Legal Theory, **25** (2019), 26–52.
© Cambridge University Press 2019 0361-6843/19
doi:10.1017/S1352325219000016

VAGUE COMPARISONS AND PROPORTIONAL SENTENCING

Jacob Bronsther*
Harvard Law School

ABSTRACT
The "small improvement problem" ("the Problem") applies when no option in a comparison is best nor, it seems, are the options equal, because a small improvement to one would fail to make it the better choice. I argue that vagueness causes the Problem, such that the options are vaguely equal or vaguely "related." I then unpack an important instance of the Problem, the comparison between a crime and a punishment upon which the ideal of a retributively deserved sentence is based. I argue that this comparison is not only vague, but remarkably vague, leading to an expansive array of "not undeserved" sentences. I conclude, however, that retributivism can only justify the *least harmful* "not undeserved" sentence.

INTRODUCTION

"The small improvement problem" ("the Problem") is ubiquitous in the realm of practical decision-making.[1] Should I get lunch with a friend or go for a run? Should we increase funding for the arts or education? Should we protect a small forest or exploit it to produce $300 million of tax revenue to be used for, say, health care? In all such instances, if we think that neither option is best, we know they are not equal, goes the argument, because a small improvement to one (making the lunch a little more delicious, making the artists a little more creative, making the forest a little bigger) would fail to resolve the comparison. If they were equal, then any improvement, no matter how small, would make it the better choice. The retributive ideal of the proportional criminal sentence, meanwhile, is based upon a comparison between an offense and a punishment. It searches for punishments that stand in some proportion to offenses in

* For incisive written comments, I am indebted to Nicola Lacey, Daniel McDermott, the late Derek Parfit, Peter Ramsay, Alex Voorhoeve, and two anonymous reviewers. I am also grateful to Professors Lacey and Ramsay for extensive discussion of the argument.

1. Ruth Chang refers to it as the "small improvements argument." Ruth Chang, *The Possibility of Parity*, 112 Ethics 659, 667 (2002). For earlier explications of the phenomenon, *see* Ronald de Sousa, *The Good and the True*, 83 Mind 534, 544–545 (1974); L.J. Savage, The Foundations of Statistics (1954), at 17.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

terms of their "moral gravity" or, as I will explain below, their "injuriousness." Classically, retributivists understand this proportion to be 1:1, such that the punishment ought to be equally as "injurious" as the offense. The resulting comparison is often a Problem case. Which is more injurious, the offense of embezzling $250,000 from a pension fund or the punishment of 25 months' imprisonment? Or perhaps they are equally injurious? If we imagine that the offender embezzled only $240,000, it would not help us to resolve the comparison. We need to unravel the Problem, then, to assess the promise of retributively proportional sentencing. Only by understanding the nature and depth of our uncertainty in Problem cases can we assess the degree to which we could employ reason—rather than some lesser guide like whim, unthinking precedent, or fad—to select a "proportional" sentence.

There are three sections to this paper. Section I argues, contra Ruth Chang, that the Problem is caused by vagueness, such that Problem options are vaguely equal or vaguely "related." Section II applies this analysis to the comparison between an offense and a punishment. The comparison is not only vague, I argue, but remarkably vague, leading to an expansive array of "not undeserved" sentences, in the language of Norval Morris.[2] What are the implications of this vagueness for retributivism as a foundation of state punishment? "Limiting" retributivists, like Morris, Michael Tonry, and Richard Frase, would claim to be unbothered.[3] They maintain that nonretributivist considerations, like the pursuit of crime deterrence, should determine our choice of punishment within the range of "not undeserved" sentences. If this range is extremely wide, they would argue, it simply affords us greater flexibility to pursue nonretributivist goods. Section III criticizes limiting retributivism, arguing that, looking only to retributivist considerations, to impose anything other than the *least harmful* "not undeserved" sentence is cruel and unjustified. If we want to impose a punishment *more* harmful than this, I argue, the offender's retributive desert cannot provide the justification; and we would need another argument that explains why harming him for the purpose of, say, deterrence, does not merely sacrifice him for the greater good.

I. VAGUE COMPARISONS

Chang explains that, when comparing two options, one option could not be simply "greater" or "better" than the other. It could only be greater or better at exhibiting a specific value—the "covering value," in her terminology.[4]

2. Norval Morris, Madness and the Criminal Law (1982), at 179.

3. *See id.*; Norval Morris & Michael Tonry, Between Prison and Probation: Intermediate Punishments in a Rational Sentencing System (1990); Richard S. Frase, *Punishment Purposes*, 58 Stan. L. Rev. 67 (2005); Richard Frase; Just Sentencing: Principles and Procedures for a Workable System (2012).

4. Ruth Chang, *Introduction*, *in* Incommensurability, Incomparability, and Practical Reason 1, 4–5 (R. Chang ed., 1997).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

The word "value" in this term can be misleading. A covering value need not relate to moral or ethical value, though, of course, it may. It is simply a predicate. It could be "baseball skill," "generosity," "meanness," "redness," etc. The comparison proceeds by considering the relative degree to which the options exhibit this predicate. In this section, I will argue that the small improvement problem is caused by vague covering values.

Let us consider a well-known Problem case. Who is more creative, Mozart or Michelangelo?[5] Neither seems more creative than the other, but neither seems less creative. It would not resolve our dilemma, however, if we imagined a "small improvement" to Michelangelo's level of creativity (Michelangelo+), whereby he is slightly better at, say, painting eyelashes. Since we would be unable to declare Michelangelo+ more creative than Mozart, we know that Michelangelo and Mozart are not equally creative, continues the argument, because between two genuinely equal options, any improvement to one option, no matter how small, would make it definitively better than the other. What is going on, then? How could we have two options, neither of which is better than the other at exhibiting the covering value, but that nonetheless do not exhibit it equally?

Joseph Raz argues that in such cases—where none of the standard three comparative relations applies (better than, worse than, equal to)—the options must be "incomparable" all the way down, such that how one option exhibits the covering value is irreducibly distinct from the other.[6] The Problem, Raz writes, "speaks not of what does escape reason, but of what must elude it."[7] Chang replies to Raz, I think successfully, that there is no reason to believe that Problem options are incomparable in this deep way. Her most convincing argument is that between any two putatively incomparable options, we could successfully compare a "nominal" version of one option with a "notable" version of the other.[8] This means that the original options are not irreducibly distinct, given that we are comparing the same set of variables in both cases. For instance, imagine that we keep stripping creativity away from Mozart until we reach the skill level of a novice, hobbyist composer ("Novice Composer"). Since we *can* compare Michelangelo and Novice Composer with regard to their creativity—declaring Michelangelo to be more creative—then we know that Michelangelo and Mozart are not incomparable all the way down. Let us assume that this argument works. Where does it leave us? If neither option is better, nor are they equal, *but the options are nonetheless comparable*, then a fourth

5. Chang, *supra* note 1, at 659.

6. JOSEPH RAZ, THE MORALITY OF FREEDOM (1986), at 321–366.

7. *Id.* at 334; *see also* Elizabeth Anderson, *Practical Reason and Incommensurable Goods*, *in* INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 90 (Ruth Chang ed., 1997) (arguing that intrinsic and instrumental goods are fundamentally incomparable).

8. *See* Chang, *supra* note 4, at 14–17; *see also* Laurence Tribe, *Policy Science: Analysis or Ideology?*, 2 PHIL. & PUB. AFF. 66 (1972); James Griffin, *Are There Incommensurable Values?*, 7 PHIL. & PUB. AFF. 39 (1977).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

comparative relation must exist, Chang concludes—beyond "better than," "worse than," and "equal to." This is "parity." On Chang's view, Mozart and Michelangelo are "on a par" with regard to their creativity, all the way down.

A. The Noncomparative Basis of Comparative Judgments

Why should we think that the Problem is caused by vague covering values (Vagueness), rather than the options being "on a par" all the way down (Parity)? The explanation lies in the relationship between *noncomparative* and *comparative* judgments. Imagine that we have two sticks in two different locations, and we want to know which is longer. This will depend, ultimately, on an independent or noncomparative assessment of each stick's length. We examine Stick 1 and see that it is 24.22 inches long. Then we examine Stick 2 and see that it is 23.95 inches long. And thus we conclude that Stick 1 is longer (exhibiting more of the covering value of "length"). If the sticks were in the same location, we could place them side by side to see which was longer, perhaps, but the question would still depend on the degree to which each stick, *considered on its own*, exhibited "length." Now consider Michelangelo versus Novice Composer with regard to their creativity. How creative is Michelangelo, exactly, just considered by himself? This is very hard. All we can say is something like, "He is unbelievably creative." We cannot determine that he has, say, 78.4 units of creativity. Nonetheless, as indicated above, we can conclude with confidence that Michelangelo is more creative than Novice Composer, who is merely, let us say, "somewhat creative." The point is that all we are doing with any comparison is examining the extent to which the options, *independently of each other*, exhibit a particular predicate. This is what we do when comparing Stick 1 and Stick 2 with regard to length, and there is no reason to believe that the comparison between Michelangelo and Novice Composer works differently, even though we cannot measure their independent levels of creativity with cardinal precision.

How does this relate to the question of whether Parity or Vagueness explains the Problem? In this way: if we can make noncomparative measurements with cardinal precision—as in the stick example—then we will never have a Problem case. We will be able to conclude that one option exhibits more of the covering value, or that they exhibit it equally. Problem cases can emerge only when we cannot measure the options independently of one another with cardinal precision. This is true even if—as in Michelangelo v. Novice Composer—not *every* comparison featuring coarse noncomparative judgments is a Problem case. The question then becomes: What causes our noncomparative judgements to be imprecise in Problem cases? Why are our measurements of Michelangelo and Mozart's creativity levels so coarse? Whatever explains this noncomparative coarseness, it seems, will *also* explain why we have a Problem case at the comparative level.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

Thus, if Parity is meant to explain Problem cases—via the conclusion that the options are "on a par" all the way down—it would have to explain our *noncomparative* imprecision in those cases as well. Why is it that we cannot precisely measure someone's level of creativity? That is the most pertinent question for Mozart v. Michelangelo and I am doubtful that Parity, with its comparative focus, has much to say on the matter. To maintain that Problem cases were *not* caused by covering value vagueness, furthermore, it seems that Chang would have to deny that vagueness causes our *noncomparative* imprecision regarding such values. Such an argument seems very unlikely to succeed. It is far more plausible that creativity is a vague concept, the vagueness of which explains our inability to make refined measurements in the noncomparative cases, which in turn leads to an inability to resolve the comparative cases. Vague covering values entail uncertain noncomparative, cardinal judgments that, in turn, entail uncertain comparative, ordinal judgments. We conclude that Mozart and Michelangelo, considered independently of each other, are both something like "unbelievably creative," but nonetheless are almost certainly not precisely equally creative, and then we are at an impasse with regard to our comparative judgement. We can thus conclude that Mozart and Michelangelo are "vaguely equally" creative.

Vagueness, importantly, does not imply that all Problem options are "vaguely equal" in this way. Chang, fully aware of Vagueness as a rival explanation of the Problem, mischaracterizes the argument on this point.[9] Sometimes the "vaguely equal" conclusion is appropriate, but in other instances, the covering value can be so vague that to imply the options are "almost" or "sort of" equal would be misleading, given our uncertainty about their relative worth. Rather than "vaguely equal," they are simply "vaguely related."[10] The covering value at the center of proportionality analysis—"injuriousness"—is one such example, as I discuss in Section II. Specific crimes (e.g., embezzling $250,000) are often only "vaguely related," rather than "vaguely equal," to specific punishments (e.g., 25 months in prison) in terms of their injuriousness.

Finally, if options being "on a par" were *not* a product of vague covering values—if it were indeed as basic a comparative relation as "better than," "worse than," and "equal to"—then we ought to see parity when comparing options with regard to nonvague covering values. But we never do. And the reason is that nonvague covering values, like length and mass, generate precise noncomparative judgments, which will always lead to precise comparative judgments of better than, worse than, or equal to. For the sake of completeness, *epistemic uncertainty* could generate a simulacrum of a

9. Chang, *supra* note 1, at 664–665, 679–688. *See also* JAMES GRIFFIN, WELL-BEING: ITS MEANING AND MEASUREMENT (1986), at 80–81, 96–98, 104 (discussing "rough equality"); THOMAS HURKA, PERFECTIONISM (1993), at 87; JOHN LAIRD, AN ENQUIRY INTO MORAL NOTIONS (1935), at 245–264.

10. *See* DEREK PARFIT, REASONS AND PERSONS (1987), at 430–431 (referring to "rough comparability").

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

Problem case with a nonvague covering value. Imagine that we do not have a ruler and cannot determine with the naked eye whether Stick 1 or 2 is longer. We would then conclude that, given the information we have, we cannot determine whether either stick is longer or if they are equally long. This is a simulacrum of a Problem case because we could make such a determination with complete epistemic information.[11]

B. Vagueness and Substantive Argument

This subsection considers Chang's central argument against Vagueness, against the idea that Problem cases are "borderline" applications of a vague predicate. Understanding how Chang's argument goes wrong, exactly, will provide us with important conceptual tools for assessing the degree to which reasons can determine the choice of a proportionately "injurious" punishment.

Chang's argument centers on a hypothetical. Pretend we have to choose whether a borderline case has passed the relevant threshold (e.g., whether Herbert, a borderline case of baldness, is "bald" or "not bald"), and then, in the context of a Problem case, pretend we have to choose whether one option exhibits more of the covering value, or whether they exhibit it equally.[12] For instance, we would have to decide whether Mozart or Michelangelo is more creative, or whether they are equally creative or—to use a different example of Chang's—whether a particular act of promise keeping or a particular act of bringing about great happiness exhibits more "moral goodness," or whether they exhibit it equally. Chang argues that we resolve the borderline and Problem cases differently.[13] While in borderline cases like bald-or-not-bald, she writes, "it is perfectly permissible to resolve the indeterminacy in favor of application or not by arbitrary stipulation," in Problem cases only "substantive" argument could resolve the comparison permissibly.[14] If people disagreed on how to classify the Mozart or promise-keeping comparisons, there would be, Chang writes, "a genuine substantive disagreement between them, not a mere clash of arbitrary

11. There are theorists who argue that epistemic uncertainty, rather than semantic indeterminism, is responsible for vagueness itself, such that there are precise, cardinal answers in terms of Mozart and Michelangelo's creativity levels, but we simply lack the epistemic resources to discern them. *See, e.g.*, Timothy Williamson, Vagueness (1994). I will continue in this paper on the assumption that vagueness is caused by semantic indeterminism, that the difficulty of Mozart v. Michelangelo is grounded in the imprecision of "creativity" all the way down. Nonetheless, in terms of the degree to which we can use reason to resolve vague practical dilemmas—which is our ultimate aim in this section—very little hangs on whether vagueness is semantic or epistemic. *See infra* note 24. On vagueness as semantic indeterminism, *see, e.g.*, Kit Fine, *Vagueness, Truth and Logic*, 30 Synthese 265 (1975).

12. Chang, *supra* note 1, at 679–686. This methodology, of pretending we are "forced" to choose, is rather probative. For in our everyday lives, we are "forced" to choose between Problem options all the time—Should I play basketball or read a novel? Should I support Candidate A or Candidate B?—with the pragmatic principle that some choice will be better than none doing the "forcing."

13. *Id.* at 682.

14. *Id.* at 682–683.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

decisions."[15] Chang then maintains that no conception of semantic vagueness explains how we could employ substantive debate, as opposed to arbitrary stipulation, to resolve vagueness.[16] Given that substantive debate will concern the meaning and interpretation of the relevant concepts, and given that vagueness entails that the meaning of such concepts will not determine the answer, what could there be to debate substantively? Thus, Chang concludes, we know that vagueness does not cause the Problem: only substantive argument could resolve such cases acceptably and substantive argument is incompatible with the resolution of vagueness.

As to Chang's first claim—that only substantive argument, as opposed to arbitrary stipulation, could acceptably resolve Problem cases—consider Herbert and Greg, two borderline cases of baldness. We want to know who is balder. While Herbert is balder in the front, Greg is balder in the back. We might imagine a small improvement problem, where making Herbert a touch balder fails to resolve the comparison.[17] Forced resolution of Herbert v. Greg would, I believe, be a matter of arbitrary stipulation, rather than substantive disagreement. If two people disagreed, they would have nothing to debate. If this holds, then we have determined, contra Chang, that resolution of Problem cases can be a matter of arbitrary stipulation.

As to Chang's more important second claim—that substantive argument is incompatible with the resolution of vagueness—I will attempt here to explain the possibility and, crucially, the limits, of providing reasons and substantive arguments to resolve vagueness as a practical matter. In the first instance, I am searching for an explanation that does not appeal to "external" considerations beyond the particular covering value, such as which choice meshes better with one's character or life plan.[18] I discuss this possibility further below. To be sure, Chang's claim was not meant to cover this possibility; her claim is that, looking only to the meaning of the covering value, we cannot use reason and substantive argument to resolve vagueness.

Chang explains that covering values have constitutive "contributory values."[19] The contributory values of "baseball skill," for instance, would be foot speed, ability to catch a baseball, ability to hit a curveball, etc. To compare two people's skill at baseball, we are said to examine the extent to which each person exhibits these contributory values. The degree to which we can employ reason to resolve a Problem case, I argue, depends

15. *Id.* at 685.

16. *Id.* at 686.

17. Johan Gustaffson introduces a baldness Problem case. Johan E. Gustaffson, *Indeterminacy and the Small-Improvement Argument*, 25 UTILITAS 433, 436 n. 10 (2013).

18. For analysis of the various methods of resolving hard comparisons beyond considerations inherent to the particular covering value, *see* Francisco J. Urbina, *Incommensurability and Balancing*, 35 OXFORD J. L. STUD. 575, 596–603 (2015).

19. Chang, *supra* note 4, at 5.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

on the nature of this covering value–contributory value relationship. When a vague covering value has one or two well-defined contributory values, as with the covering value of "baldness" and the contributory value of "hair coverage"—which we could possibly break down into "front of the head hair coverage" and "back of the head hair coverage"—there is nothing to substantively debate. We either think Herbert is bald, and balder than Greg, or we don't. We will report such beliefs as if they were scientific measurements.

However, when a vague covering value has numerous, interdependent covering values—as in the case of "creativity" and "moral goodness"—substantive debate involving the provision of reasons may *sometimes* be possible when forced to resolve the comparison. Let us call such covering values *organic* covering values. The term is taken from G.E. Moore's theory of "organic unities," whereby intrinsic value often belongs to "organic" complexes, the value of which "bears no regular proportion to the sum of the values of its parts."[20] Organic covering values are vague for a number of reasons, in particular: we cannot precisely list the relevant contributory values; we cannot precisely state their relative importance; and we cannot precisely understand how they interdepend on one another. To demonstrate the interdependence of some contributory values, notice that improving one's eyesight would fail to improve his "baseball skill"—understood as the ability to increase the chances, on average, that a randomly selected baseball team will win a baseball game—if he lacked the ability to hold up a bat or to stand up. It would be an improvement to a *particular* baseball skill—an ability to perceive balls and strikes—but in a manner that does not impact the covering value of baseball skill *simpliciter*. Certain contributory values, in this way, are relevant only in the presence of other contributory values.[21]

To see how we can sometimes provide reasons to resolve Problem cases centered on an organic covering value, consider a modified version of Raz's classic Problem case.[22] We have to choose one of two career options for Sarah: Financially Poor Clarinetist (FPC) or Miserably Bored Lawyer (MBL). The covering value at issue—"career goodness"—has numerous, often interdependent contributory values: skill in one's profession; career earnings; possibilities for advancement; work hours; types of people one will interact with, etc. That we are being forced to resolve such cases as a practical matter is important. Our real, theoretical conclusion as to whether FPC or MBL is the better career is clear: it's vague.[23] On some "weightings"

20. G.E. Moore, Principia Ethica (1903), at 27–28.

21. Below I explain how "injuriousness," the covering value at the center of proportionality analysis, is doubly organic, insofar as its interdependent contributory values each have their own set of interdependent contributory values. *See infra* Section II.C.

22. Raz, *supra* note 6, at 328.

23. As Chang writes, "[o]f course in one way, there is already a 'resolution' in a borderline case: it is borderline." Chang, *supra* note 1, at 682.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

of the contributory values of "career goodness," FPC is the better choice for Sarah, while on others MBL is better. But it seems possible to accept this conclusion and nonetheless debate FPC v. MBL substantively when forced to choose.

The explanation resides in the meaning of being forced to choose. It means, I believe, that one has to choose a "weighting" of the contributory values that will resolve the comparison. In the case of FPC v. MBL, maybe one chooses a weighting that prioritizes one's passion for her work, which leads to selecting FPC, or alternatively one that prioritizes financial security, which leads to MBL. The key is that it is possible to provide reasons for choosing one such weighting over another, and it is consistent with doing so that one accepts that each weighting on its own misses something fundamental about the meaning of "career goodness." By accepting the possibility of multiple weightings, one is not committed to their absolute and context-independent equality at capturing the covering value's meaning. And when forced to choose, substantive debate concerns the following question. Among the weightings that could resolve this particular comparison in this particular context, which does the best job of capturing the meaning of the covering value, of, say, career goodness? In this way, we might describe the debate as being over which weighting is the *least arbitrary*. An important caveat is that some weightings would fail to resolve a comparison—like, for many people with regard to FPC v. MPL, a weighting that prioritized intellectual pursuit—and would thus be ineligible for debate when one was forced to choose.[24]

This method has rather severe limitations. Sometimes, when considering an organic covering value, its contributory values will be very difficult to conceptualize, or they will be entirely unknown, such that discovering any coherent "weighting" to argue for would be impossible or artificial. I suspect that this would be the case with regard to Mozart v. Michelangelo and "creativity." While creativity is surely an organic covering value, with various considerations that interact in complex ways, the number and nature of these considerations is generally unclear. To create a list of contributory values that we ought to consider in judging someone's creativity would for the most part be forced and unhelpful—perhaps even dishonest. There are limits to the usefulness of Pro v. Con lists. At other times, one may be able to

24. As indicated above, I am focusing here on methods of resolving semantic vagueness, but this could be a method of resolving epistemic vagueness as well (if vagueness was in fact epistemic). On the epistemic view, there is a precise answer, in reality, as to whether FPC or MBL exhibits more "career goodness." The method, then, would not be about finding the weighting of contributory values that captures the greatest amount of the covering value's meaning. It would rather be about reaching some sufficient level of confidence that could serve as an *evidentiary tie breaker*, as was incisively suggested by an anonymous reviewer. So, the epistemicist would admit that she does not *know* the answer to FPC v. MBL, but given that there must in fact *be* an answer, a good argument in favor of a particular weighting would mean that there is now sufficient reason to choose either FPC or MBL, rather than to merely flip a coin.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

conceptualize the relevant contributory values, but there will be so many of them, and they will be so similarly weighted, and perhaps so vague and slippery themselves, that one simply cannot reach a considered judgment on the comparison. One's level of confidence in determining which weighting captures the greatest amount of the covering value's meaning will be low, perhaps vanishingly so. Rarely does life present vagueness as neatly as it does in some of Chang's Problem examples or in FPC v. MBL, where the considerations at issue are so limited and well-defined. In FPC v. MBL, for instance, we focused on only two contributory values, both of which have relatively clear meanings: financial security and passion for one's work.

The larger point, which I will return to when considering retributive proportionality, is that vagueness is on a spectrum, and some covering values and comparisons are vaguer than others. Substantive argument will not take us far when dealing with a certain level and type of vagueness, even if, contra Chang, it has some role to play elsewhere in the context of resolving vagueness. Even with those comparisons where substantive argument could have a role to play, like FPC v. MBL, its insights are rather coarse. For instance, if one had to choose the salary level of MBL that created an option equal to FPC in terms of career goodness, then substantive argument would not take us far; no weighting would seem to determine the matter. The comparison demanded by sentencing proportionality is structured in this manner (e.g., how many days in prison equals punching someone in the arm?).

C. External Reasons

Finally, where we have a practical dilemma centered on a vague covering value, and we cannot discern a "least arbitrary" weighting of the contributory values that will resolve the comparison, reason may have yet another arrow in its quiver when seeking the most justified choice. As suggested above, it may sometimes be possible to appeal to "external" reasons, which are relevant to the choice, but not to the options' relative covering value exhibition. Even if we cannot declare a possible clarinet career better than, worse than, or equal to a possible law career, we might be able to use reason to choose between the two options by taking into account considerations that are not themselves relevant to "career goodness" as normally understood. Imagine that pursuing the clarinet career would require one to move cities, which would in turn very negatively impact the education of one's children. That the clarinet career is worse for one's children does not make it any worse at exhibiting "career goodness," but it does provide one with *a reason* to choose the law career over the clarinet career. Of course, this reason may prove dispositive even if the clarinet career were definitively better than the law career, such that one had reason to pursue a worse career than she might have otherwise, with that reason being her children's well-being or her duties to her children. Another example of an "external" consideration is how the choice meshes with one's past

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

commitments and conception of the good life as a whole, as Charles Taylor explains.[25] Along these lines, when choosing between vaguely equal or related options, we may have reason to select the one that enables us to lead a life of integrity or narrative unity (e.g., we have thus far been committed to a life of music and, even though neither option is better *as a career*, to choose the law over the clarinet career would dishonor that past). These analyses will prove important in Section III, where I consider which "external" reasons a retributivist state might appeal to when selecting an option within the wide array of vaguely proportional sentences.

Section II has presented several arguments in favor of Vagueness, the idea that the "small improvement problem" is caused by vague covering values. I have attempted to unpack the structure of comparisons, to show how they rely ultimately upon *noncomparative* judgments; and that when we cannot resolve a comparison, it means that there is indeterminacy within our noncomparative judgments. Covering value vagueness, I have argued, generates such noncomparative indeterminacy, which means, in turn, that it generates our comparative uncertainty. I have also addressed Chang's argument that since the resolution of borderline cases must be a matter of arbitrary stipulation and the resolution of Problem cases is a matter of substantive argument, we know that vagueness does not cause the Problem. In reply, I have (1) presented an example of a Problem case, the resolution of which is a matter of arbitrary stipulation and (2) explained how reason can sometimes—to some degree—resolve a vague comparison, either by discerning a "least arbitrary" weighting of the contributory values that will resolve the comparison in a particular context, or by discerning dispositive "external" reasons.

II. PROPORTIONAL SENTENCING

Wherever the concept of "proportionality" emerges within the law—most strikingly within criminal and constitutional law—we will find a vague comparison. It is one instance of the ubiquity of vagueness and vague comparisons within the law.[26] Here I will examine the comparison at the center of the ideal of the retributively proportional criminal sentence.[27] This ideal

25. *See* Charles Taylor, *Leading a Life*, *in* INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 170 (Ruth Chang ed., 1997). *See also* Michael Stocker, *Abstract and Concrete Value: Plurality, Conflict, and Maximization*, *in* INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 196 (Ruth Chang ed., 1997) (arguing that alternatives are rarely considered in the abstract, but rather are usually considered in more concrete ways, such as whether they contribute to one's own good life).

26. *See generally* TIMOTHY ENDICOTT, VAGUENESS IN LAW (2000).

27. I will not consider here how the above sections relate to constitutional proportionality analyses, which involve "balancing" various constitutional considerations in determining the legality of state action. On this issue *see generally* Aharon Barak, *Proportionality and Principled Balancing*, 4 LAW & ETHICS HUM. RTS. 1 (2010); Timothy Endicott, *Proportionality and Incommensurability*, *in* PROPORTIONALITY AND THE RULE OF LAW: RIGHTS, JUSTIFICATION, AND REASONING 311 (Grant Huscroft, Bradley W. Miller & Gregoire Webber eds., 2014); Paul

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

provides that an offender's punishment ought to be "proportional" to her moral desert, which the "gravity" of her offense determines.[28] The principle is a jewel in the retributivist crown, as the reasons justifying the infliction of punishment seem to themselves generate sensible punishment limits: the reason to inflict punishment is to give the offender the suffering or the censure that she deserves, and an offender only deserves a punishment proportional to the gravity of her offense.[29] A purely consequentialist punishment theory, by comparison, lacks internal punishment limits tied to the seriousness of offenses, and would license extreme or degrading punishments, even for trivial offenses, should inflicting them maximize crime prevention, happiness, or whatever it may be. This section, however, uses the above analyses to question proportionality's ability to live up to its promise for retributivists.

Proportionality relies upon a very hard comparison to deliver its practical guidance: punishment v. offense gravity. I argue that the covering value at issue—"moral gravity" or, as I will explain below, "injuriousness"—is too vague to provide meaningful sentencing guidance. Related, that some jurisdictions contain complex matrices of precedents and guidelines on proportional sentencing does not diminish the importance of the questions pursued here. For proportional sentencing law presents itself as standing, to some essential degree, upon objective moral analysis about what an offender retributively deserves as a response to his wrongdoing. And if

Craig, *The Nature of Reasonableness Review*, 66 Current Legal Probs. 131 (2013); Robert Alexy, *The Construction of Constitutional Rights*, 4 Law & Ethics Hum. Rts. 20 (2010); Gregoire Webber, *Proportionality, Balancing, and the Cult of Constitutional Rights Scholarship*, 23 Can. J. L. & Juris. 179 (2010).

28. "Censuring" retributivists, like Antony Duff, as well as "traditional" or "strict" retributivists, like Michael Moore, are committed to proportionality. As Duff writes:

> A requirement of proportionality is intrinsic to any theory on which the, or a, primary purpose of punishment is to communicate the censure that offenders deserve for their crimes. We must determine not just that an offender deserves censure but how severe that censure should be: the more serious the crime, the more severe the deserved censure. That censure is communicated by penal hard treatment, and severity is a dimension of penal hard treatment as it is of censure. Thus the severity of the penal hard treatment will communicate the severity of the censure: the more severe the hard treatment, the more severe the censure it communicates. But it is then a simple requirement of justice (and of communicative honesty) that the severity of the offender's punishment (as penal hard treatment) be proportionate to the seriousness of her crime. To punish her with disproportionate severity, or leniency, is to communicate to her more, or less, censure that she deserves. This is, however, dishonest and unjust, since it is to punish her more, or less, severely than she deserves.

R.A. Duff, Punishment, Communication, and Community (2001), at 132 (internal citations omitted).

29. Since the mid-1970s this ideal has grown increasingly influential within American, British, and Nordic sentencing policy, to a significant degree displacing concerns with rehabilitation, deterrence, and incapacitation. *See* David Garland, The Culture of Control (2001); David Garland, Punishment and Welfare (1985); F. A. Allen, The Decline of the Rehabilitative Ideal (1981).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

this paper succeeds in questioning the efficacy of such analysis, then we have reason to doubt the moral foundations of the law. In presenting this view, I aim to supplement and clarify the philosophical basis of Nicola Lacey and Hannah Pickard's recent conclusion that proportionality is a "chimera," incapable as an abstract ideational device of determining sentences, as evidenced in part by the dramatic variations in punishment severity within contemporary and historical communities looking to apply retributively proportional sentences.[30] Lacey and Pickard argue that "what has been thought of as proportionality is not a naturally existing relationship, but a product of political and social construction, cultural meaning-making, and institution-building."[31]

Finally, I aim in this section to continue my critique of Parity by demonstrating its inability to deliver this criticism of proportionality analysis, due to its inability to distinguish between Problem cases. Parity deems all Problem options "on a par" full stop, while in reality—and crucially for making practical decisions based upon those comparisons—the relationship between such options varies with the level of the covering value's vagueness. Some Problem options, like a punishment and a crime with regard to injuriousness, are much less closely related than others. In Section III to follow, I consider the implications of the vagueness of retributive proportionality for retributivism more generally, concluding—contra "limiting" retributivism—that retributivism can only justify the *least harmful* sentence on the wide scale of "not undeserved" punishments.

A. Cardinal Proportionality

There are a few foundational issues to resolve before setting the comparison at the center of proportionality analysis. First is the distinction between ordinal and cardinal proportionality. "Ordinal" proportionality is the ideal that the scale of punishments ought to reflect the scale of offense severity, so that the "graver" the offense, the more severe the punishment. This is contrasted with "cardinal" proportionality, which is meant to "anchor" this scale by determining, for a minimum of one or possibly two offenses,[32] the absolutely deserved punishment, that is, the punishment deserved only by reference to the offense's moral gravity. The ordinal scale is set, in theory, by reference to these few points of cardinal judgment. Without such cardinal anchors, ordinal proportionality loses much meaning, as many commentators have noticed, since its principles would be realized even if the scale were set at an extremely high level, where minor regulatory offenses receive 100 lashes, and on up for the worst offenses, or at an extremely low level, where the worst offenses receive a $100 fine, and on down to a 10-cent fine.

30. *See* Nicola Lacey & Hanna Pickard, *The Chimera of Proportionality: Institutionalising Limits on Punishment in Contemporary Social and Political Systems*, 78 MOD. L. REV. 216, 227–228, 231 (2015).
31. *Id.* at 216.
32. *See* Greg Roebuck & David Wood, *A Retributivist Argument Against Punishment*, 5 CRIM. L. & PHIL. 73 (2011).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

Retributivists thus cannot escape the complexities of proportionality by appealing only to the less contentious ordinal conception. As Lacey and Pickard write, "[T]he very idea that a certain punishment is deserved—and with it, the whole basis for proportionality's purported capacity to set upper limits on, or to generate substantial criteria of fittingness of, punishment—rests on cardinal rather than ordinal proportionality."[33] Retributivists who accept this point, and set out to fix a point or two on an absolute scale, miss a larger criticism: if we can determine the cardinally proportional punishment for any offense, in order to set the ordinal scale, we should be able to do the same for every offense, and thus we should not require the ordinal conception at all. One might reply that the cardinal punishments for certain offenses—maybe murder—are easier to determine than others. Regardless, the takeaway point is that proportionality requires the establishment of at least some cardinal points, and so we will focus here on that possibility. Hereafter, by proportionality I refer to the cardinal conception.

B. Proportionate Injuries

How could one year in prison or a $1,000 fine be "proportionate with" or "equal to," in any sense, punching someone? In determining the magnitude of punishment necessary to "negate" a crime, Hegel attempts to provide such an explanation. He argues that the "identity" of crime and retribution "is not an *equality* in the specific character of the infringement" but an equality of "*value*."[34] Hegel argues that the equivalence between punishments and crimes can be found at an abstract level, in terms of their "character as injuries."[35] This concept, which I refer to as "injuriousness," and which is synonymous with "moral gravity," is the relevant covering value, according to Hegel. He explains that the concept has "a determinate qualitative and quantitative magnitude."[36] By "qualitative magnitude," Hegel refers to the offender's culpability level (e.g., intent, knowledge, recklessness, and negligence). The talionic, eye-for-an-eye methodology of retributive proportionality, by comparison, considers only the "quantitative magnitude" of the harm caused, at least as traditionally interpreted.[37] It would not matter whether the offender intentionally poured or recklessly spilled acid into her victim's eye; in both cases we ought to intentionally

33. *See* Lacey & Pickard, *supra* note 30, at 227.

34. G. W. F. Hegel, Elements of the Philosophy of Right (Allen W. Wood ed. & H. B. Nisbet trans., 1991), at 127 (§101).

35. *Id.* at 128 (§101).

36. *Id.* at 127 (§101).

37. *See* Black's Law Dictionary (6th ed. 1990), at 913 (Lex talionis "requires the infliction upon a wrongdoer of the same injury which he has caused to another" and is "expressed in the Mosaic law by the formula, 'an eye for an eye; a tooth for a tooth,' etc."). The Bible is in fact ambivalent with regard to the talionic ideal, calling for non-talionic punishments in a number of instances. *See Exodus* 21-22, *in* The Holy Bible: King James Version (2010) ("And he who curses his father or his mother shall surely be put to death," *id.* at 21:17).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

pour acid into her own eye in response. For modern retributivists like Hegel, though, the gravity of an offense, and thus the severity of the punishment demanded by proportionality in response, increases not only with the amount of harm caused, but also with the degree of the offender's culpability.

Proportionality demands some ratio, some *proportion*, of crime injuriousness to punishment injuriousness. Classically, the ratio is 1:1—a proportion of equality. A proportional punishment will exhibit, on this view, no more (and, ideally, no less) injuriousness than the underlying offense. Though, accepting the concern with "qualitative magnitude," we might conclude that a proportional punishment should be *less* harmful than the offense. The rationale is that the impact of culpability goes both ways, and we need to account for the fact that criminal punishments are thoroughly intended by the state. Thus, if an offender causes 10 units of harm due to her recklessness, then her proportional punishment would be less than 10 units of harm, given that these units would be delivered intentionally. We might imagine, following Camus, that even intentional crimes like murder cannot match the heavy purposiveness and deliberateness of the criminal process, such that even for intentional crimes, the amount of penal harm must be less than that inflicted on the offender's victim.[38] To be clear, on the view considered here, where we limit the punishment's harm due to these culpability issues, the ratio of punishment to offense injuriousness remains 1:1. Even if the proper ratio is greater or lesser than 1:1, such that the injuriousness of the punishment should be greater or lesser than that of the offense, proportionality still requires a consistent ratio up and down the offense spectrum. That is, an extra "unit" of offense injuriousness should lead to the same amount of additional punishment injuriousness, regardless of the starting point along the spectrum.

We now have a better grasp of the relevant comparison. In setting a proportional punishment, a decision-maker—say, a judge employing her sentencing discretion or a lawmaker writing sentencing guidelines—is tasked first with assessing the injuriousness of the offense, understood as a combination of the harm caused and the offender's culpability. The next step is to take this "value" and then compare which of the available punishments exhibit the same degree of injuriousness, or whatever greater or lesser degree is required by a different ratio.

C. Analyzing the Comparison

The comparison of crimes and punishments with regard to the covering value of injuriousness will generate Problem cases, as indicated at the outset. Which is more injurious, the crime of embezzling $250,000 from a pension fund (Emb250) or the punishment of 25 months' imprisonment (25Mo)? Or perhaps they are equally injurious, such that 25 months is the perfect retributivist punishment if the ratio is 1:1? This is a hard

38. ALBERT CAMUS, RESISTANCE, REBELLIONS, AND DEATH (J. O'Brien trans., 1974), at 199.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

comparison. As stated above, if we imagine that the offender embezzled only $240,000, it would not help us to resolve the comparison, nor would it if she received only 24 months' imprisonment.

Consistent with his analysis of other Problem cases, Raz would deem Emb250 and 25Mo incomparable, all the way down. They are injurious in such different ways, he might argue: one affects many people a little financially, while the other affects one person significantly in terms of her liberty over a period of time. Unless the crime matched up very closely with the form of punishment—e.g., the crime of false imprisonment and the punishment of imprisonment—it seems that Raz would have to deny the possibility of *any* proportionality judgments, and thus would reject the ideal of proportional punishments entirely. However, as suggested above, no two options will be incomparable in this deep way. For instance, if we take Emb250 as our "standard," broadly following John Broome, and imagine a spectrum of sentences from one nanosecond at the bottom, to life without parole at the top, there will be a point in this spectrum above which the prison terms are clearly more injurious than Emb250, and a point below which they are clearly less injurious than Emb250.[39] Emb250 is thus not *irreducibly distinct* from 25Mo, given that we can confidently compare Emb250 to options that exhibit the same variables as 25Mo (only in different degrees).

What then is going on with this comparison, where neither option is more injurious, nor are they equally injurious? We know from Mozart v. Michelangelo that the difficulty results from our uncertainty over how much injuriousness a crime and a punishment exhibit when considered *independently* of each other. How injurious, exactly, is embezzling $250,000 from a pension fund? And 25 months of imprisonment? Our final judgments will not involve any degree of cardinal precision such that we could conclude that one is injurious to, say, 58.6 degrees and the other to 42.7 degrees. This noncomparative imprecision results, in turn, from the vagueness of "injuriousness," following the above discussion.

Let us unpack its vagueness. Injuriousness, consistent with Hegel's interpretation, has two core contributory values: harm and culpability. Andrew von Hirsch and Andrew Ashworth provide a convincing theory of harm as that which reduces someone's "standard of living"—the "means and capabilities that would *ordinarily* promote a good life."[40] It is only one conception of harm, of course, but it is a thoughtful and prominent conception, and there is no reason to believe that any other honest attempt could avoid the problems that beset their theory.[41]

39. John Broome, *Is Incommensurability Vagueness?*, *in* INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 67 (Ruth Chang ed., 1997).

40. ANDREW VON HIRSH & ANDREW ASHWORTH, PROPORTIONATE SENTENCING: EXPLORING THE PRINCIPLES (2005), at 145.

41. Thanks to an anonymous reviewer for raising this point. *See, e.g.*, JOEL FEINBERG, THE MORAL LIMITS OF THE CRIMINAL LAW, VOLUME 1: HARM TO OTHERS (1987), at 31 (defining a "harm" as a "setback to interest").

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

Von Hirsch and Ashworth argue that "standard of living" has four main components: (1) physical integrity, (2) material support and amenity, (3) privacy, and (4) freedom from humiliation.[42] "Standard of living," so defined, is *doubly organic*, because while these contributory values are undoubtedly interdependent, they each in turn have their own set of interdependent contributory values. For example, many instances of material support and amenity (e.g., having a home) depend upon privacy for their impact upon one's living standard; meanwhile, material support and amenity is itself a complex, interdependent mixture of goods that depend on each other for their impact (e.g., food, shelter, clothes, liquid assets, communication devices, access to transportation, access to markets, etc.). Thus, while reduction of living standard seems to be the right metric for measuring harm, it is exceptionally vague. The room for reasonable disagreement over which offense is more significant than another in terms of its impact on one's living standard is expansive. Compare, for instance, the impact of a humiliating crime like being spat on by a stranger to that of a financial fraud that impacts one's material support, or to that of a burglary that impacts one's privacy in addition to one's material support. Finally, even if we were able to emerge from this harm analysis unscathed by vagueness—which we cannot—we would need then to combine it, somehow, with the culpability judgment. H.L.A. Hart expresses this final worry succinctly: "Is negligently causing the destruction of a city worse than the intentional wounding of a single policeman?"[43] In this manner, we see how limited proportionality is as a device for determining criminal sentences. It requires precise measurements of "injuriousness," which, due to the vagueness of that concept, cannot be determined.

We have here a vague comparison, moreover, that cannot be refined significantly by substantive argument when we are forced to decide, at least not when that argument is focused on the covering value itself, as opposed to "external" considerations. Of the contributory values that we can conceptualize, they are too numerous, too similarly weighted, and too vague and interrelated for us to be able to reach a considered judgment as to which weighting does the best job of capturing the meaning of injuriousness in a given case. We must add to this complexity a dose of humility regarding our ability to even conceptualize all the relevant considerations to be incorporated into our weighting. Any claim that a particular weighting is best—e.g., one that prioritized humiliating harms and ignored culpability—would be distinctly artificial. It is doubtful, furthermore, that any weighting could resolve the question of how many days in prison, or how large a fine, equals or is proportionate to the injuriousness of a particular offense. Substantive argument, accordingly, would not take us very far in the large majority of cases. What we are left with, if we appeal to proportionality alone, is an

42. VON HIRSH & ASHWORTH, *supra* note 40, at 145.
43. H.L.A. HART, PUNISHMENT AND RESPONSIBILITY (1968), at 162.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

extraordinarily wide zone of discretion. And we should not be surprised, then, that Swedish and American penal systems have such dramatically different sentencing scales, even though they both look to the concept of retributive proportionality for authoritative sentencing guidance.[44]

This is not to say, however, that proportionality provides no sentencing guidance or limits at all. Just as we could determine that Michelangelo was conclusively *more creative* than Novice Composer, even though creativity is a vague, organic covering value, there are undoubtedly situations in which a form of punishment (say, twenty years in prison) is conclusively *more injurious* than a particular offense (say, theft of a bicycle).[45] A twenty-year sentence would be deemed disproportionate to a bicycle theft in *any* system committed to retributive proportionality. Thus, while injuriousness is very vague, it nonetheless has *some* level of conceptual definition. Though, we should emphasize that such retributive upper limits would be largely irrelevant in the context of serious offenses, where severe or even degrading punishments would reside within the wide array of proportionally injurious sentences.[46]

44. *See* Lacey & Pickard, *supra* note 30, at 217, 227–228, 231.

45. *See* Duff, *supra* note 28, at 132 ("Such intuitions as we have about proportionality do seem to include intuitions about absolute, as well as relative, proportionality. Five years' imprisonment would be a grossly disproportionate sentence for a parking offense, and a fine of 10 dollars a grossly disproportionate sentence for a brutal murder.").

46. What about "fair play" retributivism? Perhaps that theory could provide more definite sentences? Fair play retributivists reject the traditional retributivist idea that an offender, by committing a moral wrong, deserves to suffer or to be censured "because, and only because" it is morally deserved. Michael S. Moore, *The Moral Worth of Retribution*, *in* Responsibility, Character, and the Emotions: New Essays in Moral Psychology 179, 179 (Ferdinand Shoeman ed., 1987). Rather, they understand penal desert to derive from a commitment to fairness. If we assume that an offender has benefitted from everyone else's restraint in following the law—not always a safe assumption, Jeffrie Murphy argues—then he has gained an unfair advantage by breaking the law and failing to restrain himself in turn; and the harm or suffering of punishment is thus deserved as a means of stripping away the offender's unfair gain. *See* Jeffrie G. Murphy, *Marxism and Retribution*, 2 Phil. & Pub. Aff. 217 (1973); Herbert Morris, *Persons and Punishment*, 52 Monist 475 (1968); Richard Dagger, *Playing Fair with Punishment*, 103 Ethics 473 (1993). *See also* George Sher, Desert (1987), at 69–90. There are number of well-documented, I think fatal problems with this view, and so I will not examine in detail the nature of fair play proportionality. Most importantly, fair play retributivism requires the position that *mala in se* offenders have indeed profited from their offenses, "that the unpunished rapist is really better off than men and women who resist the temptation or feel no temptation to rape," as M. Margaret Falls writes. M. Margaret Falls, *Retribution, Reciprocity, and Respect for Persons*, 6 Law & Phil. 25, 31–32 (1987). There is the further problem that we seem less tempted by *mala in se* offenses than by *mala prohibita* offenses, as Antony Duff argues. Would fair play theory entail that the former deserve less punishment than the latter? *See* R.A. Duff, Trials and Punishments (1986), at 213. Duff writes: "[T]alk of the criminal's unfair advantage implies that obedience to the law is a burden for us all: but is this true of such *mala in se*? Surely many of us do not find it a *burden* to obey the laws against murder and rape, or need to *restrain* ourselves from such crimes: how then does the murderer or rapist gain an unfair advantage over the rest of us, by evading a burden of self-restraint which we accept?" *Id.* Regardless, we should not think that fair play sentencing could proceed with precision. How many days in jail is injurious enough to erase the advantage that someone gained by punching someone? Thus, instead of comparing the injury of the offense with the injury of the punishment, fair play proportionality would demand that we compare an injury (the punishment) with a *benefit* (the offense) in

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

D. Parity and Proportionality

Parity is less capable than Vagueness of delivering this critique of retributive proportionality. While Parity *could* inform us that the zone of sentencing discretion is wide, such that an offense is "on a par" with a wide range of punishments, it would mischaracterize the nature of the comparative relationships inside the zone of discretion. For parity, as conceived by Chang—and as understood colloquially—is connected closely to a judgment of equality; it is meant to explain, in particular, cases of seemingly "rough" or "imprecise" equality, as in the case of Mozart v. Michelangelo.[47] Chang writes of parity emerging "between items that bear very different aspects of V and yet are nevertheless 'in the same neighbourhood' of V-ness."[48] But sometimes the covering value (the V) can be so vague, as indicated above, that deeming Problem options "on a par," given this connection to equality, is inapposite.[49] It fails to illuminate to deem embezzling $250,000 "roughly" or "imprecisely" equal to 25 months in prison with regard to injuriousness, given that it may also be "roughly" equal to 2 months or 10 years. The notion of rough or imprecise equality in this case—and thus of the options being "on a par"—has little purchase. The "neighbourhood" of V is so large that it does not help to determine that both options are residents.

Related, Parity is incapable of deeming options "*really* on a par" or "*a little* on a par." Parity is either on or off, full stop, just like genuine equality; options cannot be *really* equal or *a little* equal. Thus, if Mozart is only "on a par" with a handful of sculptors in terms of creativity, while Emb250 is "on a par" with a wide array of punishments in terms of their injuriousness, Parity cannot distinguish between individual Problem cases based upon these comparisons. It cannot inform us that Mozart is much more closely related to Michelangelo in terms of his creativity than Emb250 is to 25Mo in terms of its injuriousness. The picture that Parity draws is of close comparisons—*almost* equality—up and down the zone of sentencing discretion, as if comparing Emb250 with all the sentences proportional to it is like comparing Michelangelo with many slightly different versions of Mozart. Vagueness, however, explains that, depending on the covering value's level of vagueness, a comparison can be a *little* vague or, as with regard to injuriousness, *really* vague. We have little sense of how a crime and a punishment relate with respect to injuriousness, except at the far extremes. Vagueness, unlike Parity, can account for such situations, where we know

terms of their impact (positive or negative) on one's well-being or living standard. If anything, this comparison is *more* difficult than the one demanded by traditional retributivism, which at least has the advantage of comparing two forms of injury. I am grateful to an anonymous reviewer for pressing me to consider this issue.

47. *See* Ruth Chang, *Parity, Imprecise Comparability and the Repugnant Conclusion*, 82 THEORIA 182 (2016).

48. *Id.* at 193.

49. For relevant discussion *see* Joshua Gert, *Value and Parity*, 114 ETHICS 492, 506–508 (2004).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

that the options are comparable, but our judgments of their relative worth are profoundly uncertain; and, unlike Parity, it can warn us that such comparisons, when engineered to deliver ethical or practical guidance, are of limited value.

Let us take stock. I have argued, contra Ruth Chang and Joseph Raz, that the small improvement problem is caused by vague covering values, not "parity" or "incomparability." I have then applied this analysis to the retributivist ideal of the proportional penal sentence. I have attempted to unpack the comparison between a crime and a punishment upon which this ideal is based, arguing that the relevant covering value is "injuriousness," and that the vagueness of this concept generates Problem cases. I have argued that "injuriousness" is not only vague, but exceptionally vague, featuring multiple layers of interdependent covering values, such that proportionality analysis is incapable of providing meaningful sentencing guidance or limits.

III. LIMITING RETRIBUTIVISM

So what? If retributive proportionality is considerably looser than we have appreciated, what does this mean for retributivism as a justification of state punishment? "Limiting" retributivists, like Norval Morris, Michael Tonry, and Richard Frase, would claim to be untroubled, as indicated at the outset. They have always emphasized—and embraced—the vagueness of retributive proportionality, they would argue, and it does not impact their argument that it might be *extremely* vague. Morris writes:

> "[A] deserved punishment . . . does not mean the infliction on the criminal offender of a pain precisely equivalent to that which he has inflicted on his victim; it means rather a *not undeserved* punishment which bears a proportional relationship in a hierarchy of punishments to the harm for which the criminal has been convicted."[50]

Within the range of "not undeserved" punishments that bear a "proportional relationship" to the offense, limiting retributivists argue that (a) any punishment is retributively legitimate and (b) nonretributivist considerations—like deterrence, rehabilitation, incapacitation, social norm maintenance, and so forth—should determine the choice of punishment. In this way, limiting retributivists appeal to what I call the "while we're here" argument: "while we're here exacting retribution," they argue, "we might as well maximize other social goods." If the range of "not undeserved" sentences is remarkably wide, they would continue, it does not alter this logic; indeed, it would merely afford us greater flexibility to pursue those other social goods. Limiting retributivists thus employ "external" reasons to resolve proportionality's vagueness. That is, they look beyond the particular covering value

50. Morris, *supra* note 2, at 179 (emphasis added).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

when searching for a reason to select one vaguely related option over another. Above, when considering the choice between a clarinet career and a lawyer career, we discussed how the lawyer career being better for one's children could be a determinative *external* reason in favor of that option. It would not make the lawyer career better *as a career*, but it could nonetheless determine the choice. For limiting retributivists, nonretributivist considerations represent our external reasons for determining proportional sentences. Whether a sentence contributes more or less to, say, general deterrence does not make it more or less retributively deserved, but it is nonetheless a legitimate *reason* for or against selecting that sentence within the array of eligible options. Or so goes the argument.

Limiting retributivism is a hugely influential theory. It is the explicit basis for the "modified just deserts" sentencing guidelines in Minnesota, as well as the guidelines in Washington, Oregon, Kansas, and North Carolina, which were based on Minnesota's system.[51] It has been endorsed by the American Law Institute and its Model Penal Code on Sentencing, which was updated in 2017, and the American Bar Association.[52] Loose versions of the theory are implicit in all the other American guidelines systems, as well as in the sentencing regimes of most Western nations.[53] Frase writes that it is the "de facto consensus theoretical model of criminal punishment."[54] In this section, however, I will argue that the limiting retributivist attempt to move *up* the scale of vaguely deserved sentences violates an offender's inviolability. To inflict a sentence more harmful than the *least harmful* "not undeserved" sentence, for the purpose of realizing nonretributivist ends, merely sacrifices the offender for those ends. The external reasons that limiting retributivists attempt to bring into the frame to forcibly resolve the vagueness of retributive proportionality are thus illegitimate. Retributivism, I conclude, can only ground an exceptionally mild system of punishment. To justify the infliction of *more* penal harm than this, we would need to appeal to a nonretributivist justification of punishment, one that itself does not merely sacrifice offenders for the greater good.

51. FRASE, *supra* note 3, at 3–4.

52. MODEL PENAL CODE: SENTENCING §1.02(2) (2017) ("(2) The general purposes of the provisions on sentencing, applicable to all official actors in the sentencing system, are: (a) in decisions affecting the sentencing of individual offenders: (i) to render sentences in all cases within a range of severity proportionate to the gravity of offenses, the harms done to crime victims, and the blameworthiness of offenders; (ii) when reasonably feasible, to achieve offender rehabilitation, general deterrence, incapacitation of dangerous offenders, restitution to crime victims, preservation of families, and reintegration of offenders into the law-abiding community, provided these goals are pursued within the boundaries of proportionality in subsection (a)(i). . ."); AMERICAN BAR ASSOCIATION, ABA STANDARDS FOR CRIMINAL JUSTICE: SENTENCING (3d ed., 1994); AMERICAN BAR ASSOCIATION JUSTICE KENNEDY COMMISSION, REPORTS WITH RECOMMENDATIONS TO THE ABA HOUSE OF DELEGATES (2004).

53. FRASE, *supra* note 3, at 4.

54. *Id.*

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

A. Retributivism and Inviolability

Pretend that retributive desert is our singular sentencing aim. We have no concern for deterrence, rehabilitation, or whatever it may be. Only retributive desert. We are then presented with an offender who, say, stood watch while his accomplice broke into someone's house and stole $5,000 worth of goods. How much punishment does he retributively deserve for wronging the homeowner in this way? Let's say that after considering the meaning of "injuriousness," surveying contemporary and historic retributive sentencing scales worldwide, and engaging in reflective equilibrium, we decide that the range of "not undeserved" sentences for this offense is between 1 and 100 months of imprisonment.[55] One could reasonably argue for a range even wider than this, I think, but the 1–100 scale provides some analytical clarity. How do we select a punishment within this range, given that our only positive sentencing aim is retribution? The answer is clear. We ought to sentence the offender to the *least harmful* "not undeserved" sentence: 1 month of imprisonment. To sentence him any more severely would just be cruel. That is, it would inflict *wanton* harm. It would harm him to a greater degree without in any way increasing the realization of our penal aim.

Let us now return to the real-world scenario where we have nonretributive sentencing aims, as well. Let us assume that sentencing offenders to the least harmful "not undeserved" sentence is not the optimal choice for realizing these other aims. Let us assume, more particularly, that sentencing offenders in this manner will lead to a dramatic increase in crime. Limiting retributivists would then dutifully argue that we ought to increase sentences to the optimal point for the purpose of generating deterrence, so long as we stay within the retributivist range—say, 50 months for offenses as serious as the one just mentioned. In this way, where the zone of discretion is exceedingly wide, the retributivism within limiting retributivism fails to provide much of any sentencing guidance, with all the work being done by the "external" nonretributivist considerations. The notion, then, that we are punishing an offender for the purpose of giving him his just deserts is

55. Consider the contrast between the U.S. and Swedish sentencing guidelines on burglary. Depending on factors such as criminal history, the value of the property taken, damaged, or destroyed, the degree of planning, and whether the burglar possessed a dangerous weapon, the U.S. Sentencing Guidelines' recommended range for "residential burglary" is 24 to 210 months (17.5 years). U.S. Sentencing Guidelines Manual §2B2.1 (U.S. Sentencing Comm'n 2016), at 118–120. In Sweden, by comparison, "theft" has a maximum sentence of two years. Brottsbalken [BrB] (Penal Code) 8:1 (Swed.), translation available at https://www.government.se/contentassets/5315d27076c942019828d6c36521696e/swedish-penal-code.pdf. The range is different, however, for "gross theft." "In assessing whether the crime is gross, special consideration shall be given to whether the unlawful appropriation took place after intrusion into a dwelling, whether it concerned the appropriation of property borne by a person, whether the accused was equipped with a weapon, explosive or similar aid, or whether the act was otherwise of an especially dangerous or ruthless nature, concerned property of considerable value or entailed a keenly felt loss." *Id.* at 8:4. If the offense is deemed "gross theft," the sentencing range is six months to six years. *Id.*

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

almost completely diluted. Nonetheless, limiting retributivists would insist that the offender has no complaint if we increase his sentence, since we still do not punish him more than he deserves—and we might as well, because the rest of us will be much better off.

It strikes me that the offender could reasonably object that he is being sacrificed for the greater good in this instance. More particularly, he could reasonably object that he is being sacrificed to mitigate a social problem—future crime—for which he has no personal responsibility. When the state is concerned, and only concerned, to give him his just deserts he receives a one-month sentence. The only reason for inflicting a greater degree of penal harm is that it would help the rest of us to mitigate a social problem. But that is a good reason for harming someone more than we would otherwise be entitled to only if he is to some degree responsible for that problem. If he is *not* responsible, then by harming him further we fail to respect his inviolability. The principle of human inviolability just is the idea that it is wrong to harm people for the purpose of mitigating social harms or threats for which they lack responsibility (or at least the principle can be formulated as such).[56] And if there are people in society who require the threat of nonmild punishments to desist from crime, the offender could rightly wonder why they are *his* problem.[57] For without their presence, he would receive much less prison time, perhaps years less. The mere fact that he committed a wrong is not a sufficient explanation, given that, if we looked only to that wrong, he would receive a much more lenient sentence.

This argument has dramatic implications for retributivism. When determining how much penal harm to inflict upon an offender, the retributivist ought to ask herself what is the *least harmful* punishment that is "not undeserved" in terms of its degree of injuriousness. Given the remarkable vagueness of injuriousness, this principle will generate an extremely mild sentencing regime—mild by comparison to *any* contemporary system, let alone the harsh American and British systems.[58] This is a prima facie attractive conclusion in this age of penal severity and mass incarceration.

56. *See* ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA (1974), at 28–33.

57. On the question of whether harming an offender as a means of deterring would-be future offenders merely sacrifices him for the greater good, *see* Jacob Bronsther, *Two Theories of Deterrent Punishment*, 53 TULSA L. REV. 461, 462 (2018) ("[D]eterrent punishment seems to violate the non-sacrifice principle rather straightforwardly, as the state inflicts suffering upon an offender as a prudential warning to would-be future offenders, for whom the offender has no responsibility. I call this the 'Means Problem.' Why, if at all, is the state entitled to use offenders as a means of bringing about general deterrence?"); VICTOR TADROS, THE ENDS OF HARM: THE MORAL FOUNDATIONS OF CRIMINAL LAW (2011), at 113–114, 181–182, 266–267.

58. *See* Nick Cowen & Nigel Williams, *Comparisons of Crime in OECD Countries*, CITIVAS: INSTITUTE FOR THE STUDY OF CIVIL SOCIETY (April 2012), http://www.civitas.org.uk/content/files/crime_stats_oecdjan2012.pdf (comparing crime and punishment rates within OECD states).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

B. Retributivism and Deterrence

But the question nonetheless remains. Will this be *enough* punishment for the sake of our nonretributive penal aims, most importantly for the sake of crime deterrence? Crime deterrence is, of course, not a luxury item. It is not simply about reducing the aggregate amount of harm; it is about maintaining the social trust that grounds a well-functioning and cooperative civil society, upon which, in turn, the possibility of human flourishing relies.[59] While we cannot examine this issue in depth, there are at least three reasons to think that the least harmful "not undeserved" sentences *would* be enough punishment to reduce the objective threat of crime sufficiently. First, there is considerable evidence that the *certainty* of receiving some level of punishment is more important for the purpose of deterrence than the *severity* of the punishment received.[60] Second, the money saved from lengthy sentences could be spent on nonpenal crime reduction policies, such as police presence (which increases the certainty of punishment)[61] and investments in the community like early childhood development programs.[62] Third, the threat of punishment is not the only reason that people desist from crime and cooperate; there are also, for instance, noncriminal legal institutions, like tort and contract law, as well as nonlegal social norms. Whether an extremely mild system of punishment would indeed be sufficient for the purpose of deterrence is a difficult empirical question.[63]

59. On the notion that the function of the criminal law is the maintenance of a cooperative civil society, *see* Lindsay Farmer, Making the Modern Criminal Law: Criminalization and Civil Order (2016), at 37–60; Nicola Lacey, State Punishment: Political Principles and Community Values (1988), at 160–201; Thomas Hobbes, Leviathan (Richard Tuck ed., Cambridge 1996) (1651), at 86–90; Alice Ristroph, *Hobbes on "Diffidence" and the Criminal Law*, *in* Foundational Texts in Modern Criminal Law 23 (Markus D. Dubber ed., 2014); John Finnis, Natural Law and Natural Rights (2d ed. 2011), at 261–262; Neil MacCormick, Institutions of Law: An Essay in Legal Theory (2007), at 293; Hyman Gross, A Theory of Criminal Justice (1979), at 10.

60. *See* Daniel S. Nagin, *Deterrence in the Twenty-First Century: A Review of the Evidence*, *in* Crime and Justice in America: 1975–2025 199 (Michael Tonry ed., 2013); Andrew von Hirsch et al., Criminal Deterrence and Sentence Severity: An Analysis of Recent Research (1999), at 27, 36, 45; Steven N. Dulauf & Daniel S. Nagin, *Imprisonment and Crime: Can Both be Reduced?*, 10 Criminology & Pub. Pol'y 13 (2011).

61. Dulauf & Nagin, *supra* note 60.

62. There is significant evidence that early childhood development programs are effective in reducing crime. *See* James Heckman, Rodrigo Pinto & Peter Savelyev, *Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes*, 103 Am. Econ. Rev. 2052 (2013); Alex R. Piquero et al., *Effects of Early Family/Parent Training Programs on Antisocial Behavior and Delinquency*, 5 J. Experimental Criminology 83 (2009).

63. In this way, there is uncertainty with regard to deterrence proportionality (i.e., how much ought to we punish if we want X degrees of deterrence?), just as there is with regard to retributive proportionality. There are two important differences, though. First, our uncertainty with regard to deterrence proportionality is epistemic, while our uncertainty with regard to retributive proportionality is conceptual. There is an empirical answer to the deterrence question—like there was when we compared Sticks 1 and 2—and it seems that we are indeed making (slow) progress on it. But there are deep and *permanent* conceptual barriers to making progress on retributive proportionality. Second, given what criminologists now know about deterrence, the uncertainty of deterrence proportionality seems less extreme than that of retributive

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

Importantly, the answer may vary from society to society, given all the variables that impact crime rates.[64]

Let us assume, though—not outrageously—that such a mild sentencing regime will be *insufficient* for the purpose of deterring crime.[65] In that case, even if we accepted all of the retributivist claims about who deserves punishment and why, retributivism would fail as the foundation of state punishment; or at least it would fail as the sole foundation of state punishment. For it could justify the infliction of punishment, but not the infliction of *enough* punishment for the criminal law to realize its nonretributive function of maintaining a cooperative civil society.

What then for the retributivist? To inflict more penal harm than the least harmful "not undeserved" sentence, she would need to secure a *nonretributivist* theory of punishment, and one that did not itself violate human inviolability. It could not be an unadulterated utilitarianism, of course, given our refusal to merely sacrifice people to mitigate social problems for which they are not responsible. Utilitarians have a famously hard time explaining what is wrong with punishing innocent people, exactly, if doing so is an efficient means of decreasing crime, increasing happiness, and so forth. How, then, could we justify harming an offender for the purpose of deterrence if we can appeal to neither (a) the limiting retributivist

proportionality. If we polled criminologists as to what the sentencing guidelines should be if we were only concerned with deterrence—and assuming that we wanted to do this in a parsimonious way that accounted for the harm inflicted on offenders—the average range of sentences proposed for a given offense would, I think, be considerably narrower than the retributivist range for that offense. I am grateful to an anonymous reviewer for pressing me to consider this point.

64. *See generally* Nicola Lacey, David Soskice & David Hope, *Understanding the Determinants of Penal Policy: Crime, Culture, and Comparative Political Economy*, 1 ANN. REV. CRIMINOLOGY 195 (2018) (analyzing four paradigmatic determinants of penal policy—crime rates, cultural dynamics, economic structures and interests, and institutional differences—and considering the impact of race as an independent determinant of U.S. penal policies); NICOLA LACEY, THE PRISONERS' DILEMMA: POLITICAL ECONOMY AND PUNISHMENT IN CONTEMPORARY DEMOCRACIES (2008) (examining political economic, institutional, and cultural determinants of penal severity); John Pratt, *Scandinavian Exceptionalism in an Era of Penal Excess, Part I: The Nature and Roots of Scandinavian Exceptionalism*, 48 BRIT. J. CRIMINOLOGY 119 (2008) (arguing that high levels of social trust and solidarity have grounded Scandinavian criminal justice systems and considering demographic and economic factors conducive to those high levels); John Pratt, *Scandinavian Exceptionalism in an Era of Penal Excess, Part II: Does Scandinavian Exceptionalism Have a Future?*, 48 BRIT. J. CRIMINOLOGY 275 (2008) (same); JAMES Q. WHITMAN, HARSH JUSTICE: CRIMINAL PUNISHMENT AND THE WIDENING DIVIDE BETWEEN AMERICA AND EUROPE (2003) (arguing that cultural and ideological differences explain the difference between the American penal regime, on the one hand, and French and German regimes, on the other); Nicola Lacey & David Soskice, *Crime, Punishment and Segregation in the United States: The Paradox of Local Democracy*, 4 PUNISHMENT & SOC'Y 454 (2015) (arguing that local government autonomy in the United States, and the resulting fact that criminal justice policies are filtered through local electoral politics, presents unique challenges for garnering political support for integrative criminal justice policies).

65. *See* Steven Sverdlik, *Desert as a Limiting Condition*, 12 CRIM. L. & PHIL 209, 219–220 (2018) (arguing that, if the probability of punishment is low, *maximally* harmful retributive punishments may be insufficient for the purpose of deterrence).

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

argument that the punishment is not undeserved nor (b) the plain fact that doing so maximizes social utility?

In closing this section, I will sketch two examples of what such an argument would look like. The basic intellectual challenge is to explain how it is the offender's personal responsibility to decrease the threat of future crime, such that harming him for that purpose does not merely sacrifice him for the greater good. First is Victor Tadros's "duty" theory of punishment.[66] Tadros argues that the offender has a duty to rectify the wrong he has committed against his individual victim. The offender fulfills this duty, Tadros continues, by protecting his individual victim from a *future* wrong. The offender can best achieve this, in turn, by allowing himself to be harmed by the state for the sake of general deterrence, thereby contributing to an umbrella of protection that will shield his individual victim from a future wrong. Second are the "corrective justice" and "social defense" theories of punishment that I have developed in recent work.[67] I argue that the criminal law is a system of protections, upon which all citizens rely for their assured liberty, and that an offender weakens this system by contributing to "criminality," not merely the perceived but the objective threat of crime in society. I conclude that the state is entitled to harm the offender for the sake of general deterrence as a means of rectifying his criminality contributions; over time, ideally, it would be as if he had never contributed to criminality at all, in terms of the average threat of crime faced by society. Assuming that extremely mild retributivist sentences were insufficient for the purpose of deterrence, then even the most ardent retributivist would have to bolt some such theory onto their system, with the result being that the offender both (a) deserves to suffer or to be censured, irrespective of its consequentialist impact and—*potentially more importantly when it comes to justifying his punishment*—(b) has a personal duty to contribute to crime deterrence.

CONCLUSION

In this essay, I argued that the small improvement problem is caused by vague covering values, not by incomparability or parity. I then considered a particularly difficult comparison: How many months in prison is retributively proportionate to a particular offense? I argued that the covering value at the center of this analysis—injuriousness—is not only vague, but remarkably vague. Furthermore, it is not the type of vagueness that, when considering the meaning of the covering value, we can resolve with substantive argument. The result is an extremely wide array of "not undeserved" sentences. I then argued that "limiting" retributivism fails in its attempt to apply "external" reasons to resolve this vagueness. Looking only to

66. Tadros, *supra* note 57.
67. Bronsther, *supra* note 57.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press

retributivist considerations, I argued, to inflict any punishment other than the *least harmful* "not undeserved" sentence is cruel and unjustified. If we wanted to inflict a sentence more harmful than this for the sake of a nonretributivist aim, then we would need a nonretributivist explanation as to why the offender has a personal duty to contribute to that aim. Otherwise, to inflict such a sentence would merely sacrifice him to the greater good.

https://doi.org/10.1017/S1352325219000016 Published online by Cambridge University Press