LEIF WENAR* ⓘ The Value of Unity

I. INTRODUCTION

Two debates over value are nearly coeval with philosophy itself. One debate is over what is good for its own sake (intrinsically good), the other is over what contributes to an individual's welfare ("what would make this person's life go, for him, as well as possible").[1] These two debates, over "the good" and "the good-for," are distinct, yet they have been parallel both in their leading theories and their main objections. Historically, the leading theories in both debates have been hedonist, desire-satisfaction, objective list, and perfectionist theories.[2]

Hedonist theories claim that what is valuable—either intrinsically or for a person—tracks only experiential quality. One challenge for hedonist theories is that malicious or otherwise *anti-social pleasures* may seem not to be valuable in either sense.[3] Another challenge is to distinguish *higher and lower pleasures*: to explain how some activities are more valuable even if others are enjoyed just as much.[4] Hedonist theories also fail to register non-experiential

1. Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), 493.

2. For example, in Plato's *Republic* (505b-c), Socrates rejected a hedonistic account of intrinsic value because there are bad pleasures. In the *Gorgias* (491e–492c, 494a-e), Socrates objected to a desire-satisfaction account of welfare that there are worthless desires. Later, Critolaus defended an objective list of intrinsic goods against Stoic objections, but was then criticized for listing the wrong elements. Aristotle offered the classic perfectionist theory of human flourishing.

3. Fred Feldman, *Pleasure and the Good Life* (Oxford: Oxford University Press, 2004), 38–39.

4. Jonathan Riley, "Different Kinds of Pleasure," in *Mill's A System of Logic: Critical Appraisals*, ed. Antis Loizides (New York: Routledge, 2014), 170–91.

facts that seem evaluatively relevant, such as that a person has *false friends*.[5]
Similarly, hedonist theories suggest that the best life can be had by entering
Nozick's *experience machine*, which many find implausible.[6]

Desire-satisfaction theories are challenged by what Chris Heathwood
calls "defective desires."[7] For example, and parallel to hedonism, some
people have *anti-social desires*, such as to see cruelty, or to be cruel. More,
and again parallel to hedonism, some people have what might seem *trivial or worthless desires*, such as to count blades of grass.[8] Other people
have *masochistic* or *imprudent* or *adaptive preferences* that are self-destructive or self-negating.[9] Moreover, some desires incorporate *mistaken beliefs*, which appear to vitiate the value of their satisfaction.[10]

Objective list theories *lack explanatory power*. As David Brink says, a
mere list of purportedly valuable states "begins to look like a disorganized
heap of goods."[11] Ben Bradley objects more sharply: a view that "does not
tell us why those things are on the list or how to weight them. . . is not
theorizing, but a refusal to theorize."[12] There seems little that proponents
can say when their lists diverge; for example, Martha Nussbaum lists sexual satisfaction as a central functioning, while G.E. Moore describes sexual
pleasure as "evil in itself."[13] And even where lists have elements in common, such as "success" and "love," they may seem only to be enumerating judgments shared by those making the lists.[14]

5. Shelly Kagan, *Normative Ethics* (Boulder: Westview Press, 1998), 34–36.

6. Eden Lin, "How to Use the Experience Machine," *Utilitas* 28, no. 3 (2016): 314–32.

7. Chris Heathwood, "The Problem of Defective Desires," *Australasian Journal of Philosophy* 83, no. 4 (2005): 487–504.

8. Donald Bruckner, "Quirky Desires and Well-Being," *Journal of Ethics & Social Philosophy* 10, no. 2 (2016): 1–32.

9. Rosa Terlazzo, "Must Adaptive Preferences be Prudentially Bad for Us?" *Journal of the American Philosophical Association* 3, no. 4 (2017): 412–29.

10. Chris Heathwood, "Desire-Fulfilment Theory," in *The Routledge Handbook of Philosophy of Well-Being*, ed. Guy Fletcher (New York: Routledge, 2016), 135–47, esp. 138–40.

11. David Brink, "Some Forms and Limits of Consequentialism," in *The Oxford Handbook of Ethical Theory*, ed. David Copp (Oxford: Oxford University Press, 2007), 380–423, 391.

12. Ben Bradley, *Well-Being and Death* (Oxford: Oxford University Press, 2009), 16–17. But see Eden Lin, "Enumeration and Explanation in Theories of Welfare," *Analysis* 77, no. 1 (2017): 65–73.

13. Martha Nussbaum, *Women and Human Development* (Cambridge: Cambridge University Press, 2000), 78; G.E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, [1903] 1993), 258.

14. Justin Ivory and Valerie Tiberius, "Philosophy and Well-Being," in *The Oxford Handbook of the Positive Humanities*, ed. Louis Tay and James Pawelski (Oxford: Oxford University Press, 2021), 381-83.

Perfectionist theories attempt to explain the goods on an objective list as activities that develop or exercise characteristic human capacities.[15] Yet the development or exercise of some such capacities (like gratuitously inflicting pain) are *anti-social*.[16] Perfectionism also strains to capture the attractive idea that *pleasure is good*, as feeling pleasure is not a capacity that can be developed or exercised.[17] Even more challenging is the idea that *pain is bad*, as perfectionism seems stronger as an account of value than of disvalue.[18] Perfectionism may also violate a *resonance constraint*, when it classifies as valuable activities that are not "compelling or attractive" to those engaged in them.[19]

These debates over "the good" and "the good-for" have continued inconclusively for centuries.[20] Each theory has well-known strengths, yet apparently incurable defects. I believe that value theory is at this impasse because we have inherited too few theoretical options at the object level. The familiar hedonist, desire-satisfaction, and perfectionist theories are too simplistic to capture many of our confident evaluative judgments, while no objective list theory can justify the elements it includes and excludes. Because these object-level theories are inadequate, metaethical debates that refer to them (e.g., between objective versus subjective theory, or about the relationships between value and reasons) can inherit their weaknesses.

This article outlines a new object-level theory of intrinsic value by setting out a formal model that generates lists of ultimate goods and bads.[21] This formal model is extensionally more adequate than familiar hedonist, desire-satisfaction, and perfectionist theories, while also being simpler

---

15. For example, Thomas Hurka, *Perfectionism* (New York: Oxford University Press, 1993) offers a theory of intrinsic value; Richard Kraut, *What is Good and Why: The Ethics of Well-Being* (Cambridge, MA: Harvard University Press, 2009) offers a theory of well-being.

16. Dale Dorsey, "Three Arguments for Perfectionism," *Nous* 44, no. 1 (2010): 59–79.

17. See Matthew Strohl, "Pleasure as Perfection: *Nicomachean Ethics* 10.4-5," *Oxford Studies in Ancient Philosophy*, 41 (2011): 257–87.

18. Gwen Bradford, "Perfectionism," in *The Routledge Handbook of Philosophy of Well-Being*, ed. Guy Fletcher (London: Routledge, 2015), 124–34, esp. 130–31.

19. Peter Railton, "Facts and Values," *Philosophical Topics* 14, no. 2 (1986): 9.

20. Epicurus, for example, held that only pleasure is good in itself, Spinoza once identified the good with what satisfies a desire, Moore defended an objective list of intrinsic values, and T.H. Green advanced a perfectionist theory of human good.

21. I will say no more about individual welfare, except to note that the better we understand "the good," the more perspectives on "the good-for" may open.

and more fruitful than an objective list.[22] Most of the article will be devoted to defining this formal model precisely and to showing that it appears to capture central features of how we reason about value. I will also interpret this model in ways that suggest that the concept that structures our reasoning about value is unity. Ultimately, I will propose that value acts as if all wills are one, meeting the world together.[23]

I realize that this sounds obscure—and to the extent that it is clear, implausible. Yet we appear to employ specific logics of "unity" quite naturally even in complex cases. For example, even without knowing the formalism to come, one can see that caring for children (a unity) is good, while abusing children (a disunity) is bad. Now consider how easily we understand complex statements of unity and disunity, such as

I help them fight child abuse.

Here, four parties are in relations of dis/unity. The speaker is in unity with the fighters, who are disunified with the abusers, who are disunified with the children. Despite this complexity, we assess the speaker's action easily—what the speaker is doing is good. Moreover, when we flip any unity term to form another proposition, our assessment of the speaker's action instantly flips:

I help them commit child abuse.

I stop them committing child abuse, etc.

The model to come will show that the logics that explain such complex evaluations have a formalization that is compact and cognitively ergonomic. More, this new approach will overcome many of the challenges to the four traditional theories: explaining the badness of anti-sociality, the distinction between higher and lower pleasures, when to enter the

22. For a discussion of simplicity, fruitfulness, and extensional fit as criteria for theory-selection, see Leif Wenar, "The Analysis of Rights," in *The Legacy of H. L. A. Hart*, ed. Matthew Kramer, Claire Grant, Ben Colburn, and Antony Hatzistavrou (Oxford: Oxford University Press, 2008), 251–73.

23. By saying that value "acts as if" all wills are one, I mean that the logic of the formal model is what it would be were all wills one (in the specific ways to be detailed). The model itself carries no controversial metaphysical commitments. Also note that organic unity plays no role in this theory.

experience machine, and so on. Specific logics of unity appear to structure a core of human evaluative reasoning, or so I will venture.

I.A. Scope and Plan of the Article

Physical pleasure and acts of kindness are good in themselves, we believe, even when they lead to what is bad. Similarly, physical pain and acts of cruelty are bad in themselves, even if they lead to a greater balance of good. Why is this so?

As we return to the ancient question of "the good," we will need some modesty about what we can achieve. First, the main ambition of this article is only to describe a new object-level theory of intrinsic value. Indeed, as is typical with the four traditional theories, the ambition is even less. The ambition is only to describe an object-level theory of *basic* intrinsic value, to which any (non-basic) intrinsic values will stand in some valid relation.[24] To illustrate, if pleasure is a basic intrinsic value then "maximum pleasure" might be an intrinsic value that stands in a *mereological* relation to it.[25] Or again, Thomas Hurka claims that knowledge is a basic intrinsic value and that "loving knowledge" is an intrinsic value that stands in an *intentional* relation to it.[26] Other such relations have also been proposed.[27] "X is intrinsically valuable" occurs often in the literature.[28] In this article we only say that for any assertion of intrinsic value to be correct, "X" would need to refer to, or stand in some valid relation to, the basic intrinsic values described here.

24. A basic intrinsic value is such that its intrinsic value can be known without knowing the intrinsic value of anything else. Erik Carlson, "The Intrinsic Value of Non-Basic States of Affairs," *Philosophical Studies* 85, no. 1 (1997): 95–107; see also Fred Feldman, "Basic Intrinsic Value," *Philosophical Studies* 99, no. 3 (2000): 319–46. It is important for this non/basic distinction that "intrinsic value" in this article means final value ("for its own sake"), instead of value that supervenes on intrinsic properties; see Garrett Cullity, *Concern, Respect, and Cooperation* (Oxford: Oxford University Press, 2018), 99 n23.

25. Michael J. Zimmerman, *The Nature of Intrinsic Value* (Lanham, MD: Rowman & Littlefield, 2001), 131–88.

26. Thomas Hurka, *Virtue, Vice, and Value* (Oxford: Oxford University Press, 2001), 11.

27. Toni Rønnow-Rasmussen and Michael J. Zimmerman, "Introduction" to their edited volume *Recent Work on Intrinsic Value* (Dordrecht: Springer, 2005), xxix–xxxiv.

28. A survey of the literature finds, for example, that autonomy, democracy, equality, beauty, religion, consciousness, education, biodiversity, cultural forms, original artifacts, deserved punishments, public goods, and everything that exists have been called intrinsically valuable. From this point onward in the article, "intrinsic value" in the body of the text will mean "basic intrinsic value."

For the second boundary on this article's scope, our study here is value theory, not moral theory. We will draw no conclusions about what is right and wrong, praiseworthy and blameworthy, or what anyone should do. Third, we also set aside political theory. Though the intrinsic disvalue of certain relations (like domination) will be discussed, there is nothing here about what institutions there should be or how scarce resources should be distributed.

Fourth, the value theory here is not an ethical theory—a theory of the good person. An ethical theory will explain what character traits a person should and should not have (the virtues and vices). For example, an ethical theory will detail when a person is being attentive or callous: that is, when a person is being appropriately sensitive to the ends of others. Since the value theory here will not speak to what character traits a person should have, it will not aim to explain attentiveness or callousness.[29] The theory here will explain the intrinsic value of actions like *helping* someone, which an ethical theory might use to explain the virtue of *helpfulness*. The theory here will also explain the intrinsic value of actions like *taking care of someone*, which an ethical theory might use to help explain the value of *caring about* someone that, as Stephen Darwall says, "involves a whole complex of emotions, sensitivities, and dispositions."[30] Yet the value theory here can only hope to inform such an ethical theory, not to complete it.

The new theory of intrinsic value will be a species of desire-satisfaction theory. As in all such theories, good and bad will be functions of desire-satisfaction and there will be no desire-independent values. Desire-satisfaction theories of value have a distinguished historical pedigree and remain popular in philosophy and economics.[31] Part of their attraction has

29. To get a sense of the complexity of ethical theory here, consider a normal person on a normal day and how aware she is of the ends of people in her physical proximity. A surgeon, for example, might be minutely sensitive to the preferences of her children before leaving for work, then not register the suffering in the homeless camp that she drives past every day, and then be quite differentially attentive to the desires of patients and their families when she arrives at the hospital. And this does not yet reach to the surgeon's awareness of the ends of more distant people, in her city and country, not to mention in the wider world. Which ends of others a person should be aware of—as well as how much, and when—are important discussions within ethical theory, to which an account of intrinsic value can at most contribute.

30. Stephen Darwall, *Welfare and Rational Care* (Princeton: Princeton University Press, 2002), 2. Ethical theory may show that attentiveness, callousness, helpfulness, caring about and so on are (non-basic) intrinsic values or disvalues, which stand in a valid relation to the basic intrinsic values discussed in this article.

31. Heathwood, "Desire-Fulfilment Theory," 135–38.

been that, especially in matters of taste and fancy, getting what one wants often seems to be good. Desire-satisfaction theories also obey the resonance constraint, by classifying as valuable only activities that are compelling to those engaged in them.[32]

Yet traditional desire-satisfaction theories fail extensionally because they attempt to explain value with only a single formal function: good is just "whatever satisfies (informed, etc.) desires." This simplistic function leaves traditional desire-satisfaction theory vulnerable to extensional objections, especially concerning anti-social desires.

A single function cannot capture the phenomena of value, because the value of a satisfied desire depends on its object—particularly when its object is the satisfaction of another desire. What will distinguish the new desire-satisfaction theory from traditional versions is how it characterizes value as emerging from the logical relations that one's desires can have not only to the world, but also to others' desires and to one's own desires.

The formal model to come will set out the logics of desire-satisfaction along three distinct dimensions of value: value in our relations to the world, value in our relations to each other, and value in our relations to ourselves. We might call these the "extrapersonal," the "interpersonal," and the "intrapersonal" dimensions of value.

In addition to this three-dimensional structure, the formal model will be built with further definitions and axioms, some familiar and some novel. Early on, there will be a distinction between positive and negative desires. The middle of the article will draw out the implications of two axioms: *telicism*, familiar from Parfit, and *telic nesting*, which is new. As with any new model, some elements may initially seem surprising; yet, as the bulk of the article shows, the model fits the "data" of our confident evaluations, while remaining relatively simple.

Throughout, this new theory of intrinsic value will be purely formal. What is good will turn on formal relations of our desires to the world (Section II), to others' desires (Section III), and to our own desires (Section IV). By the end, these logics of unity will yield an attractive value

---

32. Some are also attracted to desire-satisfaction theories for their potential to provide a general, naturalist, reductive theory of normative reasons; see David Enoch, "On Mark Schroeder's Hypotheticalism: A Critical Notice of *Slaves of the Passions*," *Philosophical Review* 120, no. 3 (2011): 423–24. A detailed assessment of the attractions of desire-satisfaction theories is David Sobel, *From Valuing to Value: A Defense of Subjectivism* (Oxford: Oxford University Press, 2016).

pluralism. The theory will affirm that many lifestyles and cultural practices are good in themselves, while confirming that bullying and racist domination are intrinsically bad.

The formal model will turn on relations that define very specific senses of "unity." The success of this model naturally raises the question of why these exact relations seem to capture our evaluative reasoning. Toward the end, I respond to this question by sketching an interpretation of the formal model, the "many-one interpretation." This interpretation explains the structure of the formal model by drawing a parallel between interpersonal value and individual rationality. At the very end, I extend this interpretation by reworking an image from Plato's *Symposium*.

The formal model, and its interpretation, will raise many metaethical questions. For example, like all desire-satisfaction theories, the new theory will need to explain why our phenomenology often seems to favor a "desired because good" explanation within a Euthyphro problem.[33] There is also a larger set of questions for all such theories about the relations between value so understood and deontic concepts like "reasons" and "ought."[34] Instead of addressing those questions, I will conclude with the hope that philosophers will welcome having a new approach to intrinsic value to consider.

II. EXTRAPERSONAL VALUE

Extrapersonal desires define the first dimension of intrinsic value in the formal model. Extrapersonal desires have the world as their object, for example "that I eat the chocolate," or "that I go for a row," or "that the Reef will survive this century." (Here, "the world" is a technical term in the formal model, comprising any state of affairs that is not the satisfaction or dissatisfaction of a desire.) This extrapersonal dimension of value follows a distinctive logic of unity, which can be described quickly because it resembles traditional (single-function) desire-satisfaction theory.

We begin with Derek Parfit's notion of a "telic" desire, which is a desire for some state of affairs as an end, or for its own sake.[35] One axiom of

---

33. David Sobel, "How to be a Subjectivist," in *The Routledge Handbook of Practical Reason*, ed. Ruth Chang and Kurt Sylvan (Abingdon: Routledge, 2021), 307–17.

34. See Mark Schroeder, "Value Theory," *Stanford Encyclopedia of Philosophy* (Fall 2021 edition), ed. Edward N. Zalta, https://plato.stanford.edu/entries/value-theory/.

35. Derek Parfit, *On What Matters*, vol. 1 (Oxford: Oxford University Press, 2011), 58, broadening Parfit's "event" to "state of affairs." The end of a telic desire is a final end, which is wanted even if nothing else will follow from it.

unity theory is that the satisfaction of a (positive) telic extrapersonal desire has intrinsic value. Take the example of physical sensations. We define a "pleasing" sensation as one that is the object of a telic desire: a sensation that one wants to have for its own sake. It follows that having a sensation that is the object of such a desire is good. Having pleasing sensations is intrinsically good because one is experiencing bodily sensations that one wants to have for their own sake.

Many desires for bodily sensations are ingrained into our common animal nature. Other such desires vary widely across persons. On this account, the value of sensations tracks a person's desires: it is good that you experience the taste of chocolate only if you want to experience the taste of chocolate.[36] We will call a satisfied telic extrapersonal desire a "basic unity."

What of the badness of physical pain? Here it is useful to revive a concept of Bernard Williams, of a "negative desire": a desire that something not occur.[37] As a pleasing sensation is one that is the object of a telic positive desire, so an aversive sensation is one that is the object of a telic negative desire. We extend the axioms to say that, just as the satisfaction of a *positive* extrapersonal desire has intrinsic value, so the dissatisfaction of a *negative* extrapersonal desire has intrinsic disvalue. In a "basic disunity," one gets what one wants not to get. Aversive sensations are bodily sensations that, in themselves, one wants not to have. Thus the experience of an aversive sensation (like physical pain) is intrinsically bad.[38]

36. A more complex analysis will be needed for "intentional pleasures," such as enjoyment, gratification, and relief. Hurka, *Virtue*, 12.

37. B. A. O. Williams, "Ethical Consistency," *Proceedings of the Aristotelian Society, Supplementary Volumes* 39 (1965): 105. Negative desires here are conative states distinct from positive desires (I sometimes refer to them as "aversions"). See Shelly Kagan, "An Introduction to Ill-Being," in *Oxford Studies in Normative Ethics*, vol. 4, ed. Mark Timmons (Oxford: Oxford University Press, 2014), 261–88; Daniel Pallies, "Attraction, Aversion, and Asymmetrical Desires," *Ethics* 132, no. 3 (2022): 598–620; Chris Heathwood, "Ill-Being for Desire Satisfactionists," *Midwest Studies in Philosophy* 46 (2022): 33–54. Throughout, I occasionally use "desire" to cover both positive and negative desires, and "satisfaction" to cover both satisfaction and dissatisfaction.

38. The account of the value of sensations here is neutral on whether physical pleasure and pain are best characterized phenomenologically or conatively. Gwen Bradford, "The Badness of Pain," *Utilitas* 32, no. 2 (2020): 236–52. For brevity, I will sometimes refer to pleasing and aversive sensations as "pleasure" and "pain."

The zero point of extrapersonal value is the dissatisfaction of a positive desire (when one does not get what one wants) and the satisfaction of a negative desire (when one does not get what one wants not to get). In these cases, neither a basic unity nor disunity is formed, and (it is axiomatic that) no intrinsic value is present.

The satisfaction of a positive telic extrapersonal desire has value, the dissatisfaction of a negative telic extrapersonal desire has disvalue. In contrast to perfectionist theories, the most attractive theses of hedonism, about the intrinsic value of physical pleasure and pain, follow directly from the definitions and axioms so far.[39]

Extrapersonal desires need not have sensations as their objects: the objects of a person's world-oriented desires can be states of affairs beyond her current or possible experience. One can want to see the cherries blossom next spring, for example, or want that one's book stay in print after one's death. Any such desire may integrate some factually incorrect beliefs, which raises a classic problem: is what satisfies a desire based on false beliefs good or not?

Parfit advances a thesis that handles desires based on false beliefs about means. Parfit says that only what satisfies telic desires matters—only what is wanted for its own sake.[40] Call this thesis "telicism" for intrinsic value. Say a woman falsely believes that this train will take her to her lover. Then, by telicism, her taking this train will not help to satisfy her telic desire, and will have no value in itself.

And yet, as Parfit says, one may also have telic desires with false beliefs incorporated in their end.[41] Consider an ancient Canaanite, say, who wants to worship the god Baal. Or consider Arneson's wife, who (Arneson imagines) wants to construct a huge monument to his virtue, which she wildly overestimates.[42] Some theorists move from such examples to "informed desire" (or "ideal advisor") theories, which attempt to filter out

39. In this article, I assume that the bearer of intrinsic value is the "combo" of a desire and a state of affairs that satisfies it; when I occasionally speak of the "object" (the state of affairs) as having intrinsic value, it is in the derivative sense of its being a component of this combo. Eden Lin, "Two Kinds of Desire Theory of Well-Being," *Midwest Studies in Philosophy* 46 (2022): 55-86.

40. Parfit, *Reasons and Persons*, 117; Parfit, *On What Matters*, 59, translating the thesis from reasons to intrinsic value.

41. Parfit, *On What Matters*, 60.

42. Richard Arneson, "Human Flourishing versus Desire Satisfaction," *Social Philosophy & Policy* 16, no. 1 (1999): 124.

desires based on false beliefs.[43] Yet we might instead stay within actual desire theory, and look to the teleological nature of desire.[44]

Desiring entails having a disposition to attend to an end, and to realize (or prevent) it when one believes that one can do so.[45] The end of a desire sets its *telos*: its aim or object. The actions taken by the Canaanite or Arneson's wife cannot achieve the aims of their telic desires, because the definite descriptions within their ends ("the god Baal," "Arneson's great virtue") have no referents. So their actions cannot be valuable with respect to these aims, which is the correct result.

What of wishing? A wish is a desire whose end the agent believes she cannot act to achieve. When an agent has a wish, her underlying disposition may keep her attention directed at the end, even as her practical reasoning keeps churning out the result that she can take no action to achieve it. When a wished-for state of affairs comes about ("I wish the sun would shine"), a basic unity is created just as when an end is achieved through action. For extrapersonal desires, it is good when one gets what one wants, whether through action or by happenstance.[46]

How good? Degree of value varies with strength of desire. A stronger desire disposes an agent to attend more to things associated with its end; and a stronger desire motivates the agent more to realize or prevent the end, modulo the agent's subjective probability of attaining the end.[47] The pursuit of a weak desire is easily abandoned for other desires, while a life project may be pursued across a wide range of circumstances at cost to other ends. The stronger the desire, the more valuable is the achievement of the end.[48]

---

43. For example, Rawls, Brandt, and Railton offer informed desire theories of welfare.

44. Heathwood, "Defective Desires," defends "actualism."

45. See Mark Schroeder, *Slaves of the Passions* (Oxford: Oxford University Press, 2007), 156–58. "Realize" and "prevent" distinguish positive and negative desires.

46. An action is intrinsically good when it is part of the state of affairs that satisfies a telic desire. "I want to paint the house" is satisfied only when I paint the house, while "I wish that the house were painted," can be satisfied whether I paint the house or not. We might evaluate the character of a person who is "willing to" differently than a person who merely "wishes that," yet that is an issue within ethical theory.

47. This account of strength of desires draws on Neil Sinhababu, *Humean Nature: How Desire Explains Action, Thought, and Feeling* (Oxford: Oxford University Press, 2017), 23–26. Any account of strength of desire will need elaboration to account for akrasia, addiction, and so on.

48. Note that, as in hedonism, value here sums across persons: if some at a dance want a certain song to be played, while others are averse to this, the total value of the song being played will be the sum of the strengths of the (positive and negative) desires of each dancer.

This account of extrapersonal value captures what has attracted many to traditional desire-satisfaction theories. Here value tracks only what one finds compelling, so the account obeys the resonance constraint. Moreover, this account affirms that, in many cases, getting what one wants can be very good.[49] Of course, telicism says that what is truly valuable turns on what we actually want, which can be a deep fact unknown even to us. Yet if we give ourselves poetic license for a moment and call everything that is not desire-satisfaction "the world," we can say that, on this account, value flows from unities between will and world.

III. INTERPERSONAL VALUE

III.A Basic Analysis

Kindness is good and cruelty is bad—and why? These look to be objective, desire-independent judgments of value. I will suggest that while these judgments are indeed objective, analysis shows that they track purely formal relations among desires. There is a logic—or, if you like, a geometry—that defines the relations that desires can have to each other, and once one sees it, one sees it everywhere in everyday interpersonal evaluations. The "formal" structures our "ordinary" evaluative concepts, and many of our confident judgments of value have formal relations between desires at their core.

Interpersonal desires have the desire-satisfaction of others as their objects. The logics of interpersonal desire are more complex than the logic of extrapersonal desire, because interpersonal desires are all higher-order desires, and they have two possible structures. In the simple form of the first structure, the Tau- or T-structure, the object of one person's desire is another person's desire-satisfaction. In the simple form of the second

49. An Editor wonders about the value of the satisfaction of an unusual telic extrapersonal desire, such as one person's desire for the destruction of the Grand Canyon. The value of the satisfaction of this single destructive desire would hardly register in overall value, given the enormous number of people with strong contrary desires regarding the Canyon. Yet what if this destructive desire became universal? Imagine a future where many come to view the Canyon with horror, after learning of bloody massacres of ancient indigenous peoples there. Many others come to see the Canyon as an ugly gash in the beautiful skin of Gaia. Many others come to feel existential dread (as I once did) contemplating the eons of indifferent time embedded in its stratigraphy—perhaps after a gifted writer does for the Canyon what Sartre did for the chestnut tree in *Nausea*. In a future where everyone wants the Canyon to be filled in, would it not be better, *ceteris paribus*, to do so?

structure, the Lambda- or Λ-structure, two people share the end of acting together.[50] The T-structure models what we want to do *to* and *for* each other; the Λ-structure models what we want to do *with* each other.

We will examine the T-structure throughout most of Section III. This sub-section begins to set out its logic, the next two sub-sections test the power of this formalism to explain ordinary evaluative judgments. We then turn to multi-leveled T-structures with "telic nesting." After this, we explore telicism in more depth and make some basic contrasts between the right and the good, before ending this Section with a "unity" interpretation of the T-structure formalism and how it introduces the Λ-structure formalism.

Let us start on the T-structure with a simple two-person case. P (Patient) yearns for a certain pleasing sensation—a sensation like the warmth of the sun on his body. As we have seen, this means that P has a telic extrapersonal desire for this sensation: his desire is "that I feel *S*." Now consider the second person, A (Agent), who has a telic *interpersonal* desire: that P have the experience that he wants to have. The object of A's telic desire is P's desire-satisfaction: A's final end is "that P feels *S* (because he wants to feel it)." When P feels the sun's warmth, we say that P is in unity with the world and that A is (in a different sense) in unity with P. The satisfaction of A's desire generates its own positive (interpersonal) value.

To calculate the total value in a two-person case like this one, two separate analyses are needed: one "on the P (Patient) side" and one "on the A (Agent) side." It is good "on the P side" when P experiences the sensation that he wants to have (extrapersonal value). Leaving strengths of desires aside, let us stipulate that "P feels *S*" has value +1, and "P does not feel *S*" has neutral value 0. Then it is also separately good "on the A side" when A's desire that P experience this sensation is satisfied (interpersonal value). On the A side, "P feels *S*" has value +1, and "P does not feel *S*" has value 0. The satisfaction of A's *kindly* desire generates a separate interpersonal goodness, to be added to the extrapersonal goodness of P's pleasure. When P experiences the sensation, the total of value (extrapersonal plus interpersonal) is +2.[51]

50. The "T" represents one desire directed at another; the "Λ" represents two desires converging.

51. For simplicity, we assume here that interpersonal value is added to extrapersonal value, instead of being a multiplicative or even more complex function of it. Hurka, *Virtue*, discusses such functions extensively.

Say now that A wants P to feel an aversive sensation like pain. In this case, P wants "that I not feel S^*" and A wants this negative desire of P's to be dissatisfied. As above, P's feeling the aversive sensation is bad: there is extrapersonal disvalue in P experiencing sensations that he wants not to have. "On the P side," the dissatisfaction of P's negative desire (P feels S^*) has value -1 and the satisfaction of this desire (P does not feel S^*) has value 0. Yet A takes as her end the dissatisfaction of P's desire: A's will is disunified with P's will.[52] Here, the satisfaction of A's desire has value -1, and its dissatisfaction has value 0. The satisfaction of A's *malign* desire generates its own, interpersonal, badness. When P feels the aversive sensation, the total of value is -2.

Filling in more of the logical space fills out the formal model. Say that A desires that one of P's negative extrapersonal desires be satisfied: A (like P) wants P not to feel a particular pain. Then when A's *compassionate* desire is satisfied, the P value is 0 and the A value is $+1$.[53] Say now that A desires that one of P's positive extrapersonal desire be dissatisfied: A (unlike P) wants P not to experience a pleasure. Then when A's *spiteful* desire is satisfied, the P value is 0 and the A value is -1.[54]

This formal model has a neat compactness. For notice that when A wants a desire of P to be satisfied, the satisfaction of A's desire generates interpersonal value *whether P's desire is positive or negative*. That is, whether A wants P to feel pleasure, or A wants P not to feel pain, the satisfaction of A's desire is an interpersonal unity ($+1$). Similarly, when A wants a desire of P to be dissatisfied, the satisfaction of A's desire has interpersonal disvalue *whether P's desire is positive or negative*. Whether A wants P not to feel pleasure, or A wants P to feel pain, the satisfaction of A's desire is an interpersonal disunity (-1). This is a central feature of the model, which will be important for the interpretation of "telic nesting" below.

The analysis so far may bring to mind an axiological principle familiar from Hurka and others: the recursive principle ("loving the good is good,

52. As we will see in the interpretation of the formal model below, A's will contradicts P's will—as the will of a person with who satisfies a telic desire for her own pain contradicts itself.

53. If P experiences the pain, the P value is -1 and the A value is 0.

54. If P experiences the pleasure, the P value is $+1$ and the A value is 0. The full formal model contains more logically possible combinations of positive and negative desires than are discussed here.

loving the bad is bad," etc.).[55] While the recursive principle may parallel some of the logic of value described here, it will not be useful in unity theory. Unity analysis turns not on attitudes (like "loving") but on desire-satisfaction itself ("getting," we might say). Moreover, the recursive principle's reference to "the good" can blur the mathematical values at stake; it may not distinguish "feeling pleasure" ($+1$) as a good from "not feeling pain" (0) as a good. As we continue to set out the formal model, there will be ever more divergences between unity theory and recursive theories, which will be tracked in the footnotes.[56]

Even in the partial model that we have so far, the explanatory power of the formalism is emerging. For notice how cleanly the analysis captures the conviction that it can be bad when people get what they want. Within this model, for instance, to act cruelly is to thwart for its own sake certain strong desires of another person—say, their desires not to feel pain. The objective badness of cruelty follows from the formal relations between the desires, and indeed the analysis says that the more a person wants to act cruelly and does, the worse things get. We reach this result directly, without needing to "launder out" cruel desires with an informed desire theory, and without simply stipulating that cruelty is bad with an objective list theory. In this model, the objective badness of cruelty is captured by the same formal analysis that captures the objective value of kindness, of compassion, of spitefulness, and so on. The model shows directly when it is bad that people get what they want, using the same analysis that shows when it is good that people get what they want. This formal model is, like perfectionist theories, attempting a systematic account of the elements on an objective list. Yet the model is achieving a better extensional fit with our evaluative judgments.[57]

III.B The Formal in the Ordinary

The model so far appears to track some of the thin structure of our reasoning about interpersonal value. Moreover, notice how the theory is finding this

55. Brentano, Rashdall, Moore, Ross, and Nozick all invoke recursive principles; Hurka, *Virtue*, 23–28.

56. For example, recursive theory seems saddled with a sharp distinction between its "base-level" and recursive (intentional) accounts of value (Hurka, *Virtue*, 37–55). Unity theory is more unified, as all three of its dimensions explain value in terms of desire-satisfaction, albeit in different ways. Moreover, the interpretation of the formal model below gives a unifying explanation of the relationships between the model's three dimensions of value.

57. As above, the model outperforms perfectionism by easily capturing the value of pleasure and the disvalue of pain. The model also correctly identifies as bad the characteristic human activity of gratuitously inflicting pain, which perfectionism finds harder to explain.

logic embedded in our thick evaluative concepts as well. Acts of *kindness* and *compassion* are good in themselves, while *cruel* and *spiteful* and *malicious* acts are bad in themselves. These firm judgments of value are being explained in terms of merely formal relations between desires. The more ordinary value-laden concepts we can account for in this way—in terms of formal relations between desires—the more evidence we will have that this model tracks the underlying structure of our evaluative thought.

Sadistic torture, for instance, is bad. On a formal analysis, sadistic torture is bad because of its combination of cruelty and pain. The torturer takes as her target the victim's aversions: the torturer wants to force her victim to experience sensations that he cannot help wanting not to have.[58] The cruelty "on the A side" combines with the pain and other aversions "on the P side" to produce an action that is highly negative overall. Turning to the positive end of the value spectrum, to *take care of* someone involves preventing and alleviating his pain, both of which increase interpersonal and extrapersonal value.

Let us for the rest of this sub-section focus on cases where the satisfaction of the patient's desires will produce positive value (e.g., the patient will achieve some telic extrapersonal end). In these cases, many ordinary interpersonal value concepts can be analyzed as modalities by which an agent furthers the satisfaction of a patient's desires. For example, successful acts of *aid* and *support* help the patient to achieve his ends. *Facilitation* and *encouragement* are interpersonally good for the same reason. *Generosity* and *beneficence* provide another with what he desires, which is in these cases good.[59] All of these ordinary value concepts yield to analysis in terms of purely formal relations between desires.

We can continue to find "the formal in the ordinary" with analyses of familiar cases where an agent succeeds in frustrating a patient's desires. Start with gratuitous *deception*. Here A instills a false belief in P in order to thwart P's end-achievement. Say P is in SoHo, trying to get home to Brooklyn and walking toward the Brooklyn Bridge. P stops to ask A whether he is on the right route. Then A's willfully deceiving P will be straightforwardly bad, as A intentionally thwarts P's desire-satisfaction. Again, the badness of A's deception is "on the A side"—to be added to the

58. David Sussman, "What's Wrong with Torture?" *Philosophy & Public Affairs* 33, no. 1 (2005): 1–33.

59. An Editor notes that a beneficent act can provide another with what he *will* desire at the time he gets it; for example, one can leave an inheritance for a grandchild.

separate loss of value "on the P side" of P's getting lost in the city. It is bad in itself for A gratuitously to mislead P.

The analysis of this last case says that the badness of A's lie "on the A side" turns on P not achieving P's (good) end. And this is like the badness of other things that A might do to P in order to frustrate P's ends. For example, A might *obstruct* or *impede* or *injure* or *kill* P. Here A may be aiming to cause P pain (which we know is bad); or, as with deception, A may be aiming to frustrate P's achievement of his ends. With deception, it is the false belief; with obstruction or impediment or injury or killing, it is the physical incapacity. In all these cases, A's telic desire is to cause P not to achieve P's ends, and A's success will be bad ("on the A side") over and above the badness of P's lost end-achievement ("on the P side").

A similar analysis holds of telic coercion. In any coercion, the agent's successful threat will result in the patient satisfying at least one fewer of his desires (e.g., his desire for life only, instead of his desires for life and money). So when A coerces P for its own sake, this will be interpersonally disvaluable "on the A side," while the total calculation of the value of A's coercion will also include its impact "on the P side" in the reduction of P's desire-satisfaction. Telic coercion is *domination*, where one party subordinates the other as an end in itself. Thus domination is another ordinary concept that has an interpersonal disunity of desires at its core.

III.C The Formal in the Ordinary, Continued

The reader may be wondering how this theory handles instrumental (not telic) deception, injury, coercion, and so on. We will take up this question in sub-section III.E, once more of the model of interpersonal value is in view.

First, let us collect more benefits from the analyses of deception and coercion—further examples of "the formal in the ordinary." The concepts of *manipulation* and *subjugation* will yield to the same formal analyses as deception and coercion. *Harassment*, *bullying*, and *molestation* have the same formal structures at their core, with their conceptual peripheries tailored to specific interpersonal contexts.

On the positive side, successful *advising* is the opposite of successful deceiving: it is the transmission of beliefs that help the other achieve his goals. And *offering* is the opposite of coercing: here the agent changes the patient's choice set so that the patient can achieve a more-desired set of ends.

Maleficence and *munificence*, *malice* and *altruism*, *mauling* and *tormenting* and *healing* and *nurturing*—very many value-laden

interpersonal concepts yield to formal analysis of the kind we have seen. While each thick concept is a distinctive blend of the descriptive and the evaluative, the badness in the negative concepts always comes from one person wanting another not to get what he wants, or to get what he does not want—while the goodness in the positive concepts always comes from one person wanting another to get what he wants, or not to get what he does not want. So long as these interpersonal desires are telic, their satisfactions are intrinsically bad or intrinsically good.

Finding the formal in the ordinary so extensively is further evidence for the model. What our everyday judgments are showing us, we might say, is that interpersonal disvalue flows from disunities of desires—and value, from unities of desires.

III.D Telic Nesting in Multi-Level Structures

Part of what we are seeing is how unity theory meets the challenge of anti-sociality. Cruelty, malice, domination—all of these types of anti-sociality can be analyzed formally as disunities between desires. We can also add *intolerance* to this list. If one person wants to read a certain book, or to perform a traditional ritual, then it will be bad when a second person satisfies her telic desire to stop the first from doing what he wants to do. Moreover, the "target" desires for intolerance can also be desires for shared action (which will be modeled later with the Λ-structure). So if two people want to perform a ceremony together, or to give each other pleasures, then it will be bad when intolerant outsiders with telic desires to stop them succeed.

The formal analysis correctly explains many of the touchstone cases of intolerance in the philosophical literature: the Nazi persecutions of Jews, the puritanical majority that hates a minority or what they do, and so on.[60] The analysis also explains why, for example, the genocidal campaign by Hutus against Tutsis in Rwanda was worse than an epidemic would have been that killed the same Tutsis. The killings added a great deal of (interpersonal) badness to the badness of the deaths themselves. The formal model accounts for the intrinsic badness of anti-sociality

---

60. An Editor asks about cases where majorities suppress minorities not from intolerance but purely instrumentally—say, to prevent public disorder. Discussing such cases would lead us into political theory; here we only note that the belief that suppression is required to maintain order is often false. Where it is not, responsible political theorizing will begin by weighing the values at stake.

without simply stipulating it, which has been perhaps the greatest challenge for the four traditional theories of value.

So far, all of our interpersonal examples have involved two parties, A and P. We can extend the logic of the model to an indefinite number of parties by examining "telic nesting" in a multi-party chain of desires, using a familiar example of anti-sociality as our starting point.

Tertullian tells of the public torture of a Christian named Perpetua in a Roman amphitheater in 203 AD. Gladiators first formed a line to whip Perpetua, then set wild beasts on her, before ultimately killing her with sharp instruments.[61] During this torture, let us say that Perpetua wanted not to feel pain, that a gladiator wanted to cause Perpetua pain, that a spectator wanted the gladiator to satisfy his desire to cause Perpetua pain, and that all of these desires were telic. This example illustrates "nesting" telic desires: telic desires that are specified in terms of the dis/satisfaction of another telic desire (which may be specified in terms of the dis/satisfaction of another telic desire, etc.). The gladiator's desire is specified in terms of Perpetua's desire-satisfaction, and the spectator's desire is specified in terms of the gladiator's desire-satisfaction.

Call Perpetua's desire the extrapersonal "anchor" of this nesting desire-chain.[62] Then we can calculate separately the value of the satisfaction of each of the telically nesting interpersonal desires (the gladiator's desire and the spectator's desire). In this analysis, each interpersonal desire follows the "two party" (A and P) logic relative to the anchor desire, with any intermediate links in the chain determining whether the interpersonal desire in question is directed at the satisfaction or the dissatisfaction of the anchor desire. So the *sign* of the value of the satisfaction of such an interpersonal desire depends on whether it "affirms" or "negates" the anchor desire. The interpersonal desire "affirms" the anchor desire if there is an even number (including zero) of negations in the desire chain leading to the anchor, and its satisfaction will produce positive value. The interpersonal desire "negates" the anchor desire if there is an odd number of negations in the desire chain leading to the anchor, and its satisfaction will produce negative value. (In either case, the *magnitude* of the value produced by the satisfied interpersonal desire in question will be the magnitude of the strength of that desire.)

61. David Bomgardner, *The Story of the Roman Amphitheatre* (London: Routledge, 2000), 142–43.
62. Perpetua's desire is the anchor of the chain; it is not part of the chain.

In our example, the gladiator wants a negation: he wants Perpetua's desire to be dissatisfied. So there is an odd number of negations (one) between the gladiator's desire and the anchor. The spectator's desire is positive: he wants the gladiator's desire to be satisfied. So again there is an odd number of negations (one) between the spectator's desire and the anchor. Therefore, each of these interpersonal desires negates the anchor desire, so the satisfaction of each will produce a separate, interpersonal disvalue. Thus the analysis finds negative value at all levels of this example: the pain of Perpetua, the action of the gladiator to inflict pain, and the fulfillment of the spectator's wish for the gladiator to inflict pain. Moreover, the analysis also finds that the more such acts of cruelty are performed in the amphitheater, the worse things become. And the more spectators are there to watch the cruelty, and the more those spectators want to watch the cruelty, the worse things become. These are all the correct results.[63]

As we vet more examples, we find that this model of telic nesting appears not only to be getting the right results, it appears to be tracking the way we think. Its compact logic seems to account for the easy intuitive evaluations we make of even complex, multi-leveled propositions such as those we started with.

Consider, for example, "A helps them fight child abuse." In this example, we see how the even/odd tally cancels out pairs of negations in a telic desire chain, leaving only the fundamental affirming or negating relation between the interpersonal desire in question and the extrapersonal anchor. The abusers' desires negate the children's desires (one negation), the fighters' desires affirm the children's desires (two negations), and A's desire affirms the children's desires (two negations); thus the satisfaction of these desires will be bad, good, and good, respectively. These are also the correct results, and we reach them easily.

Table 1 shows the analyses of the telic desires that nest within "A helps them fight child abuse," separating out the distinct scenarios where there are one, two, three, and four parties.[64]

---

63. Hurka's recursive theory allows that a cruel achievement of sufficient magnitude can be not only good in itself, but can make the world better overall, even considering the victim's pain. The recursive structure of the theory also means that loving such cruelty can be good, etc. Hurka, *Virtue*, 144–52. The model here avoids these counterintuitive results.

64. The child's desire is a negative extrapersonal desire—say, not to feel pain. As before, we use dummy values for the strengths of desires.

Table 1 Nesting Telic Desire-Chains Within "A Helps Them Fight Child Abuse"

Number of parties	Highest-level desire	A (level 4) Affirming desire	Fighter (level 3) Affirming desire	Abuser (level 2) Negating desire	Child (level 1) Extrapersonal anchor	Total value
1	Satisfied				0	0
	Dissatisfied				−1	−1
2	Satisfied			−1	−1	−2
	Dissatisfied			0	0	0
3	Satisfied		+1	0	0	+1
	Dissatisfied		0	−1	−1	−2
4	Satisfied	+1	+1	0	0	+2
	Dissatisfied	0	0	−1	−1	−2

The ease with which we evaluate these increasingly complex relations is noteworthy. Moreover, as we continue to form new examples by flipping these interpersonal desires between affirming and negating, our evaluations immediately flip to track the relation between each interpersonal desire and the extrapersonal anchor. For example, in "A stops them fighting child abuse," the satisfaction of A's desire will be bad. In "A stops them fighting child care," the satisfaction of A's desire will be good.

Our ability quickly to evaluate such complex, multi-leveled propositions suggests how deeply embedded the logic of interpersonal value is in our everyday thinking. Even people with no formal training whatsoever use this logic quite expertly. Nor is it the case that we have a general proficiency for understanding such complex propositions beyond the realm of value.[65] This formal model is cognitively ergonomic, and it appears to model an impressive axiological facility in humans.

With telicism and telic nesting we have the two distinctive axioms of the formal model of interpersonal value. We next explore telicism in more depth, especially its two corollaries of neutral value in satisfied instrumental interpersonal desires. We then quickly survey some differences between the right and the good, which will lead to the interpretation of the T-structure of interpersonal desires and then the Λ-structure.

65. Compare, for example, "A sketches them drawing landscape painters."

III.E Neutral Unities

By Parfit's telicism, only the satisfaction of telic desires generates intrinsic value. Therefore, merely instrumental unities of ends are "neutral" in value. Say that A is behind P in line to be served. A might desire that P gets what P wants—that P be served quickly—but only because A herself wants to be served more quickly. Or say that A is drafting P in a bicycle race. Like P, A might desire that P pedal fast—but A only wants P to pedal fast so that the two of them can outdistance a third competitor, after which A plans to race in front of P and win. In such cases, A does have a desire that P's desire be satisfied. But A's desire for P's desire-satisfaction is merely instrumental, so its satisfaction generates no interpersonal intrinsic value.

These kinds of neutral unities of ends are very common. You want the driver to succeed in maneuvering through traffic, so that you can get to the restaurant on time. The waiter wants you to enjoy the meal, so that you will leave a bigger tip. Most market exchanges, and some casual sexual encounters, also create merely neutral unities. In such cases, "on the agent's side" the satisfaction of the patient's desire will be a (valueless) means to the achievement of the agent's final end.

III.F Neutral Disunities

We are within a purely formal, "geometric" value theory. As neutral unities have zero value, neutral disunities have zero value as well. This is because of telicism: only the satisfaction of telic desires matters for intrinsic value. Thus instrumental lying, obstructing, injuring, killing and so on will generate no disvalue on the agent's side. Even torturing—if done, say, as a means to gain information—will not be bad on the agent's side. This is parallel to "neutral unities" where one person satisfies another's desires as a means (for example, the waiter gives good service to get a better tip). In neither kind of case will one person's instrumentalization of another be either good or bad on the agent's side.[66]

66. As mentioned at the outset, the value theory here is only a propaedeutic to ethical theory. Within ethical theory we might judge that a torturer's indifference to the victim's ends shows her to be a callous person. This is an ethical judgment, about what desires the torturer should have. An Editor describes a related case that shows objectionable indifference through omission instead of commission. Here A is enjoying a book in the park when someone nearby is badly injured. A has no desire to help the injured person, but rather only wishes that his screaming would stop. Any plausible ethical theory will condemn A, because she lacks a desire that she should have, and some ethical theory may go on to demonstrate that her character is intrinsically bad. The value theory here can only inform, and not complete, such ethical judgments.

Of course, on the patient's side, all of these disruptive actions will typically have some, and sometimes overwhelming, disvalue: pain, injury, death, and so on. Moreover, patients may also have intense telic aversions to being treated in certain ways (like being tortured). The negative value on the patient's side usually makes these acts bad overall, and sometimes extremely bad. Indeed, cases where agent value is swamped by patient disvalue define a second sense of "anti-social action," which supplements the telic sense above. Such actions can be *ruthless* or *selfish* or *self-serving*. And yet, even when such acts are very bad overall, it will not be disvalue on the agent's side that makes them so.[67]

This raises the concern that there will be misalignment between "the right" and "the good." It may seem that the new value theory—because it is indifferent to instrumentalization on the agent's side—must judge as good some acts that we would intuitively regard as wrong.

There is indeed a range of such cases. What is interesting is how narrow this range is, and that it is bounded by alignment between the right and the good on both sides. The divergence between value and morality is limited, and examining it appears to confirm part of what morality is about.

67. An Editor asks about a case where A has no telic desire to harm P, yet has a telic desire (stronger than all of P's harm-related desires combined) that is tightly connected with P's suffering harm, such as to hear P's screams. Would not the satisfaction of A's desire then yield a net gain in value? The structure of the case is correct; the challenge for making it into an objection is to find an example realistic enough to test our intuitions. Setting aside the disvalue on P's side, what are we imagining A to be like?

Characters from fiction may come to mind: the murderer in Süskind's *Perfume*, Hannibal Lecter or Buffalo Bill in *The Silence of the Lambs*. In real life, the serial killers Brady and Hindley made a recording of a 10 year-old girl's pleas and cries as they tortured her to death. Yet I doubt we are here imagining people who have no telic desire to harm or control their victims. To be certain that we are imagining a person for whom harm or control are only instrumental, it seems that we need to picture a person who would be "just as happy" to achieve their final end through non-harmful, non-controlling means (such as knowingly to hear perfectly simulated screams). That is, it would have to be idiomatically correct to say of such a person that they have no desire to hurt anyone. If such a person can indeed be imagined, I suspect that our main response would be ethical revulsion: such a person would have an appalling character.

The Editor also asks about comparing a less harmful sadistic torture to a more harmful instrumental torture. The theory's response here is straightforward: the former case can be worse. This comparison offers a good opportunity to test one of this article's fixed points: that acts of cruelty are bad in themselves.

III.G The Right and the Good

Intuitively, the right and the good align substantially. Acting kindly can be the right thing to do, and a good thing to do; acting cruelly can be the wrong thing to do and a bad thing to do, and so on. The model of intrinsic value here gives the formal explication of the "good" and "bad" sides of these alignments. In cases where acting kindly, cruelly, and so on describe the agent's final end, the intuitively right act is good and the intuitively wrong act is bad.

In cases where an agent thwarts a patient's desires instrumentally, there is still extensive alignment between common-sense morality and the formal model of value. For while instrumental injury, killing and so on will have no disvalue on the agent's side, these acts will typically not only be painful and aversive for the patient, they will also cut off large branches of the patient's desire-satisfaction in the future. So whatever final end A is pursuing by means of injuring or killing P would have to be very valuable "on the A side" to outweigh the likely large negative value "on the P side." When the cousin kills the heir for the sake of the inheritance, for instance, this will be both intuitively wrong and formally bad.

Attention to the patient's side also confirms substantial alignment between the right and the good in cases of instrumental deception. When a candidate lies to the voters to get elected, this is likely both intuitively wrong and formally bad. When we raise value on the agent's side while lowering value on the patient's side, we continue to find alignment between the right and the good. A's harmless "white lie" to save face is intuitively permissible, and in unity theory the positive value of the white lie on the agent's side combines with the "no harm" on the patient's side to produce positive value overall. More, an act of harmless deception for the sake of achieving *the patient's* ends is positively valenced in both morality and value. Throwing a surprise party for someone who enjoys surprises is both permissible and good, even though it involves deceiving the surprisee.

The clearest divergences between ordinary morality and formal value are familiar cases like "killing x people to save x+y people." Here the value theory will find more patient value on the "x + y people" sides than patient disvalue on the "x people" sides. And on the agent's side, the satisfaction of the agent's telic desire to save lives is good, while her killing as a means will not register as bad: those disunities are neutral. Thus in some of these cases, unity theory does mark as good acts that morality rejects.

Even here, though, note that this divergence is bounded by alignment between ordinary morality and formal value on both sides. Common sense

and unity theory both give negative judgments on cases of killing x to save fewer than x. And nearly everyone agrees that it is permissible to kill x people to save x + y people, if y is a large enough number. (As Kagan says, moral constraints are relaxed beyond some threshold.[68]) It is only for middling values of y where our intuitive judgments of rightness do not align with positive value in unity theory. Why this is so is for moral theory to explain. What morality is about, at least in part, is the wrongness of treating others merely as a means, in a range of cases where doing so will make things better.

There is a great deal more to be said about the relations between the right and the good. Let me simply note another difference between the two areas of normativity, which is how value binds us together, while morality keeps us apart.

In detecting no disvalue in the instrumentalization of one person by another, the theory here "does not take seriously the distinction between persons."[69] From the perspective of value, your inflicting pain on me as a means to your end is (all else equal) no worse than your inflicting pain on yourself as a means to your end. Similarly, paternalistic deception is interpersonally neutral—your switching me to a better means to my end is no worse than my switching myself to a better means to my end.

Much of our evaluative reasoning, it appears, does have this structure—a structure that is familiar to those in close relationships, in which, as it is said, "all things are common." Within healthy relationships of friendship and solidarity and love, we do not take so seriously the distinction between persons. We may draw on the resources of others, and they on ours, instead of insisting rigidly on our separateness. One might think of the spirit of Cohen's camping trip.[70] Or consider very close relationships and the sincere, self-sacrificial attestations made by parents and comrades and lovers ("I would die for you").[71]

---

68. Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), 5.
69. John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 27.
70. G. A. Cohen, *Why Not Socialism?* (Princeton: Princeton University Press, 2009).
71. An Editor rightly says that this point should not be overstated. Lovers cannot be entirely constituted by their relationship: there needs to be some separation between persons in order for each lover to have a distinct person to love. Moreover, as we will see in the next Section, each of us can create value by "being our own person," through our prudence, integrity, and achievements. Still, we moderns might also pause to reflect on the degree to which we—imprisoned in our thick-walled selves and drenched in individualist ideologies—do yearn for deeper interpersonal unities.

These close relationships can go wrong, of course, and become exploitative. And outside of close relationships, people are all too willing to instrumentalize others while ignoring the negative impacts on them of doing so. Perhaps one function of morality is to prohibit instrumentalization in such situations, where people will be tempted to make things worse overall for the sake of their own desire-satisfaction. Within value theory, it is as if we are all already in the very best relations, freely taking from others and being taken from, transcending the distinction between persons. In this sense, from the perspective of the good, all are one.

III.H The Many-One Interpretation

We have completed the formal model of the T-structure of interpersonal desires and compared it briefly with commonsense morality. The main advertisement for this formal model of the T-structure is its extensional fit with our firm evaluative judgments. Yet it is natural to ask *why* this model works: why these tightly-defined "unity" relations consistently yield the correct results.

Let me now sketch an interpretation of what the model is showing us, which will make more precise the thought that "all are one." This will be no more than a sketch of an interpretation, which will touch on many complex topics, and the validity of the formal model will not depend on it. The crux of this "many-one" interpretation is that value between persons is homologous to rationality in a person. To epitomize the negative side of this interpretation, "bad" is like "mad." Seeing how this is so will lead us into the formal model of the Λ-structure of interpersonal desires, which I will set out in the next sub-section.

Extrapersonal desires have the world as their object; interpersonal desires have the desire-satisfaction of others as their object. Interpersonal desires may layer many levels up, yet all interpersonal desires, when fully specified, must be directed at an extrapersonal desire at level 1. Interpersonal desires that loop in referring to each other cannot be satisfied (think of a couple stuck in, "I want to do what you want to do"). Extrapersonal unity is the basic unity; all satisfiable desire-chains must ultimately be directed at the world.

So far there have been two axioms that have defined the logic of interpersonal value: telicism and telic nesting. What we have been finding in

telicism, with its corollaries of neutral unity and disunity, is that the value calculations of desire-satisfactions between persons have the same structure as the value calculations of desire-satisfactions within a rational agent. Just as a rational agent will satisfy her stronger extrapersonal desire at cost to her weaker extrapersonal desire, so more value will be created—counting both the A and P sides—when one person satisfies her stronger extrapersonal desire at cost to another person's weaker extrapersonal desire. As it is not irrational to limp on a sprained foot to escape with one's life, so it is not bad to step on another's foot to escape with one's life. In this many-one parallel, it is as if all of the extrapersonal desires of persons are the desires of a single person, and what matters is not what means are used to satisfy them but only the value of the ends achieved.

In itself, telicism does not take us beyond a traditional, single-function desire-satisfaction account of value. What cements the many-one parallel is telic nesting, which we saw in the examples of the gladiators and the fighters of child abuse. The logic of telic nesting shows that the badness of anti-sociality between persons has the same formal structure as the deepest form of individual practical irrationality, which is practical contradiction.

For an individual, it is not irrational to have contrary desires, such as wanting both to exercise and not to exercise. After all, it may be possible to sequence these ends so that they are satisfied at different times. It is not even irrational to have incompossible desires, indexed to the same time, such as wanting both to exercise and not to exercise right now. The existence of incompossible desires merely places limits on the value that can be achieved.

What is truly irrational is for an individual to thwart her own world-directed desires: that is, to satisfy telic desires that her telic extrapersonal desires be dissatisfied. Here the agent's will opposes itself directly, counteracting what the agent wants to do in the world or how she wants the world to be. As Harry Frankfurt says, "The essence of rationality is to be consistent; and being consistent, in action or in thought, means proceeding so as not to defeat oneself."[72] This irrationality is clearest in highly pathological cases, such as extreme forms of masochism where a person satisfies telic desires for her own pain. It is also seen in some cases of self-

---

72. Harry Frankfurt, "The Faintest Passion," *Proceedings and Addresses of the American Philosophical Association* 66, no. 3 (1992): 7.

mutilation and suicide, as well as in self-harming forms of addiction (as in Kipling's verse, "the burnt Fool's bandaged finger goes wabbling back to the Fire").[73]

This structure of contradiction between an individual's telic desires has its logical parallel between persons. Recall our earliest discussion of the two-person interpersonal case. There we saw that when A wants an extrapersonal desire of P to be dissatisfied, the satisfaction of A's desire produces interpersonal disvalue *whether P's desire is positive or negative.* Whatever telic extrapersonal desire P has, there is disvalue when A's telic desire that it be thwarted is satisfied. Here one person's will opposes another's directly: the agent pits her causation against the patient's causation, and fights against it. The two structures are identical: what is mad in one person is bad between two.

Higher levels of telic nesting then extend this parallel. The analysis of telic nesting resolves every many-person case into a two-person case, by testing whether an interpersonal desire at any level ultimately affirms or negates the patient's extrapersonal anchor desire. The satisfaction of any nesting interpersonal desire that negates the anchor desire is bad, just as if there were only two parties involved. So again, we have a many-one parallel of an opposing higher-order desire.

Another way to see this many-one parallel in a nesting telic desire chain is to notice that *there are no "patient side" calculations within the chain.* For example, in "They fight child abuse," the fighters thwart the desires of the abusers—but this does not register within interpersonal value.[74] This parallels individual rationality, as can be seen in a case where an individual successfully struggles against self-abuse. Here, the individual thwarts one of her own desires (to abuse herself)—but this does not register as irrational. In the many-person case, it is as if all of the nesting interpersonal desires are the desires of one composite individual, and their significance for value is only whether they are directed at furthering or thwarting that individual's world-directed desires.

Thus value between persons has the same structure as rationality in a person; "bad" is like "mad." This interpretation illuminates several

73. Rudyard Kipling, "The Gods of the Copybook Headings," www.kiplingsociety.co.uk/poems_copybook.htm.

74. If the fighters thwart desires of the abusers that are outside the desire chain—for example, causing them pain by breaking their legs—this does register as patient-side disvalue.

features of the formal model that might have otherwise seemed arbitrary, such as why "the world" is defined as states of affairs that are not the dis/satisfaction of a desire, and why the values for dis/satisfied higher-order desires are different than those for the equivalent extrapersonal desires. Unities between will and world are the basic unities, and the value—like the rationality—of higher-order desires turns on whether they aim to realize or prevent such basic unities.

This many-one interpretation of the formal model of the T-structure will be reinforced as we find more parallels in individual rationality, next specifically regarding intentions. This discussion will lead to the formal model of the Λ-structure, which fills the rest of the logical space that interpersonal relations can take. The two formal structures together will then capture the two highest points on the spectrum of interpersonal value: the two conative dimensions of love.

III.I The Two Formal Structures of Interpersonal Unity

To continue the many-one interpretation of the T-structure, consider your practical reasoning when you commit to a world-oriented action. Say, you decide to go for a row on the lake. You begin to form a plan based on the resources you have (a strong body, oars, money). You then divide this plan into sub-plans (get out of bed, get the oars, rent a boat) in ways that follow the norms of practical rationality, like means-ends coherence. These sub-plans then mesh to link your actions together in ways that get you into a boat on the lake, pulling the oars with both hands.[75]

Now let us ask: what would it be for *multiple* agents' wills to be unified in the same way that your will can be unified with itself? If it is useful, we might imagine you entering a Branch Line Teleporter: two identical copies of you emerge, with the rowing equipment divided between You_1 and You_2.[76] Yet this thought-experiment is not necessary; a friend will do just as well. In either case, there are two structures that unify the wills of the two agents: (1) one of the agents wants the desire of the other to be satisfied (the T-structure) and (2) both of the agents want to engage in shared

75. The Bratmanian analysis of intentions here is well-known, but nothing hangs on its details; see Michael Bratman, *Planning, Time, and Self-Governance: Essays in Practical Rationality* (New York: Oxford University Press, 2018).

76. Parfit, *Reasons*, 201.

action (the Λ-structure). Either your friend wants you to satisfy your desire to go rowing, or the two of you want to go rowing together.

Structure (1) is the T-structure. Here your friend's desire is directed toward the satisfaction of your desire to row, and as your rational agencies engage to achieve this end, you both begin to form plans based on the resources that each has (you have the oars, she has the boat). These plans will divide into subplans (you ask for her boat, she kindly gives it to you) that follow the same norms of practical reason as above, so that these sub-plans will mesh to get you into a boat on the lake with the oars in your hands. Your friend's will is unified with yours: it is as though you two are one agent, acting to satisfy your extrapersonal desire.

Most of the telic interpersonal desires we have seen so far have the T-structure. For example, domination and intolerance exemplify disunified desires, while kindness and generosity exemplify unified desires. Here we also find the "caring" aspect of acts of love, where one person aims to alleviate the other's suffering, or to help further the other's most important world-directed ends. This caring aspect of love is the formal antithesis of cruelty; as cruelty is at the negative pole of interpersonal value, so care is at the positive pole.

Structure (2) is the Λ-structure, where agents want to act together. Again the template is the unified will of a single agent; as Bratman says, there are deep structural similarities between individual and shared agency.[77] When you row alone, you satisfy your desire by pulling the two oars in unison. When you row with your friend, you both satisfy your desires by sitting side-by-side and pulling the two oars in unison. In the Λ-structure, it is as though you two are one agent, acting to satisfy a desire that you share.

The formal model of the Λ-structure sets out the logic of value in shared agency. Merely instrumental shared agency ("I want to *phi* with you") is just another form of neutral unity and so generates no intrinsic value ("I want to row with you to get home"). It is only the telic form of shared agency that generates interpersonal intrinsic value; we can mark it with the locution "I want us to *phi* together."

The logic of the Λ-structure telic desire is distinctive. Here value acts as if each agent has a familiar (T-structure) interpersonal desire, yet the object of this desire is the satisfaction of the desire of a group agent, of

---

77. Michael Bratman, *Shared Agency: A Planning Theory of Acting Together* (New York: Oxford University Press, 2014), 4.

which the agent is a part. So, A's telic desire "I want us to row together" analyzes as A wanting the satisfaction of P's extrapersonal desire to row, where P is a group agent consisting of A and her rowing partner. Thus, value acts as if four-limbed A wants eight-limbed P to row, and the satisfaction of A's desire generates interpersonal $(+1)$ value. When many agents each have a telic desire for shared action ("I want that we play the quartet together"), the satisfaction of all their desires generates additive interpersonal $(+4)$ value.

In this formal model of the Λ-structure telic desire, value acts *as if* a group agent exists; the model itself carries no controversial metaphysical commitments. And, once this posit is in place, the rest of the formalism refers only to the desires of individual agents. For instance, the model accepts "mixed pairs," where, say, A_1 has a telic desire to row together with A_2, while A_2 has only an instrumental desire to row with A_1 (say, to get home). More, all satisfiable desires for shared action must be desires to perform some specific action: A_1 wants "that we *row* together." If A_1 has an additional telic extrapersonal desire to do her part in the shared action (here, to pull her oar), then the satisfaction of that additional desire will obey the usual rules. So the satisfaction of A_1's additional telic extrapersonal desire to pull her oar will also generate extrapersonal value $(+1)$. To take a different example of shared action, if A_1 has an additional telic interpersonal desire to do her part in an anti-social action ("I want us to mug strangers together"), then its satisfaction will also generate interpersonal disvalue (-1).

In the telic T-structure, it is as if multiple agents are one agent wanting to satisfy the desire of one of them. In the telic Λ-structure, it is as if multiple agents want to join their wills together to act on a shared desire. Telic desires for shared action are familiar—say, to sing together or do philosophy together for its own sake. In larger groups, we find desires for fellowship, fraternity, camaraderie, society. In couples, we find desires for closeness, togetherness, intimacy. Desires for shared action also define the second aspect of loving actions: the "sharing" aspect. This sharing aspect of love is seen in families who "do everything together."[78]

Successful telic T-structure action can produce positive value both "on the A side" and "on the P side." Successful telic Λ-structure action can produce positive value on each side of the shared agency. Thus pure altruism and

---

78. J. David Velleman reviews the caring and sharing aspects of love in "Love as a Moral Emotion," *Ethics* 109, no. 2 (1999): 338–74.

pure shared action both hold more potential for value than pure competition, because in the former cases both parties' desires can be satisfied.

Many positive relations analyze as complex blends of T-structure and Λ-structure desires, mixed with desires from the other dimensions of value as well. Consider a Rawlsian citizen's desire for *reciprocity*. This citizen wants to cooperate under principles that benefit all citizens, partly for its own sake, partly for the sake of others, and partly for her own sake.[79] A purer blend of positive interpersonal desires is *solidarity*. In Sangiovanni's analysis, those acting in solidarity are willing to incur significant costs for each other as they work together to achieve their common goals.[80] T-structure and Λ-structure desires, as well as blends of the two, will also define the logic of the third and final dimension of intrinsic value, intrapersonal value.

IV. INTRAPERSONAL VALUE

Interpersonal value emerges from a logic of desires. This is why the language of interpersonal value translates so easily across places and times—why we can understand the value judgments of the ancient Hebrews and the ancient Chinese, for instance. To evaluate what people are doing to or for or with each other, we need only look at the formal relations among their ends, not at the contents of their ends, which vary enormously between individuals and eras. Cruelty is cruelty, and help is help, whatever humans are wanting to do.[81]

So far, we have located a variety of thick evaluative concepts by plotting various formal relations between people's desires. We can test this approach as rigorously as possible by applying it to the analysis of the third type of desires: intrapersonal desires. If the formal structures the ordinary, then transposing the logic of interpersonal relations to the intrapersonal dimension should also find significant thick evaluative concepts.

Intrapersonal desires have one's own desire-satisfaction as their object. In a T-structure, one is one's own "patient," and when life is going well, one's will is unified with itself.

---

79. John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), 48–54.

80. Andrea Sangiovanni, "Solidarity as Joint Action," *Journal of Applied Philosophy* 32, no. 4 (2015): 340–59.

81. Note also that the objects of interpersonal desires can be the desire-satisfactions of non-human beings.

We have already met synchronic intrapersonal desires in extreme masochism: wanting to experience unwanted sensations. Such second-level intrapersonal desires are isomorphic to second-level interpersonal desires, and our evaluations of their satisfaction are parallel. Just as the anti-social is bad, so is the "anti-personal."

Diachronic intrapersonal desires are more interesting, because of how they transpose the logic of the interpersonal to the self across time. Just as one can aim to satisfy or thwart another person's telic desires now, so one can aim to satisfy or thwart the telic desires of one's own past and future "person-stages." The parallel to altruism is most obvious here. Just as one can give up one's own desire-satisfaction for the sake of another person's ends now (say, to avert their pain), so one can give up one's current desire-satisfaction for the sake of one's future ends (say, to avert one's own pain). This is the domain of *prudent action*. The associated ordinary terms indicate how valuable this relation to one's own ends can be: it is good for one's actions to be *moderate*, *temperate*, and *foresightful*, instead of being *foolhardy*, *reckless*, or *rash*.[82]

Prudent action is an intrapersonal transposition of the T-structure "caring" form of interpersonal value: one takes care of oneself over time.[83] For our test to go well, we should also find a transposition of the Λ-structure "sharing" form, where one acts with oneself over time, as if one were an intertemporal group agent. The structure here is where each person-stage has a telic desire for a state of affairs produced by the interrelated efforts of the person-stages over time. In other words, one's desires are to achieve a goal, complete a project, or (more grandly) pursue a career or embody an identity.[84]

---

82. Imprudent action is the intrapersonal parallel of the second sense of "anti-social action" above: here, positive present value is swamped by negative future value.

83. An Editor asks about Parfit's Future Tuesday Indifference (Parfit, *Reasons*, 123–24), which can have two versions. In one version, the person in question has a telic desire, which is a conditional but telic desire for pain (i.e., that any future pain be "Future-Tuesday pain"). Satisfying such a telic desire for pain will be, as we know, bad. In Parfit's own version, the person lacks a telic desire: the person has no telic desire regarding pains on future Tuesdays, which is what explains his preference for having pains on those days. This lack-of-desire seems irrational, because the person could have a more valuable life if he instead had the familiar, prudent intrapersonal desire to avoid future pain proportionate to its intensity, on whatever day it will occur.

84. Bratman draws parallels between diachronic self-governance and shared intentional agency in *Planning*, chaps. 5, 11.

If this idea of intrapersonal group agency is not just metaphorical, there should be concepts parallel to "solidarity" for the self. And so it seems there are. Let us first set up a case of interpersonal solidarity as a template for intrapersonal solidarity. On the day of the Boat Race, it may be hard for the stroke to row because she has a calf cramp. The stroke bears her (extrapersonally bad) discomfort for the sake of the (interpersonally good) shared end of the crew. Now consider an intrapersonal parallel. It may be a major goal of yours to write a book, but on Chapter Five you develop writer's cramp. You overcome your (extrapersonally bad) discomfort for the sake of your (intrapersonally good) temporally-shared end of writing the book. Your current self, we might say, acts in solidarity with your past and future selves.

The ordinary value terms associated with this kind of intrapersonal unity are many. Your actions are determined, dogged, resolute, steadfast, tenacious, committed—instead of being impulsive, impetuous, or fickle. One apex of intrapersonal unity is self-mastery, one nadir is self-sabotage. Integrity, in its non-moralized sense, also fills some of this space of intrapersonal unity. Going deeper, having plans that connect person-stages is necessary for having a self at all, lest one become unhinged or deranged. Here, value flows from the unity of one's will over time.

One way for a life to have intrinsic value is for it to be oriented around "ground projects," where what one does now also helps to satisfy the telic desires of one's past and future selves. One wants, across time, to have an academic career, or to be a philosopher. And it is noteworthy that the successful execution of such intrapersonal projects will have even more value the more the activities involved also satisfy extrapersonal or interpersonal desires at the same time. Success in living the life of a philosopher will be even more valuable the more one also enjoys doing philosophy, hour by hour. This parallels the interpersonal: it is good to help others—and it is even better to enjoy the activities through which one helps others, such as preparing meals or reading books for them. In such cases, the activity is "prolific": the activity satisfies several telic desires at once.

Finally for our test, we should also discover neutral unities and disunities along the intrapersonal dimension of value; and so we do. Today you may fulfill a telic desire to run a little farther (extrapersonal value), but relative to your (intertemporally shared) goal of running a marathon, today's run is merely a means—it has no extra, intrapersonal value. For a neutral disunity, we sometimes think it good to thwart some of our future desires as a means to satisfying others. Tonight one turns on the program

that blocks the web browser, because one knows that in the morning one will be tempted to open the browser when doing research—but that one will then be distracted by email. "Tying oneself to the mast" self-management allows more to be achieved.

In examining the three dimensions of intrinsic value, we have found that value tracks the unification of desires, both across persons and across times. What is ultimately good, it seems, flows from unity between will and world, and between will and will—as if all wills, at all times, meet the world as one.

V. FURTHER CHALLENGES

With the full formal model in hand, we can frame responses to the remaining challenges to the four traditional theories of intrinsic value with which we began. Consider the person who is chronically self-sacrificial.[85] With the resources we have now, we can see that a self-sacrificial person is (perhaps blamelessly) overfocused on interpersonal value, at the expense of intrapersonal and extrapersonal value. This person would have a more valuable life if she could devote more time to her own health and projects, instead of putting her own needs aside to satisfy her husband's whims. As the relationship becomes more limiting or abusive, her life gets worse.

Self-sacrificial desires may also be adaptive preferences. Understanding the normative concern about adaptive preferences requires broadening the focus to include the environments that the preferences are adapted to. What is concerning about adaptive preferences is not that people's desires have adapted to their environments (which is quite common). It is that some people's environments are unjustly impoverished or discriminatory, resulting in their having desires that realize unjustly limited value.[86] The challenge of explaining adaptive preferences is thus a challenge within political theory, not within value theory proper. The challenge for political theory is to explain which environments are unjustly impoverished or discriminatory, which will reveal which preferences have been "causally shaped by injustice."[87]

85. For example, the young Vasanti in Nussbaum's *Women and Human Development*.
86. See, for example, Serene Khader, *Adaptive Preferences and Women's Empowerment* (New York: Oxford, 2011).
87. David Enoch, "False Consciousness for Liberals, Part I: Consent, Autonomy, and Adaptive Preferences," *Philosophical Review* 129, no. 2 (2020): 161.

The challenge of worthless or trivial desires is squarely within value theory; the best-known example is Rawls's talented mathematician who can only want to count blades of grass.[88] Should such a person exist, I believe that our regret would not be with what he does, but with what he cannot do. Many activities, like bird watching and stamp collecting, are not so different from grass counting, except that these are "prolific activities," where a single activity registers on multiple dimensions of value at once. A success in such activities can be satisfying (extrapersonal value), at the same time as it contributes to socializing (interpersonal value) and a life project (intrapersonal value). Our disappointment with the grass-counter is that his psychology demands repeating an activity that satisfies only a single desire. If only he could use his mathematical skills, say, to design better educational software—or to further the grand human project of mathematical inquiry by constructing new proofs.[89]

Prolific activities also answer the classic challenge of higher and lower pleasures. Any "higher" pleasure will satisfy several telic desires at once. For instance, compare eating a meal; eating a meal with friends; and eating, with friends, a meal that one has long been perfecting. Prolific activities also illuminate other familiar challenges to hedonism, such as "Truman Show" cases of false friends, and Nozick's experience machine. First, it is a familiar point that many of the desires of the person in these constructed environments—such as desires to spend time with friends—are not actually satisfied and so generate no value.[90] Yet there is then a further question: should one plug in to the experience machine and program it to give one the optimal series of physical pleasures?

Unity theory says that the answer depends on one's options.[91] Some people's options are so limited that a series of physical pleasures (from the experience machine, or from certain drugs) will be their best choice.

88. Rawls, *Theory*, 424–33.

89. Rawls's grass-counter presents a carefully designed example of an *isolated* activity, which is an activity that satisfies only one desire. We react negatively to isolated activities because they sacrifice so much value, compared to the familiar prolific alternatives. Indeed, as above, a person whose isolated activities show no regard for her past or future desire-satisfaction may even seem unhinged or deranged. Nearly all of our activities are prolific to some degree, as can be seen by examining one's own quotidian projects.

90. Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 42–45.

91. Robert Nozick, *The Examined Life: Philosophical Meditations* (New York: Simon and Schuster, 1989), 105.

Yet most people are more fortunate: they can engage in prolific activities that satisfy interpersonal and intrapersonal desires as well as extrapersonal. Indeed activities can be prolific within dimensions as well as across them: volunteering at the homeless shelter may satisfy telic desires both to help the residents and to help one's community; a walk in the mountains may stimulate many creative thoughts. Prolific activities are the richness of life.[92]

VI. VALUE PLURALISM

One corroboration of unity theory is how extensively it explains our firm evaluative judgments. Judgments that might have appeared to be perceptions of desire-independent value turn out to be tracking formal relations among desires. Consider the many thick evaluative concepts that have yielded to the analysis over the course of this investigation. The "formal" lives within the "ordinary," and the logics of desire explain a wide range of convictions about what is good and bad in itself.

The badness of anti-social actions has been a particular challenge to all four traditional theories of value. The new theory explains this badness as disunity, and shows how this structure also explains the badness of "antipersonal" actions like masochism. This structural approach yields an attractive value pluralism. In our relations with the world, value is maximally pluralistic. Every state of the world that people want for its own sake registers on the extrapersonal dimension of value. Because humans are so varied in their constitutions and concerns, an extraordinary variety of activities can be valuable. More, when people join together for world-oriented activities, these can also register on the interpersonal dimension and so be more valuable still. Yet, because of the logics of interpersonal and intrapersonal value, the satisfaction of desires that are hateful or self-destructive will be bad.

So unity theory is maximally value-pluralistic, within the bounds of decency and prudence. Innumerable "experiments in living" can be good in themselves, as can an enormous range of cultural practices—while intolerance and racist domination are basically intrinsically bad. Overall, the theory says that value will be greatest when people do what they want

92. The theory here also explains multiply-aversive "abysmal conditions." See Leif Wenar, "The Development of Unity: The Martha Nussbaum Lecture 2019," *Journal of Human Development and Capabilities* 21, no. 3 (2020): 211–22.

in the world, which can be as diverse as human imagination, while everyone at least tolerates the diversity of others, and, at best, enjoys it.

VII. CONCLUSION

Most of this article has set out the formal model of basic intrinsic value, along the three dimensions. The article's main aim has been to show that this new object-level theory of value is a live alternative to the four traditional theories, because it better meets traditional criteria for theory-selection (simplicity, fruitfulness, extensional fit) and because it overcomes challenges that have stymied the other theories.

Along the way, I offered a partial answer to the question of why these formal structures work to systematize our evaluative judgments. This is the "many-one" interpretation of the T-structure and the Λ-structure of interpersonal desires, which draws a parallel between interpersonal value and individual rationality. I realize that some readers may have found this interpretation helpful, while others may have had their doubts. At the risk of provoking this latter group even further, let me here offer one last suggestion about why we find these two specific formal structures when we analyze value. After all, it might still be wondered why the T-structure and the Λ-structure *in particular* define the whole space of interpersonal value. This last suggestion will be even less than a sketch of an interpretation—in fact, I can only offer two images within a myth.

The reader will recall the Myth of Aristophanes in Plato's *Symposium*.[93] In this myth, human beings were originally eight-limbed creatures, whom the gods later split in half, which explains why lovers now yearn to reunite with their "other halves." We can put this myth to our purposes here by imagining a rational "unsplit" eight-limbed creature, who wants to go for a row. Our eight-limbed creature gets a boat and sets out on the lake, pulling each oar with two of their four hands.

Consider now that a split of this eight-limbed creature into two four-limbed creatures can happen in two ways. In the first (T-structure) way of splitting, the first four-limbed creature has the desire to row and the second has the boat. Here, value is generated when the second four-limbed creature *gives* the boat, in the same way that the rational eight-limbed creature *gets* the boat. In the second (Λ-structure) way of splitting, the two

93. 189c–193e.

four-limbed creatures each have a desire to row, and they have the boat, and value is generated by the creatures joining their wills to row together side-by-side. The first way of splitting is "horizontal," separating the desire to row from the means to row. The second way of splitting is "vertical," turning a single desire to row into two desires to row together. The T-structure and the Λ-structure are thus the two ways that a willing creature can be "divided," and each structure then indicates a way that value is generated when two wills are "united" (or, as in the Platonic myth, "reunited").

Even if these interpretative suggestions pan out, many metaethical questions will remain about the new theory. My hope is that enough has been said to make these questions worth considering, and that philosophers will welcome having a new object-level theory of value to explore.

The root of the word "good," etymologists tell us, is a proto-Indo-European word that means "to unite."[94] The thesis of this article has been that that goodness comes from unity: unity with the world, unity with each other, and unity within ourselves.

94. Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981), 418.