane systems of criminal law and punishment. Retributive theory can in fact provide a bulwark against many such ills, and can serve as a resource in the fight for their eradication.

However, reasons to protest and the targets of protest can and have unfortunately come apart. Not for the first time in history, the interest in care, compassion, mitigation, welfare, and well-being has led to an attack on retributivism. Retributivism is targeted as an ill or, more brazenly, a cause to be eradicated if we are to move forward toward a more humane criminal law.

Some of these attacks are attacks against rhetoric: that the language and dissemination of retributive ideas and concepts have generated harsh systems of criminal punishment, independent of the merits of retributive views. James Whitman and Erin Kelly, for example, attack the language of blame and retribution, arguing that “crying blame” does more harm than good, and is responsible for the excessively punitive state of American criminal justice.¹¹ They predict, though it appears no more than a mere (yet adamant) guess, that abandoning such language will make the system more humane and moderate. Here the argument is empirical, though not backed by empirical evidence beyond gestures toward correlations between the two, without any consideration of how the harshness of punishment correlates with other legitimating narratives and rhetorics of punishment.¹² But the merits of these arguments are to be empirically tested and are not the subject of this article.

More to the point, the arguments are also often an attack on substance: that the substance of retributivism and what it has to offer as a justification of punishment is at odds with valuing welfare, and that, as above, if human welfare or well-being is valuable, retributivism *must* be abandoned.¹³ It is to this set of arguments that this article turns to demonstrate that the constructed opposition has no basis. This point should be evident from the fact that many contemporary retributivists would demand the very reforms that critics seem to believe retributivism must be laid waste to, to promote. But it is not. This paper is therefore intended to set the record straight.

At the base of the paper lies the insight that promoting the well-being of persons, on the one hand, and retributive responses to wrongdoing, on the other, can arise from the same normative commitments, and should be viewed as complementary projects that emerge from a single worldview. The paper is directly occupied,

¹¹James Q. Whitman, *A Plea Against Retributivism*, 7 *BUFF. CRIM. L. REV.* 85 (2003); KELLY, *supra* note 1.

¹²E.g., the idiom of dangerousness, which might be supposed to be similarly “responsible” for the current state of the system; (though one should doubt whether either is responsible if responsibility requires a causal relation between the two). See, similarly, Douglas Husak, *Retributivism and Over-Punishment* 41 *Law and Philosophy* 169 (2022). See also Chad Flanders, *Retribution and Reform*, 70 *MD. L. REV.* 87, 106–107 (2010).

¹³Note that this article is not engaged with a strict welfarist (or utilitarian) objection to retributivism, which argues, based on a strict commitment to the exclusive value of welfare, that only contributions to welfare have the capacity to justify punishment. Instead, the identified line of objection relies on a more expansive approach, what might be thought of as “welfare orientation” to reject retributivism: it argues that to the extent that human welfare is an important moral good, to be valued, preserved, enhanced, promoted, maximized, etc.—i.e., to the extent that welfare promotion and humaneness are important moral aims—retributivism must be rejected. This is a far more threatening line of objection given that few would reject the moral significance of human welfare. It is this opposition that I claim is false and with which I take issue here.

however, with engaging with particular objections to retributivism and demonstrating that they fail, independent of whether the deeper normative picture presented is endorsed. It parses out the ways in which retributivism and welfare orientation are, contrary to what seems to be common belief, quite simply happy bedfellows.

III. Two False Objections

The project of this article is thus to demonstrate the consistency of welfare orientation and retribution. This consistency is denied at a number of points throughout the anti-retributive literature, which are not always kept apart. It will be important to clarify each of the relevant objections, which offer different focal points for grounding the rejection of retributivism, in order to adequately address the concerns each may raise. To anticipate, these include the claims that:

1. promoting (ex ante) welfare-oriented goals and projects requires rejecting retributivism;
2. opposing welfare-destroying and unjust practices, particularly mass incarceration, requires rejecting retributivism;
3. a commitment to human welfare excludes the possibility of endorsing retributivism.

In this section, I address the first two lines of argument, which I take to be weaker arguments against retributivism. I begin by illustrating examples of prominent work that promotes these arguments, taking some time to demonstrate the way the constructed opposition figures in arguments about punishment, before explaining the fallacies and misunderstandings that sustain them. In the next section, I identify what I take to be deeper underlying concerns about the tensions between retributivism and welfare-orientation that may underlie such rejections of retributivism. I argue that when pressed, these too come undone, and explore possible tensions that remain. Ultimately, I aim to demonstrate that this line of objection does not provide grounds for rejecting retributivism.

A. Retribution vs. Welfare-Promotion

1. Substantive Inconsistency

One claim prevalent in the contemporary antiretributive literature is the argument that to endorse welfare-oriented projects, we have to give up on retribution. Martha Nussbaum's *Anger and Forgiveness* is a prime example of the literature that trades in this false dichotomy. The book presents promoting welfare and justice, on the one hand, and retribution, on the other, as mutually exclusive options that are in direct conflict with one another, where opting for a future, welfare-oriented approach requires abandoning anger and retribution in favor of love and an exclusively future-looking program.¹⁴ While Nussbaum is interested in the emotional

¹⁴Nussbaum's angst is directed not only at anger as an emotional state but at retribution and the desire for retribution from the very first page. The terms are almost interchangeable for Nussbaum; or, more precisely, retribution is conceptually part of the idea of anger on her conception. NUSSBAUM, *supra* note 1.

side of anger, integral to her argument is a rejection of the normative basis of retributivism, which she takes to be an integral part of the concept of anger—“the idea that it would be a good thing if the wrongdoer suffers some bad consequence.”¹⁵ She conceives of forward-looking punishment as a *replacement* for anger and retribution, where deterrence is to be pursued “*rather than* payback.”¹⁶ The abandonment of the project of retribution is, on her view, necessary for the progression to a just, welfare-oriented resolution to the problem of wrongdoing.¹⁷

Vincent Chiao’s *Criminal Law in the Age of the Administrative State* also understands retributive commitments to be at odds with a welfare-oriented approach. While the monograph’s positive proposal is concerned with promoting a particular justice-oriented approach to criminal law, its negative proposal presumes that the adoption of any such approach requires the abandonment of retributive and other such desert-oriented, “private morality,” “backward-looking,” or “strict deontological” approaches to punishment.¹⁸ This is evident, for example, in Chiao’s analysis of *DeShaney v. Winnebago County Department of Social Services*,¹⁹ with which he opens and closes the first chapter.²⁰ In *DeShaney*, the Court addressed the case of

¹⁵*Id.*

¹⁶*Id.* at 4. [Italics not in the original.] Per Nussbaum, retribution makes care and love impossible. Whether Nussbaum intends to locate this mutual exclusivity within the moral goods at stake or within human psychology is never quite clear. The latter option will be discussed below.

¹⁷*Id.* at 3. For Nussbaum, the options of retribution and care are fundamentally opposed, whether as a conceptual, normative, or practical matter. Consider an example in which this presumed mutual exclusivity comes to the fore. Nussbaum envisions a case in the personal sphere (which she takes to be parallel in many respects to the realm of justice) wherein there is a stranger rape, and a friend of the victim has to decide how to react. Nussbaum describes the friend as facing a “three-pronged fork-in-the-road”: she can take one of two retributive options, the “road to payback” (referring to the fantasy of punishment securing restoration of the damage done) or the “road to status” (referring to an obsessive focus on the status of the offender and at downgrading the offender’s status relative to the victim)—two exhaustive variations on the retributive theme—or adopt a third alternative, which consists in what would be “really helpful going forward.” This third option consists in “creating future welfare,” e.g., by supporting her friend through therapy, raising campus awareness about safety precautions, etc. A combination of acting in ways that are supportive of one’s friend and of a retributive response is no option on the Nussbaumian picture. One has to choose: one is either a compassionate fan of welfare or a vindictive seeker of retribution. No balanced respondents, who can recognize both the independent significance of addressing the past as well as the significance of acting in view of the future are anywhere on the horizon. *Id.* at 28–29. There is much to say about what makes these arguments false. The options of the “magical” road to payback and the “narcissistic” road to status are clearly not the only options on the retributive table. They are themselves rather mocking caricatures of any retributive view that might be in the vicinity of the kinds of thinking she wants to reject. (Nussbaum later addresses what she calls “nuanced retributive” views, but ultimately takes these to rely on one of these two paths as well. *See id.* ch. 6.) But I leave it to particular defenses of retributive theories to lay out the more defensible options that might be available to the retributivist. For one such proposal, *see* Dahan Katz, *supra* note 3.

¹⁸While in one instance Chiao leaves open the possibility that there may be some residual role for retributive theory to play within the context of criminal law, the general tenor of the book as well as its particular arguments (especially in Chapters 1, 2, and 4) presume such opposition throughout. Nonetheless, Chiao is far more interested in his positive proposal and proposed theory of just institutions than in the negative proposal that sets strict deontological theories aside, and so may be more flexible on the question of retributivism than other critics who take targeting retributivism to be a primary objective of their projects.

¹⁹489 U.S. 189 (1989); 812 F.2d 298, 301 (1987).

²⁰CHIAO, *supra* note 1.

a boy who had been physically abused by his father over a period of time during which he was not removed from his home, at the end of which the boy was severely injured. The boy and his mother sued the state of Wisconsin for its failure to remove the boy from the father's home, despite prolonged evidence of abuse, and its failure to protect the boy from abuse. The Court, in both instances, found for the defendant, alleging that the state had no duty to protect the child, though it did have a duty to punish the perpetrator *ex post*. Chiao rightly criticizes the Court for its austere approach, defending only the state's duty to punish without recognizing its duty to protect or provide other *ex ante* services, which per the Court are provided by the state at its own discretion. Chiao, however, allies retributive views with the *DeShaney* Court's position, positing a conflict between retributivism and welfare provision.²¹ The analysis polarizes the views by positing that theoretically (not just budgetarily) one has to choose—either one thinks *ex post* punishment is what is needed, or one recognizes the state's *ex ante* welfare obligations of protection. On Chiao's analysis, one cannot defend both.²² A focus on *ex ante* schemes for the promotion of welfare is assumed to be inaccessible to a view that would offer a famed “backward-looking” perspective on the justification of punishment.

This mutual inconsistency of retribution and welfare-promotion is similar to that posited by Nussbaum, according to which either we care about the social determinants of crime and its alleviation in the form, *inter alia*, of early intervention in education, nutrition, welfare provision, etc. or we care about holding people retributively accountable for their wrongs. This, however, is to demonize retributivism in unwarranted ways that hark back to the critique of retributivism as committed to the inflation of wrongdoing: for, or so the argument goes, if the suffering of the deserving is intrinsically good, then the retributivist should be committed to hoping for more crime to allow for more retribution.²³ Clearly, however, the initial wrongdoing is as detested by a retributivist as it is by the welfarist or pluralist-consequentialist (and possibly more so, if we focus not merely on the objection to loss but to the wrong as well). To the extent that the wrongfulness and harmfulness of wrongdoing can be avoided by investing in welfare-enhancing projects that not only help persons facing difficulties, but have the further advantage of preventing future crime and alleviating social ills, this certainly can be desirable on a retributivist picture. A commitment to holding persons accountable *ex post* in no way undermines the grounds for being committed to addressing human needs and challenges *ex ante*. (In fact, the grounds of the former may dictate that we ought to promote the latter, as I shall suggest in [Section III.C](#) below.)

What is crucial to undermine this line of attack against retributivism is to see that retributive commitments do not interfere with the case for expanding public protection, education, health services, etc., or with a general pro-welfare-provision approach to state responsibilities. While the backward-looking approach to punishment is indispensable on any retributive view, the retributivist can also endorse

²¹*Id.* ch. 1, §4.

²²Arguments with a similar tenor are made throughout, for example, where Chiao argues that one can oppose mass incarceration or endorse retributivism, but, once again, one cannot do both. *Id.* ch. 4, §1.

²³*Cf.* C.W.K. Mundle, *Punishment and Desert*, 4 *PHIL. Q.* 216, 223 (1954).

Nussbaum's charge that "society should take an *ex ante* perspective, analyzing the whole problem of crime and searching for the best strategies to address it going forward." What will determine whether one endorses this view will depend, however, not on one's retributive commitments, but on one's conception of the proper scope of the state.

This line of objection thus seems to confuse the retributivist commitment to punishment with a political conception of the state as restricted to the function of punishing. That is, the view that the authors appear to be justified in opposing is one that restricts the legitimate scope of political intervention to punitive measures, whereas, on their conceptions of justice, the state should be engaged in a far broader set of welfare and justice-promoting projects beyond those that can be achieved by the punitive. In other words, they oppose a minimal, or night-watchman state. Importantly, however, there is no clear allegiance between a minimal conception of the state and a retributive conception of punishment. If anything, one would think that a minimal state would steer clear of morally charged interventions that depend on a moral rather than political conception of wrongdoing, or that aim at aligning the lives of wrongdoers with an ideal moral state (whereby wrongdoers "get what they deserve").²⁴ It would rather restrict itself, one supposes, to a consequentialist defense of social stability—of "keeping the peace"—a project that appears far more consistent with a political conception that endorses a minimal state (aimed merely at securing the basic conditions for civil coexistence).²⁵

More importantly, the commitment to retributive punishment does not ground a commitment to a minimal state. Rather, it leaves open the question of the scope and authorities of the state, while retributivism can be consistent with even a radically expansive conception of the state that sets its sights on bettering the opportunities and situations of all those in need, so long as the duty of retributive punishment is among the responsibilities that the state is understood to be required to bear and carry out.

To put the point in terms more particular to the allegations made, while certain punitive means might be ruled out by retributivism (e.g., incarcerating the innocent) to achieve benefits such as protection against victimization,²⁶ the retributive commitment to the punishment of the deserving *does not* rule out most of the penal interventions defended by Chiao, including public policing, investment in prevention, social services monitoring, and early intervention, all of which might be provided to protect against victimization. Nor does it rule out the broader welfare-oriented interventions supported by Nussbaum, including nutrition, education, health care, housing, and employment. In fact, the best retributive views can be understood to be grounded in the same principles that underlie these projects, in terms of the need to respect the equal worth of persons: the former via *ex ante* means of providing

²⁴While prominent libertarian Robert Nozick endorsed both a minimal state and retributive punishment, the joint adoption of these two views is to be puzzled at rather than assumed. ROBERT NOZICK, *PHILOSOPHICAL EXPLANATIONS* (1981).

²⁵Or at most one that aims at appearing to advance folk conceptions of justice (that satisfy the public).

²⁶And thus one might wonder whether it is the negative retributive thesis (that bars the punishment of the undeserving) that is more at odds with welfare-oriented approaches than the positive retributive thesis (in favor of punishing the deserving), which receives all the "bad press."

the opportunity to live in ways that are consistent with such worth, and the latter via ex post appropriate reactions to the actions of others that violate such worth. Thus the first argument against retributivism fails.

2. *Inconsistent Psychologically?*

While Nussbaum's argument is plausibly read as an argument about the substantive inconsistency of retribution and welfare orientation, there is an alternative version of the argument that may yet survive: one that focuses on human psychology rather than substantive inconsistency. That is, her argument can be read as proposing that as a matter of human psychology, one can either focus on the future, love, and welfare, or focus on the past, anger, and retribution, but one cannot contain both. On this understanding, even if there is no principled reason not to view welfare-promotion and retribution as complimentary projects in an ethical system or a system of justice, human psychology renders the commitment to these projects mutually exclusive, i.e., people are either angry or compassionate, interested in retribution or welfare. This worry is reinforced by a simple glance at the political arena, wherein "tough-on-crime" and social-welfare camps are generally politically opposed. Thus, given the choice between these two psychologically exclusive moral outlooks, we should endorse the latter and abandon the former (presumably as a necessary psychological step toward fully embracing the latter).

This psychologically grounded argument seems at least to have more plausibility than the substantive variant. Yet it relies on an empirical picture that Nussbaum does not exert effort to support, and that seems to be undermined by daily experience. That is, perhaps Nussbaum is right about her conjecture (and I expect that every reader will have had acquaintance with people who appear to polarize in this way). But perhaps she is not; perhaps these are caricatures that fail to appreciate the deep tensions that can be and are borne by the human condition and multi-textured human psychologies. Even in the political context, major political movements in fact endorse both modes of interaction, though often with respect to different target groups. Consider feminist concerns for care and compassion that are coupled with calls for increasing punitive accountability with respect to sexual and domestic crime. The insistence on promoting well-being and the insistence on accountability are not mutually exclusive psychological options, but perspectives appealed to where we perceive a lack of responsiveness to each. Both environment (or predicament) and agency are important factors that play a role in the way we respond to wrongdoing. Though we may instinctively focus more quickly on the "wrongdoer as victim" in some contexts and "wrongdoer as agent" in others, the capacity to be responsive to both these perspectives is apparent, while a prereflective attraction to one perspective rather than another provides no reason to abandon the project of aiming to address all morally relevant aspects of wrongdoing, requiring the appreciation of both. Each brings important moral factors to bear on the way in which we respond to others: those that call for holding offenders to account for their culpable behavior, and those that require the appreciation of the social factors and patterns associated with crime, and of the persons disadvantaged by these as victims, even if also perpetrators of wrongs against victims.

In fact, personal reflection of the kind invited by Nussbaum on the many relationships we have with others would seem to further demonstrate that we are far more nuanced and psychologically pluralistic than the picture she presents would allow. Many people are not “angry, period;” or “compassionate, period.” Susan Wolf’s compelling depiction of “blame, Italian-style” is an expression of the central place anger can have in relationships structured by love and care, which themselves can elicit anger when those whom we love act in ways we cannot accept.²⁷ This depiction demonstrates that Nussbaum’s suggestion that anger and retribution must be displaced to clear the way for love and compassion straightjackets the human emotional and moral life into a singular picture that is far from persuasive.²⁸

If then human psychology does not rule out pursuing a moral worldview that endorses retributivism and welfare orientation, a defense of the latter does not serve to displace the former.²⁹

²⁷Susan Wolf, *Blame, Italian Style*, in REASONS AND RECOGNITION: ESSAYS ON THE PHILOSOPHY OF T.M. SCANLON 332 (R. Jay Wallace, Rahul Kumar, and Samuel Freeman eds., 2011). See further Strawson’s suggestion that retributive sentiments are integral, necessary features of valuable relationships. P.F. Strawson, *Freedom and Resentment*, in 48 PROC. BRIT. ACAD. 1 (1962). But see Gary Watson, *Responsibility and the Limits of Evil: Variations on a Strawsonian Theme*, in AGENCY AND ANSWERABILITY: SELECTED ESSAYS 219 (2004).

²⁸It appears to reflect more of a Christian ethic than a descriptive account of human psychology. Further, even if persons are, individually, so limited, there are further reasons to doubt that a state, which is not in itself a “psychological entity” (though human psychology has a clear impact on state decision making), would be subject to the same psychological limitations as individuals.

²⁹This does not resolve the question of practical inconsistencies. Given limited resources, investment in one of these projects will be at the expense of the other. Thus, even if substantively and psychologically consistent, punishment and welfare-oriented projects will come into conflict. Consequently, one might suggest that given the choice, we should give up on retribution and invest in forward-looking projects. This is yet another way in which we may interpret Nussbaum’s contention that to move forward we must—as a matter of practical necessity (in addition to substantive or psychological inconsistency)—give up on retribution. Notice, however, that devoting resources to detecting, prosecuting, and punishing is an inevitable expenditure on any but an abolitionist view (while the critics under discussion do not take abolitionism to be the best path forward). Nonabolitionists all endorse expending substantial resources on punishment that could have been allocated to projects such as education, health care, housing, etc. (I thank an anonymous reviewer for highlighting this point.) Nonetheless, in the critic’s case, such expenditure is *in service of* welfare promotion, whereas on a retributive view, resources are diverted away from welfare-oriented goals, demanding that we trade off welfare promotion against alternative pursuits. I.e. though consistent, there will be genuine *de facto* conflicts to resolve. While the project of working out the balance between different state pursuits is beyond the scope of this paper, it is worth making a number of points in this regard. First, this is a challenge that arises with respect to any pluralistic endeavor, and is not a concern distinctive to the question of how to marry retribution and welfare orientation. (That is, unless one endorses the view that the retributive project trumps all others. This would generate reason to reject the view on welfare-oriented grounds. Yet I doubt any retributivist has ever held this view.) Further, on the question of the strength of the reasons and duties involved, a commitment to retributivism would generally commit one to the conclusion that at least in an important set of instances, the value (reasons in favor or duty) of retribution outweighs that of practically competing projects, else the moral force of desert would be rendered practically inert. (But see Husak, *supra* note 3.) Finally, as will be developed, on the view presented, this is less welfare-threatening than might appear at first glance, since retributivism is not the foundational norm, but emerges from more foundational normative principles that ground other goods and duties, many of a humanistic bent. Thus, *it would be inconsistent with the underlying tenets of retributive justification to pursue retributive goods or duties to the excessive detriment of all others.*

B. Retributivism and Mass Incarceration

1. Placing Blame

Another set of arguments marshalled against retributivism revolves around the alleged relation between retributivism and harsh and unjust punishment. Arguments in this regard have centered around the phenomenon of mass incarceration, which, per some critics, is a product of retributivism, and per others, retributivists cannot oppose, such that to move forward and overcome excessive harshness and mass incarceration, we must, once again, abandon retributivism.

Mass incarceration is a phenomenon that refers, *inter alia*, to the problem of the sheer number of persons incarcerated in the United States, which is higher in absolute and per capita numbers than in any other nation in the world (at a rate that is five to ten times higher than that of other liberal democracies, and five times higher than that in the United States a generation ago). But it only begins there. The racial core of mass incarceration and its profound impact on young, black men is well documented, as is the impact of mass incarceration on black communities, the social stigmatization of color, and its role in the relegation of entire swaths of the population to a social underclass, which predicts the repetition of current patterns of social disadvantage.

The first of this set of arguments alleges that retributivism is *responsible* for harsh, excessive, oppressive aspects of criminal justice systems, including mass incarceration. These and a host of other ills are attributed to retributivism's influence on the criminal "justice" system. Thus, it is argued that to move forward to a more humane, welfare-oriented, just system of punishment, we *must* eradicate the retributive idea and its influence on systems of criminal punishment.

Erin Kelly defends such a claim, contending, at times explicitly, at times implicitly, that retributivism is responsible for many of the ills she identifies in the criminal justice system, including "massive suffering," "excessive and degrading" punishment, and racial discrimination.³⁰ She blames retributive and blame-oriented theories as inadequate justifications of punishment and the language and public narrative of blame used to legitimize punitive practices. Her criticism of these is not unconnected: she thinks of the logic of blame and retribution as "stretched" to justify a massively overreaching punitive system (and as a natural inflation of these substantive ideas rather than a case of their illegitimate usurpation for independent political purposes). Kelly advances a concrete causal claim: that the language and theorizing of retribution and blame are responsible for many of the system's injustices. Kelly asserts that were we to eradicate these ideas, the system would be improved and become more humane and moderate.³¹ Thus a commitment to humaneness and moderation begs the rejection of retribution.

³⁰KELLY, *supra* note 1, at 183. Kelly, for her part, advances an alternative reason for opposing retributive theory: she argues that political institutions *cannot* track moral desert. In addition, Kelly trades on the assertion that retributivism and welfare are conflicting goals and that a blame or a moral desert-oriented approach to punishment should be ruled out, *inter alia*, given their contribution to punitive excess, the significant injustices they produce, and their being opposed to a welfare-oriented, caring, or just approach to wrongdoing and its "resolution."

³¹*Id.*

Kelly here relies on James Whitman's attribution of American harshness and punitiveness to retribution,³² and her work is later echoed, for example, by Ekow Yankah, who takes the charge a step further. Yankah takes retributivism to be the dominant justification of American criminal justice, a fact he takes to be revealed, *inter alia*, in none other than the phenomenon of mass incarceration itself.³³ "The American experience with mass incarceration," he writes, "vastly disproportionate to any sensibly consequential or rehabilitative value of prison, reveals the retributivism at the heart of our system," an allegation made in apparent obliviousness to its glaring disproportionality within a retributive theory system of thought.³⁴ He takes retributive thought, at the heart of the system, to be responsible for much of the havoc wreaked on individuals, communities, and minority groups by way of the criminal legal system, and unavoidably so: mass incarceration, he argues, is both an "inexorable" and "predictable" outcome of a commitment to a retributive system of thought.³⁵ Echoing sentiments present in the work of Nussbaum, Kelly, and Chiao, he writes: "The simple point is that without addressing our retributivist commitments, there is no way forward on mass incarceration."³⁶

These are tall accusations, and if successful, would provide grounds for rejecting retributivism.³⁷ But there are reasons to doubt the plausibility of such claims.³⁸ For one, empirical evidence is not offered to back the causal claims, beyond motioning

³²Central to Whitman's multifaceted argument is his claim about the fundamental harshness of retribution and retributive defenses of punishment, which must, on his view, be abandoned if we are to pursue a humane program of punishment. Whitman, *supra* note 11. It warrants mentioning that while Nussbaum's work has always taken an interest in the promotion of human welfare, understanding, and care, she has in the past simultaneously endorsed retribution as an apt response to wrongdoing. In her recent monograph, she credits Whitman's attack on retributivism as playing a role in leading her to adopt a position against retributivism and anger, which she now takes to be opposed to these goods.

³³Ekow Yankah, *Punishing Them All*, 97 RES PHILOSOPHICA 185, 194 (2020). Retributive thought is, per Yankah, also to be faulted with stunted US equal protection jurisprudence. Yankah uses the famed *McCleskey* case, wherein the Court denied the relevance of the Baldus study (which demonstrated that racial factors had a significant impact on death penalty imposition) and held that an equal protection claim requires evidence of intentional (rather than institutional) racism in the particular case, to demonstrate the havoc that retributivism has wreaked on American criminal justice.

³⁴*Id.* But see Larry Alexander, *Retributive Justice*, in OXFORD HANDBOOK OF DISTRIBUTIVE JUSTICE 177 (Serena Olsaretti ed., 2018).

³⁵Here, one might suggest that Yankah and others must concede the desert of mass incarcerated wrongdoers (on a retributive perspective) for their criticism to succeed, since if those incarcerated are not deserving, their punishment would be *unjustified* per retributivism: a failure of retributive justice rather than a consequence of it. I thank an anonymous reviewer for raising this point. As developed in the next section, there are ample reasons to worry about the desert of the incarcerated and to object to mass incarceration on retributive grounds. Often, however, critics allege that adopting a retributive view leads to harsh justice independent of whether retributivism itself would endorse the implications to which it leads (because, e.g., of an inherent tendency toward harshness that exceeds the defensible implications generated by the view itself). For nonretributive views to offer a viable corrective to this problem, however, it must be demonstrated not only that alternative views provide grounds for criticizing mass incarceration, but also that adopting an alternative will, independent of what substantively follows from the view, in fact moderate the harshness of criminal justice. This article has raised doubts about whether we should believe that modifications of the type envisioned could have any such impact.

³⁶Yankah *supra* note 33, at 196.

³⁷Or at least for rejecting its open and public endorsement.

³⁸See further Husak, *supra* note 12.

in the direction of correlations between “the rise of the retributive justice movement” and the exacerbation of such ills.³⁹ Empirical evidence in such contexts is of course notoriously difficult to establish, yet there are further reasons to seriously doubt that retributive thought and justification and the legitimating narrative of retribution are in fact to blame for the harshness of American criminal justice. First, despite the rise of retributive theory within contemporary Anglo-American *philosophical* circles, retributive thought has not had nearly as significant an impact on American law and criminal doctrine as such criticisms suppose.⁴⁰ To cite a number of examples, the influential Model Penal Code places greater emphasis on the standard of dangerousness than on an ethic of desert,⁴¹ while the Sentencing Guidelines, famed for their influence on excessive and lengthy criminal sentences, fail to take account of the many retributive factors that would affect (and mitigate) sentencing, while attributing decisive significance only to criminal record (primarily as an indicator of dangerousness) and severity (which is relevant on almost any view, retributive, deterrence-based, or otherwise).⁴² In other words, retributivism *is not* the dominant system of thought that underlies American criminal law, even if it is a salient rhetoric utilized in the political sphere, and thus would be hard pressed to be responsible for a system that has largely ignored its contentions.⁴³ By comparison, the German criminal legal system, which is doctrinally more committed to culpability-based principles of punishment, has emerged as a far more moderate system, further undermining the contention that retributive principles inevitably produce harsh justice.

Though a question of political discourse rather than punishment theory, there is also reason to doubt that retributive language and rhetoric are in any way responsible for the current state of affairs. Kelly is right that retributive rhetoric has been mobilized

³⁹KELLY, *supra* note 1.

⁴⁰Michael Tonry, *Can Twenty-First Century Punishment Policies Be Justified in Principle?*, in *RETRIBUTIVISM HAS A PAST: HAS IT A FUTURE?* 3 (Michael Tonry ed., 2011).

⁴¹MODEL PENAL CODE (AM. L. INST. 1962). “Allusions to retributive ideas appear only three times, and faintly. Retributive ideas were almost absent from the most influential American criminal law document of the twentieth century . . .” Tonry, *supra* note 40. The 2007 Model Penal Code revision gave pride of place to proportionality and blameworthiness of offenders, MODEL PENAL CODE §1.02(2) (AM. L. INST. 2007), echoing the rise of retributivism in penal theory; however, the state of American criminal justice has not similarly transitioned. See Matt Matravers, *De-Moralising Retributivism*, in *THE ROUTLEDGE HANDBOOK OF CRIMINAL JUSTICE ETHICS* 144, 145 (Jonathan Jacobs & Jonathan Jackson eds., 2016).

⁴²Paul H. Robinson, *Punishing Dangerousness*, 114 HARV. L. REV. 1429 (2001). See also U.S. SENT’G GUIDELINES MANUAL §289 (U.S. SENT’G COMM’N 1999). On the hostility of desert theories to the consideration of criminal record, see, e.g., George P. Fletcher, *The Recidivist Premium*, 1 CRIM. JUST. ETHICS 54 (1982); Alexis M. III Durham, *Justice in Sentencing*, 78 J. CRIM. L. & CRIMINOLOGY 614 (1987); Jesper Ryberg, *Recidivism, Multiple Offending and Legal Justice*, 36 DANISH Y.B. PHIL. 69 (2001); Mirko Bagaric, *Double Punishment and Punishing Character*, 19 CRIM. JUST. ETHICS 10 (2000); Richard L. Lippke, *The Ethics of Recidivist Premiums*, in *THE ROUTLEDGE HANDBOOK OF CRIMINAL JUSTICE ETHICS* (Jonathan Jacobs & Jonathan Jackson eds., 2016). But see U.S. SENT’G GUIDELINES MANUAL §4A, introductory commentary (U.S. SENT’G COMM’N 2008).

The Federal Guidelines were in full force from the second half of the 1980s until well into the 2000s; while the Guidelines have since ceased to be mandatory (see *United States v. Booker*, 543 U.S. 220 (2005)), they continue to exert substantial influence on contemporary sentencing, and are only one example of influential factors that shaped the face of American criminal justice, and which fail to reflect the retributive commitments taken up by philosophers of criminal law.

⁴³Kelly concedes that criminal law is often at odds with retributive contentions and does not reflect the tenets of retributivism.

in the political arena to advance the case for harsher punishments. Yet the trope of dangerousness has similarly played a salient role in mobilizing American punitiveness (a point that need not be established as against, e.g., Kelly herself, who refers repeatedly to the centrality of idioms of dangerousness and images such as that of the to-be-feared “Super-Predator” in the American legal context).⁴⁴ In fact, the “lock ‘em up” mentality Kelly hopes to destabilize as well as the policies she attacks, from three-strikes laws to incarceration as a standard mode of punishment, are even more consistent with a logic and language of dangerousness than that of blame and retribution. Why—one might ask—would we throw people in a cell and throw away the key? One answer that might present itself is: because “they” are not moral agents like the rest of “us” to be engaged with and held accountable as per retributive theory, but rather, radical “others,” too dangerous to be allowed to remain on “our” streets. Kelly rightly objects to categorizations such as “social psychopath” and “pedophile,” and one might add “criminal” and “felon” to the list of essentializing categorizations that enable excessive punitiveness. But these categories mark out wrongdoers as stigmatized others regardless of whether it is their being “to blame” (deserving) or “evil,” or alternatively, “dangerous” or “sick,” that demands their removal. One wonders therefore whether the current state of American criminal law tells more of the hydraulics of aggression than the malignance of retribution.

In fact, retributivism is opposed to many of the very features that Kelly opposes: excessive and degrading punishment, racial discrimination, and the relegation of wrongdoers to a social underclass are all inconsistent with a retributive ethic, which defends the imposition of proportionate punishment only on the deserving and only to the extent of their desert. This congruity seems to offer ripe grounds for theorists of all stripes, nonretributivists and retributivists alike, to view each other as allies in a joint enterprise of opposing the ills of the system, which they mutually regard as targets, rather than as foes engaged in mutually exclusive projects. This is yet another opportunity for complementarity where the resources of welfare-oriented and retributive approaches can both be drawn upon to limit excessive punitiveness and take punishment “forward.”⁴⁵ Ironically however, the choice has been to cry blame, pointing a finger at retributivism rather than joining forces to combat these ills.⁴⁶

2. Barring Objections

Another rendition of the connection between retributivism and harsh and oppressive punishment alleges that retributivism, even if not responsible for mass incarceration,

⁴⁴See also Matravers, *supra* note 41, at 149.

⁴⁵See, e.g., Brian Murray, *Are Collateral Consequences Deserved?*, 95 NOTRE DAME L. REV. 1031 (2020).

⁴⁶Illustrative is Nussbaum’s treatment of *lex talionis*, which she characterizes as retributive, rather than more accurately as a genealogical predecessor of retributivism and a possible supplementary principle of quantification. Even if her characterization were correct, she must be well aware that, barbaric though it may seem today, *lex talionis* was introduced as a *limiting* principle, to bar imposing damage exceeding that which had been caused—a principle that *moderated* social practice, rather than serving as a source of excessiveness and cruelty. See SAINT AUGUSTINE, ANSWER TO FAUSTUS, A MANICHEAN 254, 408–410 (Boniface Ramsey ed., Roland Teske trans., 2007). Similar misrepresentations are reproduced in other contemporary allegations against retributivism, which also miss its limiting potential and instead insist on its role as an escalator of violence.

cannot oppose it. That is, it is claimed that another good that is unavailable to the retributivist is the ability to criticize oppressive and unjust punishment (other than where the injustice inheres in a deviation from retributive principles).

Chiao, for example, argues that one can either affirm that mass incarceration is a problem, or endorse deontological theories such as (strict deontological) retributivism, but one cannot do both.⁴⁷ On Chiao's conception, this is because it is plausible that each person incarcerated in the United States deserves punishment in the retributive sense, while retributivism does not have the resources to recognize that punishment can be oppressive under conditions where, individually, each person is receiving their "just deserts."⁴⁸

If Chiao's claim, a claim echoed in other contemporary attacks on retributivism,⁴⁹ were true, and one could not endorse a critical view of mass incarceration, this would once again be a real blow to retributive theory.⁵⁰ However, first, Chiao underestimates the resources available to the retributivist to object to mass incarceration. Per the analysis offered, retributive theory cannot point to any injustice that might lie in the sheer quantity of persons incarcerated in the United States, a significant feature of the phenomenon of mass incarceration, though by no means its most objectionable. Yet while a retributive theory of the type envisioned by Chiao cannot capture all of the injustices encompassed by this complex phenomenon, it can at least explain more than Chiao is willing to concede. For example, noting the relativity of the case for injustice in the numbers, as well as the historical patterns of discrimination reproduced in mass incarceration, can easily shed light on how retributive theories alone can point to the injustices latent in the system.⁵¹

The very idea that there are too many people incarcerated in the United States is not an absolute claim, but a relative one. Nearly every empirical rendering of the ills of mass incarceration, including that presented by Chiao, highlights the quantity and ratio of incarcerated persons in the United States *as compared to* the numbers of incarcerated persons in other liberal and nonliberal states. What is horrifying, or at least one thing that is horrifying, is US exceptionalism in this regard. If one presumes that at least some other systems are broadly invested in the significance of desert, noting this difference is reason to suspect that the United States may be going very wrong even exclusively on a retributive account.⁵²

⁴⁷"[Y]ou can either affirm that the US currently incarcerates too many people, or you can affirm a strictly deontological theory of punishment. But it is doubtful you can affirm both." CHIAO, *supra* note 1, ch. 4, §1. Chiao addresses a set of theories he terms "strict deontological" views (which are not coextensive with "retributivism" as a category of views that justify punishment in terms of the desert of the agent). *See id.*

⁴⁸For a partial endorsement of this view *see* Alexander, *supra* note 34, at 179, who takes the view that a deserving punishee can have no valid complaint against their own punishment (though others may).

⁴⁹*See, e.g.*, Yankah *supra* note 33.

⁵⁰For an endorsement of this view (in the political sphere), addressed by Yankah, *see* JAMES COMEY, A HIGHER LOYALTY (2018).

⁵¹For an argument for why certain retributive views would object to the sheer numbers of incarcerated in a system of mass punishment, even under idealized nondiscriminatory conditions, *see* Hamish Stewart, *The Wrong of Mass Punishment*, 12 CRIM. L. & PHIL. 45 (2018).

⁵²Comparing the United States to other jurisdictions can of course point in other directions—perhaps the United States is off the mark, but perhaps everyone else is (or perhaps the United States is plagued by higher rates of criminality).

Moreover, as noted in Chiao's work, much evidence demonstrates that patterns of incarceration in the United States mimic patterns of plantation enslavement and historical patterns of discrimination against the black population.⁵³ The reproduction of these patterns is reason enough, even on a strict deontological view, to suspect that the US system is rife with injustice.⁵⁴

But suppose that Chiao is correct—that such theories cannot point to any injustice in the quantity of persons incarcerated in America. Even if we accept that all the incarcerated are also robustly guilty as proposed by Chiao, is this a problem for retributive theory or reason to reject it? Chiao insists that it is.⁵⁵ Yet retributivism does not offer the comprehensive moral account that the argument against retributivism presumes. Retributivism aims to offer a justification of punishment, that is, an account of the rightness or goodness of deliberately imposing harm on others in response to their wrongs that turns to the desert of the agent to explain how the *prima facie* objection to the deliberate imposition of harm in the context of punishment is overcome, while reasons in favor of imposing such punishment are generated. But it does not aim to offer a comprehensive account of all moral value or duties, and so needs not have the resources to address every possible moral calamity or offer grounds for criticism thereof, even in the context of punishment. Imposing undeserved punishment simply is not the only thing that can go wrong with punishment. Deserved punishment can be manipulatively imposed, exploitative, oppressive, discriminatory, etc.

The possibility of these forms of injustices is not ruled out by retributive theory, nor is an account of them provided by retributive theory. To obtain a comprehensive answer to the question of when and under what conditions punishment is justified one needs not only a theory of desert, which licenses punishment and generates reason in favor of punishing, as well as a *pro tanto* duty to punish, but also theories of exploitation, oppression, discrimination, and other wrongs, the violation of which might render even deserved punishment unjustifiably imposed, as well as an account of how different moral duties interact. Developing the contours of any of these may be seen as complementary to the project of developing the contours of retributivism, such that if retributive theory fails to capture what is wrong with some aspects of the phenomenon of mass incarceration, this need not be to its detriment. Other ills may occur.

C. Taking Stock and Moving Forward

What both of the above sets of failed arguments against retributivism appear to assume is ethical or political monism: that on a retributive view the *only* thing that can matter is retributive right or good, to the detriment of all else, possibly even where the pursuit of other aims does not conflict with pursuing retribution.⁵⁶

⁵³See, e.g., MICHELLE ALEXANDER, *THE NEW JIM CROW* (2012). See further JAMES FORMAN JR., *LOCKING UP OUR OWN* (2018).

⁵⁴Chiao's argument against strict deontological views is that they cannot explain the injustice in the numbers/rates of incarcerated. He does not allege that they cannot explain *any* injustice (e.g. those related to racial disparities).

⁵⁵See, similarly, Yankah *supra* note 33.

⁵⁶Alternatively, the view presupposed may posit that even though there may be other values or moral concerns, the retributive trumps all else.

Against this caricature of the retributivist position emerges the claim that if we care about welfare, being humane, oppression, or injustice, we must abandon retributivism. But retributivism is not committed to ignoring all other moral concerns.⁵⁷ (Nor are any retributivists committed to ignoring all other moral concerns.) In fact, its commitment to punishment may be understood to emerge from the very normative grounds that underlie welfare orientation and what Nussbaum refers to as “care.”

While the refrain has been that to care about persons we must eschew retributivism, it is in fact care about persons and an unwillingness to fall to indifference and immunity in the face of attacks against persons, their welfare, and other humanistic goods and values (e.g., autonomy, equality, rights, etc., all of which Nussbaum and others take to be part of an expansive conception of welfare) that provide the best grounds for retributive punishment.

The retributive charge to punish is grounded in recognition of the special moral concern with culpable wrongdoing, that is, primarily, with manifest disrespect for persons, their welfare, their interests, and their rights.⁵⁸ In this sense, it might be understood to be an *ex post* analog of the *ex ante* commitment to promoting such goods and values. To view each and every person as an individual bearer of irreducible, equal worth, is to take human well-being and its cognates to be ends worth protecting and enhancing.⁵⁹ In a parallel vein, to view each and every person as an individual bearer of irreducible, equal worth, would, one supposes, also lead to taking special interest in voluntary attacks against these goods and values and attempts to interfere with human flourishing.⁶⁰ This, one might suggest, is precisely what lies at the root of the call for retributive punishment, a response that rejects and responds negatively and burdensomely to actions that fly in the face of respect for persons, their well-being, or their interests. Blame and retributive punishment can thus be ways of valuing and caring about persons and humanistic goods and values,⁶¹ while retributivism and welfare orientation can be understood to be grounded in

⁵⁷While antiretributivists may insist that this is what the retributivist is committed to, often reverting to a single paragraph in Kant’s work (while it is likely distorting to view Kant as endorsing a simple retributivist account), it is significant that in rejecting monistic retributivism, the antiretributivist rejects no currently defended retributive account. Thus the argument is moot if intended to undermine retributivism in any of its contemporary forms.

⁵⁸I do not intend to exclude *mala prohibita* offenses, though I will not develop their relation to this formula here. On *mala prohibita* (and retributivism), see Douglas Husak, *Malum Prohibitum and Retributivism*, in *DEFINING CRIMES* 84 (R.A. Duff & Stuart Green eds., 2005). Further, one might be concerned that the arguments herein are unnecessarily anthropocentric. To the extent this is the case, the arguments developed can be easily extended to accommodate the welfare of nonhumans, with necessary modifications.

⁵⁹This is not a picture that would be endorsed by a classic welfarist-utilitarian, but our interlocuters are nonutilitarians, who rely on the equal worth of persons as central to their conceptions of welfare orientation and justice.

⁶⁰*Cf.* Morris, *supra* note 3; Hampton *supra* note 3.

⁶¹See Christopher Evan Franklin, *Valuing Blame*, in *BLAME: ITS NATURE AND NORMS* (Justin Coates & Neal Tognazzini eds., 2012) for a defense of the view that blame is a mode of valuing required by the standards of value. On the distinctiveness of care vs. respect, see Stephen Darwall, *THE SECOND-PERSON STANDPOINT* (2006); Stephen Darwall, *Sentiment, Care, and Respect*, 8 *THEORY & RSCH. EDUC.* 153 (2010).

the same normative source materials, such that if one allows a nonminimal state, taking a position in favor of one is reason to endorse the other.

The classic consequentialist position has of course been to insist that in the face of crime, prevention and addressing the future is sufficient, to the exclusion of a retributive response. Critics insist that we should address the problem of wrongdoing *ex ante* and invest in projects that would make everyone better off, rather than focusing on the past. And undoubtably, prevention of future losses and wrongs is of the utmost concern. Yet the retributivist is well positioned to adopt just this concern for the future and prevention as part of her broader picture, while arguably it is deep care for persons and their well-being that calls *not only* for looking forward and preventing future crime, but insisting on the irreducible significance of the attacks themselves. Notice that forward-looking responses only address attacks as *occasions* or opportunities to engage in projects that we already have reason to pursue (presumably, to take Nussbaum's example, we should be raising campus awareness about sexual assault, but this is true independent of whether a sexual assault has just occurred). *Exclusively* addressing the future, on the other hand, means that we do not take the attacks against human goods and values—that is, *this* attack, *this* wrong, *this* failure of regard—to be independently worthy of our attention and care.

Yet it is unclear why the time barrier should afford such immunity. True enough, we cannot change the past. But it is misleading to suggest that the only thing relevant to the future is the future. Our present and future are shaped by the way we understand and relate to the past, where we can change the meaning of a past event for the future by addressing it as a collectively recognized culpable wrong perpetrated by a responsible agent. The animating idea behind retributivism is that if we truly care or attribute significance to what *has happened*, we must respond negatively to the culpable wrong, holding the culpable wrongdoer responsible for her disrespectful action, and not merely look to dealing with the future.

This is not an argument for any necessary connection between welfare orientation and retributivism. The question of what it takes to address past wrongs adequately is a fraught one, and Nussbaum, for example, asserts that while accountability for the past is necessary for moving forward, retribution is not.⁶² The nonretributivist may thus reject the further commitments that may be necessary to arrive at a defense of retributive punishment.⁶³ The point is rather that that the two might be understood to be derivable from common normative precepts that allow positing a deep affinity between the two rather than the necessary opposition posited by the discussed line of objection.

⁶²Nussbaum means by this a response that is forward looking and deterrence-oriented, but that acknowledges that a wrong has occurred. She offers that this includes putting into place *ex ante* mechanisms intended to prevent wrongdoing, and announcing that wrongdoing has occurred (coupled with deterrent punishment). Yet how this achieves accountability remains unclear, as the relation of the *ex ante* mechanisms to accountability is dubious, while “telling the truth” of wrongdoing fails to engage the agent (while further ignoring culpability). This leaves only deterrent punishment to bear the burden of accountability; yet it has long since been taken to be a competitor rather than instantiation of accountability (and certainly fails at that task as it has been developed here, which precisely seeks something other than exclusively forward-looking preventive punishment). See NUSSBAUM, *supra* note 1, ch. 6.

⁶³Which may relate, for example, to the significance of human agency.

(One suspects that what often lies behind the rejection of retributivism is not, then, the contrived conflict between welfare orientation and retribution, but rather a rejection of the premises of retributive punishment, namely moral responsibility and the culpability of the agent being blamed and punished. While the free will debate is beyond the scope of this paper, suffice it to say that while absence of responsibility is reason to reject retributive punishment, welfare orientation is not.)

D. Summation

Whether or not one accepts the argument for the normative connectedness of welfare orientation and retribution, we have thus far demonstrated that the welfare-oriented arguments raised by critics against retribution fail: the retributivist is able both to support welfare-promoting projects and to criticize harsh, oppressive practices.

IV. Retribution and Suffering

A. Retributivism and the Intrinsic Value of Suffering

While the above accusations against retributivism fail to provide grounds for its rejection, there nonetheless seems to be a deeper tension between retribution and welfare that underlies the line of objection that posits their incompatibility. This tension might be captured by the less patently false accusation that it is the commitment to human welfare that excludes the possibility of endorsing retributivism. Here the problem lies in the prosuffering attitude attributed to retributive theory. As addressed earlier, a welfare-oriented approach is committed to the value and promotion of human welfare, whereas retributive theory is understood to attribute intrinsic value to human suffering. If this is the case, then welfare orientation, even if it does not monistically attribute value *only* to welfare, may exclude the possibility of endorsing retributivism given the direct conflict between retribution and the welfare of wrongdoers. It is worth noting that it is not the very justification of punishment that raises these worries, as all nonabolitionist, welfare-oriented approaches, including the approaches endorsed by all the critics discussed here, endorse the institution of punishment. That is, they too offer accounts that take the deliberate imposition of burdens, deprivations, or suffering to be a justifiable enterprise. Yet retributivism alone is taken to offer a justification of punishment that is founded on the intrinsic rather than instrumental value of human suffering.

We thus arrive at the concern that is at the heart of many classic and contemporary objections to retributivism and may supply the greatest motivation to reject retributivism as a theory of punishment. It is what underlies Hart's concern with retribution as "a mysterious piece of moral alchemy, in which the two evils of moral wickedness and suffering are transmuted into good."⁶⁴ It is the concern that underlies John Gardner's skepticism with respect to retributive justification, Thomas Scanlon's relation to retributivism as "morally repugnant[t],"⁶⁵ Victor Tadros's attack against

⁶⁴H.L.A. Hart, *Prolegomenon to the Principles of Punishment*, 60 PROC. ARISTOTELIAN SOC. 1 (1960).

⁶⁵T.M. SCANLON, *MORAL DIMENSIONS* (2009).

retributivism as “barbaric,”⁶⁶ as well as many similar charges.⁶⁷ As straightforwardly articulated by Tadros, retributivism is necessarily in conflict with “a basic concern that I hold dear—that the lives of all people . . . go well.” Articulating a deep skepticism about the moral credentials of retributive theory, he continues: “Few claims need more defence than the [retributive] claim that the suffering of other human beings is impersonally valuable.”⁶⁸

Suffering, one might say, is tolerated by other views, whereas it is *valued* by retributivism, at least in the context of punishment. Herein lies the heart of the problem: that to endorse retributivism requires accepting a view that takes human suffering to be “good” or of value, which is precisely what those committed to the value of human welfare would deny.

Or so the oversimplified rendition of retributivism alleges. Careful attention to contemporary retributive theory demonstrates that this simple articulation of the conflict between retributivism and welfare fails to capture the richness of retributive theory and may misunderstand it altogether. In either case, the ousting of retributivism given its conflict with welfare would turn out to be unnecessary or unmotivated, whether because the relevant objection applies only to a subset of retributive justifications of punishment, leaving viable options on the justificatory table, or because contemporary retributivism turns out not to be vulnerable to the relevant charge at all.

To anticipate, in the following I offer three related clarifications intended to unravel this direct conflict between retribution and welfare orientation. First, I clarify that retributive justification can rely on deontic rather than axiological claims, that is, the duty to punish or reasons to punish rather than on any value attributed to suffering. Second, I clarify that axiological retributivism often attributes intrinsic value to an object other than suffering. Third, I clarify that even those views that attribute intrinsic value to “deserved suffering” ought not to be understood as attributing value to human suffering simpliciter.

I anticipate that some will take these clarifications to be “acrobatics” intended to formally escape a conclusion that remains substantively true: that however the retributivist may sugarcoat it, she is ultimately committed to the goodness of human suffering. The problem with this response, however, is that it fails to appreciate that few contemporary retributivists endorse the view that is attacked by this line of objection. If that is right, the objection strikes down a caricature of retributive thought (though not necessarily a strawman, if one takes account of all retributive theories) *and makes no progress with respect to the question of whether we should endorse or reject retributivism and retributive punishment*. This I take to be sufficient to undermine this relevant line of objection.

One point worth highlighting before we proceed: antiretributivists tend to be highly concerned with the value and glorification of *suffering* attributed to retributivism, where suffering presumably refers to a negative, subjective experience associated with a diminution in well-being. Yet many contemporary retributive accounts directly

⁶⁶TADROS, *supra* note 6, ch. 4.

⁶⁷See, e.g., HONDERICH, *supra* note 7; GOLASH, *supra* note 7.

⁶⁸TADROS, *supra* note 6.

justify the imposition of *burdens and deprivations* rather than suffering, where these are not conceived as instruments for the imposition of suffering, but rather as that which is to be imposed. For example, on some accounts, criminal punishment is after deprivations of liberty and other central rights, rather than the suffering that might accompany such deprivations, while other accounts take censure and condemnation to be that which is deserved.⁶⁹ Thus, the charges made against at least some retributive theories are uncharitable. While important retributive accounts do endorse a suffering-oriented articulation, globally rejecting retributivism on grounds of its prosuffering attitude is often inaccurate if not altogether invalid. (Criticisms may of course be rearticulated in terms of objecting to burdens and deprivations, yet at least some of the visceral purchase of objecting to “suffering” would presumably be lost.) Nonetheless, I proceed in terms most charitable to the critic to see if the arguments against retributivism survive even if it is articulated in terms of suffering.

1. Deontic Retributivism

Many retributive views rely on an intrinsic goodness claim to justify punishment, whether about the intrinsic goodness of deserved suffering or punishment. On such views, whereas we have *prima facie* reason to believe that a practice that involves the deliberate imposition of suffering is morally bad, in fact, it is morally good. This is an axiological claim, i.e., a claim about value, which is generally taken to have deontic implications, that is, implications for the question of how we ought to act (e.g., for the permissibility of punishment, for reasons in favor of punishment, or for a duty to punish). For example, consider the retributive theory defended by Michael Moore.⁷⁰ Per Moore’s account, the punishment of the deserving is intrinsically good. (Moore is perhaps the foremost contemporary defender of this claim and its centrality within retributive thought.)⁷¹ Further, one ought to promote this retributive good and impose deserved punishment where there are no constraints against its imposition (e.g., liberty constraints, which limit the permissibility of state promotion of this good). In the absence of such constraints, Moore takes the failure to impose deserved punishment to be not only bad, but derivatively unjust. In other words, Moore offers an instrumentalist rendition of retributivist justification, whereby the retributive duty to punish is grounded on an intrinsic goodness claim.

⁶⁹On central accounts, what one deserves and that which justifies punishment is not suffering but rather, e.g., the humbling of one’s will (Herbert Fingarette, *Punishment and Suffering*, 50 *PROC. & ADDRESSES AM. PHIL. ASS’N* 499 (1977)), that one’s life go less well (Berman, *supra* note 3), or one’s loss of liberty (see Dan Markel & Chad Flanders, *Bentham on Stilts: The Bare Relevance of Subjectivity to Retributive Justice*, 98 *CA. L. REV.* 907 (2010)), while other accounts focus on censure and condemnation (Markel & Flanders, *supra*; ANDREW VON HIRSCH, *CENSURE AND SANCTIONS* (1993); DUFF, *supra* note 4). Jeffrie Murphy has proposed that though he takes the retributive desert object to in fact be suffering, suffering in this context is wrongly understood as intrinsically related to pain, and should rather be identified with the original meaning of the word: “to *endure* something that is not within the control of one’s own will.” Jeffrie G. Murphy, *Last Words on Retribution*, in *ROUTLEDGE HANDBOOK OF CRIMINAL JUSTICE ETHICS* 28, 33 (Jonathan Jacobs & Jonathan Jackson eds., 2016).

⁷⁰MOORE, *supra* note 3.

⁷¹See also Alexander, *supra* note 34; LEO ZAIBERT, *PUNISHMENT AND RETRIBUTION* (2006); Lawrence H. Davis, *They Deserve to Suffer*, 32 *ANALYSIS* 136 (1972).

On some retributive views, however, the justification of punishment does not rely on an axiological claim about the goodness of punishment or suffering. Rather, punishment is taken to be right or just, while the burdensomeness or suffering involved in punishment is taken to be constitutive of the right, and thus rightly imposed.⁷² In other words, on these thoroughly deontic versions of retributivism, a claim about the rightness, justice, reason in favor of or duty to punish is endorsed, yet it is not derived from an axiological claim about the intrinsic goodness of suffering or punishment.⁷³ Thus, one can endorse a retributive account whereby there is reason or a duty to impose punishment (1) *even though* suffering is bad, (2) *even if* suffering is bad, and (3) altogether *independent* of the question of the badness of suffering, since such reason or duty is grounded in further deontic claims (or in axiological claims that are not about the value of suffering). In any case, on these accounts, retributive punishment is justified without being based on a commitment to the relevant “prosuffering” attitude alleged to be at the foundation of any retributivist account. They of course endorse punishment; but recall: so does our critic. Thus the rejection of retributivism tout court on grounds of the conflict between welfare orientation and retribution fails. Retributivism does not necessarily rely on a conception of human suffering as intrinsically valuable.

2. Intrinsic Goodness of What

Taking this last point seriously should neutralize the relevant worry for some, who are attracted to the force of desert, yet hesitate to adopt a retributive perspective on punishment owing to worries about entertaining a view that sees suffering as “good.” It warrants mentioning, however, that this option will not satisfy (or be available to) consequentialists insofar as they are committed to the priority of the good over the right (though the objection to retribution would no longer inhere in the so-called perverse presumption that human suffering is good). Further, the question arises as to whether all axiologically grounded versions of retributivism, which seem the more prevalent, are vulnerable to the relevant charge.

Here, a second clarification is in order: retributive theories may attribute value to an object other than suffering, and this has implications for our evaluation of the “moral credentials” of retributivism.⁷⁴ As Berman following Feinberg has pointed out, desert is a three-part relation whereby (1) the desert subject is said to deserve (2) the desert object, in virtue of (3) the desert basis.⁷⁵ In the context of punishment,

⁷²See, e.g., Morris, *supra* note 3; Dan Markel, *Are Shaming Punishments Beautifully Retributive? Retributivism and the Implications for the Alternative Sanctions Debate*, 54 VAND. L. REV. 2157 (2001); Dahan Katz, *supra* note 3. One might alternatively take the more modest position that there are reasons to impose retributive punishment (to treat wrongdoers as they deserve). See Husak, *supra* note 3, at 11–12. (It is not clear, however, that Husak means to endorse such a thoroughly deontic defense of retributivist reasons to punish, as he also entertains the possibility that such reasons are grounded in axiological claims.)

⁷³See further Berman, *supra* note 6, at 452, rejecting Zaibert’s claim that deontic retributivists must be committed to an axiological claim about the value of deserved suffering; ZAIBERT, *supra* note 71, at 214.

⁷⁴See, e.g., Berman *supra* note 6, at 452; Husak, *supra* note 3, at 7.

⁷⁵Berman, *supra* note 6; JOEL FEINBERG, *DOING & DESERVING; ESSAYS IN THE THEORY OF RESPONSIBILITY* (1970). A fourth dimension of the relation, not highlighted by earlier work, is the desert agent—she who gives the desert subject what she deserves. See Alec Walen, *Retributive Justice*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Summer 2021), <https://plato.stanford.edu/archives/sum2021/entries/justice-retributive/>; Dahan Katz, *supra* note 17.

the desert subject is uncontroversially understood to be the wrongdoer. The desert basis is generally understood to be culpable wrongdoing. Yet what wrongdoers deserve might or might not be suffering.⁷⁶ Other alternatives include that what wrongdoers deserve is to be treated or responded to in some particular way, whereby, for example, it is the responsive act of punishing rather than the experience of suffering that is deserved,⁷⁷ and on axiological versions of the relevant claim, valuable.⁷⁸

Thus, to the extent that the objection to retributivism is grounded in presuming suffering to be the desert object, and suffering to be that which is good, it is important to see once again that the discussed antiretributive argument does not succeed in rejecting a potentially substantial subset of retributive theories, even among those grounded on axiological claims. It thus does not constitute an argument against retributivism tout court.

3. “Deserved Suffering” and Its Value

Having demonstrated that a broad set of retributive views are immune to the proposed line of objection, we could stop here, inviting critics to take proponents of these views as fellow travelers in the pursuit of a just criminal law. More, however, can be said with respect to those retributive views that attribute intrinsic value to the suffering of the deserving. I suspect that many take such “suffering-oriented” retributive views to be the only views on offer, equating this option with retributivism itself.⁷⁹ The above should be sufficient to demonstrate that this is false. And that even a successful attack on suffering-oriented retributivism is not a successful attack against retributivism tout court.

Restricting our view to “suffering-oriented” retributivism, does this finally bring us to a retributive outlook that is at odds with welfare orientation? That is committed to the intrinsic goodness of human suffering? And that must be rejected if we are committed to the value of human welfare? I suggest not. Or at least, not yet. There may very well be retributive accounts that are to be rejected on grounds of welfare

⁷⁶For a view that takes suffering to be the desert object, see, e.g., Alexander, *supra* note 34; LEO ZAIBERT, *RETHINKING PUNISHMENT* (2018).

⁷⁷For alternative desert objects see note 69. It is a matter of debate whether punishment has been imposed in cases in which the recipient does not find the infliction unpleasant or suffer. While many take the view that treatment must be *subjectively* unwelcome, it is possible that treatment can be relevantly burdensome irrespective of the particular punishee’s subjective experience: a homeless offender may welcome the bed and meal offered when imprisoned and fail to value the liberty she has been deprived of; still, the burdensomeness of the treatment need not be undermined by the idiosyncratic reception of such response on the part of the punishee. Rather, insofar as the response is intelligible to the parties as intentionally burdensome and unwelcome, and partly appropriate in virtue of such burdensomeness, this can be sufficient for the response to count as burdensome in the relevant sense. See Leora Dahan Katz, *Private, Public and Punitive Blame*, in *FROM MORALITY TO LAW AND BACK AGAIN: LIBER AMICORUM FOR JOHN GARDNER*, (Michelle Madden Dempsey & François Tanguay-Renaud eds., 2023).

⁷⁸Berman, *supra* note 6. See Leora Dahan Katz, *Relational Conceptions of Retribution*, in *PALGRAVE HANDBOOK ON THE PHILOSOPHY OF PUNISHMENT* 101 (Matthew C. Altman ed., 2023). As Berman notes, it is not clear how many retributive accounts can be attributed to each alternative since it is not clear that retributivists have paid much attention to specifying the desert object in their accounts of punishment.

⁷⁹This is perhaps owing to the centrality of Michael Moore’s influential account, which relies on the intrinsic goodness of the suffering of the deserving as the central justificatory retributive claim. See MOORE, *supra* note 3.

orientation, but not, I contend, merely on grounds of their attribution of value to “deserved suffering.” Let me elaborate, at risk of belaboring a central retributive point.

While retributive theories of punishment are repeatedly framed as “backward-looking” theories of punishment, many fail to fully appreciate the way in which the past is integral to the retributive idea and the very grounds of the justification of punishment it offers. On both retributive and nonretributive accounts, the world is a better place in the absence of wrongdoing. On both retributive and nonretributive accounts, in the event of wrongdoing, the world can be made better if punishment is introduced into the mix. The accounts diverge on how it is that the world is made better and how it is that value is added through punishment. On a critical rendition of retributivism, the retributivist takes suffering per se to be valuable, such that adding suffering (of the deserving) is what makes the world better. Hence the resistance to retributivism as attributing goodness to human suffering. Yet this is false. In fact, on the retributive picture, we must begin by looking at the past, since insofar as retributive accounts rely on an intrinsic goodness claim, they import the past into *the very description of the object of evaluation*.

On a nonantagonistic rendition of retribution, a world in which there is wrongdoing left unaddressed is worse than a world in which there is wrongdoing, properly addressed. That is, while we are better off in the absence of wrongdoing, if an agent *has* culpably wronged, a world in which the wrongdoer suffers their “just deserts” will be better than one in which wrongdoing is left “unmet,” that is, in a state of injustice.⁸⁰ But notice that per this view, suffering is not (or need not be) attributed independent value. Rather the desert relation between the wrongdoing and suffering is the focal point for explaining how the world is made better (or less bad): value is attributed to a complex *whole* that includes both wrongdoing and suffering, rightly related,⁸¹ rather than independently to suffering, “when deserved.”⁸² To miss this is to enable the error of thinking that retributivism endorses suffering per se, whereas in fact, the shorthand “deserved suffering” refers to a complex phenomenon that imports into the value-making features of punishment a relation and reference to the past that attributes positive moral valence to the complex whole that correctly aligns between the two. To understand retributivism as a theory that attributes positive moral valence to suffering simpliciter, and then to recoil from retributivism on this ground, is thus to refuse to acknowledge the content of the view.⁸³

The critic essentially confuses the content of punishment with the source of its value. Punishment consists in the deliberate imposition of something negative and undesirable on the punishee: burdens, deprivations, or suffering. It cannot generate the relevant value without such imposition. Yet punishment derives its value not from its content, as assumed by views that understand retributivism to endorse “suffering for suffering’s sake,” but from its *meaning* and from the right *relations* it

⁸⁰ZAIKERT, *supra* note 71. Markel, *supra* note 72, at 2196. See further Walen, *supra* note 75, at §4.3.1.

⁸¹This is the very mistake that enables some to think that retributivism would endorse more wrongdoing to allow the “value gains” achieved by more just retribution.

⁸²See ZAIKERT, *supra* note 71.

⁸³That is, retribution should not, or at any rate, need not be understood as a view that attributes intrinsic value to a particular subset of “sufferings” (deserved sufferings).

generates between wrongdoing and deprivations or suffering.⁸⁴ To appreciate how value can inhere in the relations between parts rather than exclusively in the constituent parts of a whole, Leo Zaibert offers a Husserlian example from the field of aesthetics to move the debate about retributive punishment forward: that of a musical masterpiece, in the first instance, in its original form, and in the second, freely rearranged. As Zaibert articulates, in rearranging the parts of the masterpiece “what I will thereby produce will be unfathomably less valuable . . . Independently of the parts of any whole themselves, the specific *ordering* of these parts generates – or fails to generate – value in different degrees.”⁸⁵ Returning to the context of punishment, the responsiveness involved in wrongdoing being met by punishment is a relational feature that can bear value, without the necessary ascription of value to the burdens and suffering there involved. Thus the intrinsic value (or rightness) associated with retributive punishment does not inhere in the *intrinsic features* of suffering, but in its *extrinsic features*—the relations generated between wrongdoing and such burdens, the meaning and value of which can be cashed out in different ways (in terms of fairness, justice, etc.).⁸⁶ As for the intrinsic features of suffering, these, it emerges, are only valuable in service of a further object of value, whether the erasure of immunity, fairness, moral alignment, dissociation, or otherwise, depending on one’s particular retributive view.

One might put the point as follows: if we isolate the question of the evaluation of human suffering, even within the retributive outlook, suffering alone is only *instrumentally or quasi-instrumentally* valuable. While on noninstrumentalist axiological retributivism, punishment does not derive its value from its downstream consequences, but from that of which it is a constitutive part, it is still the case that *intrinsic value is not independently attributed to suffering*. Rather it has value *only in service of some other object*. What is distinctive about retributivism, then, is not that it takes suffering to be intrinsically good, but rather that it takes suffering, burdens, or deprivations to be a *constitutive* element of the good.⁸⁷

The retributivist thus *is* committed to the thought that a world that includes the suffering of punishment is better than one without. But like on competing consequentialist accounts, adding *mere* suffering makes things morally worse. As above, more must be added to the mix: in the retributivist case, it is the constitution of the desert relation, while in the consequentialist case, it is the addition of welfare or other goods that makes the moral difference. And while a welfarist may want to insist that the only kind of addition that can make this moral difference is a contribution to human welfare, that is, of course, to fall back on a different debate—about monism vs. pluralism with respect to value, and whether anything other than welfare can be of value (which, if one takes rights, equality, etc. to be independently valuable,

⁸⁴See Hart, *supra* note 64.

⁸⁵ZAIBERT, *supra* note 76, at 41, relying on Edmund Husserl, *Vorlesungen über Ethik und Wertlehre 1908–1914*, Husserliana XXVIII (Ullrich Melle ed., 1988), at 96, to demonstrate Brentano’s view that “the value of a whole is not a function merely of the parts of the whole. Goodness also lies in the relations which are exhibited within the whole.” RODERICK M. CHISHOLM, BRENTANO AND INTRINSIC VALUE (1986).

⁸⁶For the distinction between intrinsic and extrinsic value, see Christine M. Korsgaard, *Two Distinctions in Goodness*, 92 PHIL. REV. 169 (1983).

⁸⁷Though there may be thoroughly instrumentalist versions of retributivism.

already puts one on the antimonistic side of the debate). This, however, is not the disagreement at hand, and in any event not the disagreement that leaves retributivism on the “repugnant” side of the conflict.

B. Remaining Tensions

One might yet worry that the above means that retributivists lack a commitment to the minimization of suffering, and thus there remain welfare-oriented grounds to reject retributivism. Yet this is the case with respect to the alternative justifications of punishment countenanced as well (i.e., the promotion of welfare and the minimization of suffering come apart). There are many circumstances within which suffering can be avoided, though retributive theory would endorse the deliberate imposition of (burdens, deprivations, or) punitive suffering. But this is the case with respect to any view that does not take the minimization of suffering to be the trumping moral criterion. It is not unique to retributive theory but common, for example, to any rich consequentialist vision that takes suffering minimization to have competitors (equality, rights, justice, etc.). In fact, Benthamite utilitarianism can be similarly criticized as it seeks not suffering minimization but welfare maximization. If there are two courses of action, one that imposes less punishment and suffering and leads to a net increase in aggregate utility, and one that imposes a far greater amount of punishment, yet leads to a greater net increase in aggregate utility, the utilitarian (and some welfarists) are committed to the moral superiority and duty to engage in the second, suffering-enhancing course of action. Of course, one could adopt a more sophisticated approach that avoids such outcomes, but the point is that the allegation that retribution cannot be countenanced by a welfare-oriented approach because of its failure to be committed to the course of action that minimizes human suffering cannot be a basis for arguing for the incompatibility of welfare and retribution, if strict welfarist approaches themselves lead down the same path.

The conclusion that may be drawn from the above is that there are tensions between welfare and retribution, and sometimes retribution recommends moving to a world that includes more suffering than otherwise exists (for reasons that are not simply translatable into the intrinsic goodness of suffering). Yet this is not because of some value inherent in promoting suffering “for suffering’s sake,” or because of a failure to appreciate the value of human welfare. Rather (in a Strawsonian vein), it is because in a nonideal world, in which culpable wrongdoing is a given, responsive actions and practices that are themselves valuable can involve negative reactions and suffering.⁸⁸ These have an indispensable role to play in just, moral relations.⁸⁹ Further, the aspects of welfare orientation that generate these

⁸⁸This defense aligns with a defense of the moral significance of guilt, a negative emotion that has an important moral role to play in first-person responses to culpable wrongdoing.

⁸⁹One might further argue that insofar as this is the case, the practice of retributive punishment *contributes to human welfare* as it contributes to a “life worth living.” Thus, the retributivist can also make a case for the contribution of retributive punishment to human welfare, e.g., in terms of this contribution to (and necessity for) a “life worth living.” This, however, is not the strategy that this article has been engaged in defending in seeking the compatibility of retribution and welfare orientation.

tensions, if taken to their end, recommend abolitionism rather than the displacement of retributivism in favor of a consequentialist alternative. If one's welfare-oriented approach countenances punishment, it provides no solid grounds to reject retributivism.

V. Conclusion

This paper has aimed to demonstrate that prominent reasons to reject retributivism, voiced regularly in contemporary arguments about punishment, fail, and that retributive punishment and welfare orientation are broadly consistent projects that can and should be pursued concomitantly. Neither generates substantive reason to reject the other, unless one is tempted toward an abolitionist or pacifist ethic, the unattractiveness of which lies beyond the scope of this paper. While keeping both welfare-oriented goals and accountability in our sights may be demanding, it seems not exaggerated to say that justice demands that we abandon neither of these projects, the merging of which will prove the best way forward toward a humane, defensible criminal law.

Acknowledgements. My sincerest thanks to Fernando Braccini, Miri Gur-Arye, Ryan Liss, and Alec Walen for their excellent comments on previous drafts. Thanks also to Joanna Langille and the participants of the University of Western Ontario Legal Philosophy Research Group Workshop, as well as participants of the Hebrew University Criminal Theory Workshop for helpful discussions. Research for this paper was supported by the Israel Science Foundation (Grant No. 2003/21).

Competing interests. The author(s) declare none.