2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.82063

Editorial: Sin and Vice

MARIA SILVIA VACCAREZZA
University of Genoa
mariasilvia.vaccarezza@unige.it

MICHELE PAOLINI PAOLETTI
University of Macerata
m.paolinipaoletti@unimc.it

Albeit often overlapping, sin and vice are two related yet clearly distinguished concepts. It would require an extremely long detour to sketch the history of such thick, central notions in Western culture, as well as their similarities, differences, and overlappings. Sin is a polysemic notion. The word refers, primarily, to an *action* whereby an agent intentionally fails to live up to the will or commands of God (see Stump, 2018). However, sin can be also considered as a *disposition* of the will, i.e., an inclination to engage in sinful actions (Plantinga 2000). Finally, according to parts of the Christian tradition, sin is a metaphysical state, i.e., a condition of "uncleanness" which marks a fundamental ontological difference between human beings and God (M. Adams 1991, 20f). While according to the Christian thought vice can be a cause of sin, and a sinful disposition may closely resemble a vice, the two concepts shouldn't be confused. First, while sin pertains to a theological vocabulary, talking of vices is perfectly compatible with a secular philosophical discourse, as the recent success of virtue ethics and virtue-vice epistemology testifies. This comes as no surprise, if one thinks that the very origin of the concept is rooted in the secular philosophical tradition tracing back to Plato and Aristotle. Secondly, within character-based moral and epistemic theories, a vice is, unequivocally, a trait of character. As such, it is part of an agent's psychological makeup, which means that it is a stable feature of their moral psychology, independently of its specific behavioral manifestations. And while an action may well be vicious, it can be so only insofar as it springs from a vicious trait, which has—so to speak—ontogenetic, conceptual, and even normative priority over the actions it elicits.

In this special issue, we aim to further clarify the nature of sin and vice and the connection between sin and vice. The issue includes seven contributions. Three contributions concern topics in philosophy of religion, two further contributions deal with metaphysical issues and the last two contributions mostly focus on ethical problems.

Kevin Timpe, in his "The Inevitability of Sin", explores the idea that all humans (except Christ and maybe Mary) cannot freely avoid sinning, at least

following the Christian doctrine of original sin. Timpe models this idea using Lewis' theory of counterfactuals. Indeed, the inevitability of sin is a modal notion: roughly, if a human were in specific conditions, s/he could not avoid sinning. However, to understand this notion and evaluate the relevant counterfactuals connected with it, it is necessary to restrict the set of accessible worlds (from the actual world in which humans cannot avoid sinning) by introducing specific restriction condition(s). The author then shows how such restriction conditions should be singled out from the standpoint of two theological doctrines, i.e., theological determinism and Molinism. On theological determinism, the restriction condition is a certain volition of God, i.e., that each human in the relevant world commits at least one sin. On Molinism, the restriction condition is Plantinga's transworld depravity thesis. Finally, Timpe faces libertarianism. Libertarianism seems to be at odds with the inevitability of sin. However, the author suggests two options. The first option consists in accepting two theses: that someone may freely and responsibly do something even without having alternative possibilities and that all humans are (partly) identical with the first sinner by virtue of some volition of God. The second option amounts to operating on the restriction conditions for the accessibility relation connected with the relevant counterfactuals.

In "On the Privation Theory of Evil: A Reflection on Pain and the Goodness of God's Creation", Parker Haratine reconsiders the privation theory of evil in connection with two theological theses, i.e., that God is the creator of everything and that being and goodness are interconvertible. The privation theory of evil seems to follow from the conjunction of these theses. According to it, evil cannot have a positive existence or, better, it is entirely ontologically dependent for its existence on something else, which is good. Thus, evil is the privation of the latter. Otherwise, *qua* existent, evil would turn out to be good. However, Haratine examines pain as a classical counterexample that stands in the way of the privation theory of evil. Pain is evil, but it is not necessarily a privation. Pain has a positive reality. The author defends this thesis against three main responses on behalf of privation theorists: that pain is not real, since it is an evaluative mental state; that pain is not evil, since it has some function or utility; that pain actually is the lack of something, so that it squares well with the privation theory. Haratine's conclusion is twofold. First, even if pain is a positive evil and the privation theory of evil is false, the privation theory does not actually follow from the theological theses mentioned above. For those theses are actually compatible with the non-existence of evil, which is incompatible with the privation theory. Secondly, one may accept an alternative view of evil: the opposition view. According to the latter, evil is good insofar as it enjoys absolute existence—which is good by itself. However, evil is bad insofar as it does not have kind-existence,

i.e., it does not have any proper nature or end. In this respect, evil is nothing but the opposite of something which enjoys kind-existence.

Helen De Cruz and Johan De Smedt, in their "Schleiermacher and the Transmission of Sin: A Biocultural Evolutionary Model", explore Schleiermacher's account of the transmission of sin. They first point out that the traditional Augustinian account of original sin—which includes one specific 'primal sin' episode and some biological mechanism of transmission of sin—is affected by at least two drawbacks: it is not supported by current biological evidence as concerns the existence of a single ancestral pair and it does not explain why transmission mechanisms were implemented by God in the first place. Schleiermacher's account is based upon God-consciousness, i.e., self-consciousness and species-consciousness coupled with a feeling of absolute dependence upon God. God-consciousness emerges in humans from biological mechanisms, when humans start to perceive their relative dependence upon natural resources and then recognize that God is the source of nature itself. Sin enters the stage when there is a mismatch between one's own God-consciousness and one's own social and bodily self-consciousness. This mismatch is due, on the one hand, to our 'seeds of sin', i.e., our biological tendencies to commit evil acts, and, on the other hand, to the negative impact upon ourselves of the cultural communities where we live, that contribute to transmitting sinful tendencies. Therefore, Schleimermacher's model is a 'biocultural' model of the transmission of sin. The authors then present empirical evidence for the existence of God-consciousness and for the social transmission of sin, through the influence of peers, parents and social models. Finally, they adapt the cultural Price equation in order to show how changes in adhering to social norms from one generation to another turn out to depend upon the number of cultural descendants of an individual (divided by the population mean number of cultural descendants), the perceived cultural prevalence of that trait (i.e., adhering to social norms) and some cognitive attractors that work as distorting biases.

In their "Vices, Virtues, and Dispositions", Lorenzo Azzano and Andrea Raimondi draw some interesting connections between the metaethical inquiry into the nature of virtues and vices and the metaphysics of dispositions. They first point out that virtues are endowed with genuinely dispositional nature: one has a given virtue insofar as one is disposed (to a certain degree) to behave in given ways. Indeed, like dispositions, all virtues have typical manifestations and triggering circumstances. Like a disposition, each virtue may be possessed even if and when it gets unmanifested. And it may come in degrees and be interfered with. However, vices need not be thought of as dispositions. But the metaphysics of dispositions may be exploited in order to understand their nature. Indeed, the authors distinguish between three types of vicious persons: the incontinent, the malevolent and the indifferent. These types of persons may produce the same

behavior, e.g., failing to help someone in need. Yet, they may produce the same behavior in different ways. The incontinent possesses the virtuous disposition to help someone in need, but s/he does not exercise it because of systematic interferences due to negative character traits and/or negative external circumstances. In this respect, the incontinent only mimics the exercise of a vicious disposition, without possessing the latter. On the contrary, the malevolent actually possesses vicious dispositions and s/he regularly exercises them by deliberation. Finally, the indifferent lacks both virtuous and vicious dispositions. Thus, when s/he finds herself in the typical circumstances in which a given virtuous disposition should be exercised (e.g., meeting someone in need), those very circumstances cause the loss of the disposition. They act as 'finks'.

In "Presentism, Timelessness, and Evil", Ben Page reflects upon divine timelessness and evil from the standpoint of presentism. Divine timelessness seems to be incompatible with the defeat of evil at some time in the history of creation. Since the latter is part of Christian teaching, divine timelessness seems to be incompatible with Christian teaching as well. To account for the defeat of evil, one could resort to presentism. Yet, presentism needs to reject divine timelessness. Against this line of reasoning, Page argues for two theses. The first thesis is that it is possible to make sense of divine timelessness even from a presentist perspective. For one may figure out a possible world with two island universes that are temporally disconnected from each other. Both universes are presentist. In one of such universes, there is only one instant, that it is 'eternally present', so to say. It is the universe of God. In the other universe, there is only one present instant (at a time) but there is also flow of time. This is the universe of creation. Some complications are in order, since tenses and existence become relative to island universes. Moreover, one may introduce some tenseless notion of existence and argue that, even in the universe of God, the instants and the entities living in the universe of creation enjoy tenseless existence, so that evil enjoys tenseless existence as well. At any rate, evil may get defeated in the universe of creation as time passes by and it stops existing. But Page also argues for a second thesis. According to it, evil never gets completely defeated in the 'presentist' universe of creation. Indeed, consider a time at which evil does not exist anymore. Still, at that time, there are past truths about past evils, which are made true by presently existing truthmakers. Such truthmakers, even if they do not bring about pain or suffering and even if they may only consist of vivid memories in God's mind, still reintroduce evil in the universe of creation.

Ian James Kidd, in his "From Vices to Corruption to Misanthropy", explores the connection between failings/vices, corruption and misanthropy. "Failing" is a term broader than "vice", since it is not necessarily connected with the Aristotelian tradition. Kidd recognizes that failings are diverse, that some of them are temporary, whereas others are persistent and/or linked to specific

worldviews. Some failings only involve individuals, whereas others also or only involve collectives (e.g., institutions, companies, societies, and so on). Corruption results *from* failings and results *in* failings. More precisely, being exposed or subjected to corrupting conditions either results in the erosion of virtues or in the introduction or the strengthening of vices/failings. The author singles out different types of corrupting conditions and different ways in which such conditions work (e.g., through the acquisition of new failings, through the activation of latent failings, and so on). In turn, generalized corruption may lead to misanthropy. Kidd accepts one revisionary account of misanthropy—inspired by David Cooper's works—according to which misanthropy is a negative appraisal of the collective character and performance of humankind. Misanthropy may bring about different manifestations, stances and behaviors. It may come together with different affects. At any rate, it is typically produced by the experience of failings that are ubiquitous and entrenched in humankind. Kidd concludes that there is also room for Christian misanthropy, insofar as Christians recognize that humankind lives in a sinful condition, i.e., in a condition characterised by fundamentally disordered tendencies when they pursue goodness apart from God.

Charles Taliaferro and Emily Knuths, in their "How Sinful Is Sin? How Vicious Is Vice? A Modest Defense of the Guise of the Good", defend the guise of the good thesis, according to which everyone acts upon what s/he conceives of as good, or as least evil. They tackle three counterexamples against this thesis, i.e., that one may disapprove of one's own actions and still feel compelled to make them, that one may act upon conflicting and irrational urges and impulses and that one may even seek annihilation or self-annihilation. Interestingly enough, in the latter case, they argue that one still acts upon what counts as least evil (i.e., destroying something that is perceived as bad) or that one pursues the good by identifying it with punishment or revenge. Taliaferro and Knuths then claim that this thesis favors moral realism over moral subjectivism. Moral subjectivism is incapable of explaining our moral attitudes when we do something because we believe it is good. And it is incapable of providing reasons for abstaining from some profoundly evil actions when we feel compelling urges towards them. They also deal with two objections against moral realism, i.e., that it makes rightness and wrongness mysterious and that it is at odds with the higher number of existing moral disagreements. Finally, the authors examine the figure and the actions of Darth Vader in light of the guise of the good thesis.

Bibliography

Adams, M. M. 1991. "Sin as Uncleanness". *Philosophical Perspectives*, 5: 1–27. https://doi.org/10.2307/2214089.

Plantinga, A. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press. https://doi.org/10.1093/0195131932.001.0001.

Stump, E. 2018. *Atonement*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198813866.001.0001.

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.65953

The Inevitability of Sin

KEVIN TIMPE
Calvin University
kevin.timpe@calvin.edu

Abstract: Part of the traditional Christian doctrine of sin is the claim that, due to the effects of original sin, acts of sin are inevitable. Of course, our reflection on sinful actions is shaped by how we think about human freedom and divine providence more broadly. Some have argued that libertarians have a difficult time accounting for the inevitability of sin. This paper uses David Lewis's work on counterfactuals and possible worlds to give an account of how the inevitability of sin can be understood. It then shows how both theological determinism and Molinism can give an account of sin's inevitability so understood. In doing so, I show that sin's inevitability only follows if we restrict our focus to certain possible worlds that share certain features with the actual world. But once such a restriction is made explicit, I then develop a framework for how the libertarian non-Molinist can also use a similar restriction to give an account of the inevitability of sin given original sin.

Keywords: Sin, Inevitability, Original sin, Libertarianism, Theological determinism, Molinism

Introduction

The Christian doctrine of original sin is understood as a kind of state or "condition" that humanity is.[1] In his work, Jesse Couenhoven describes original sin as original in the sense that "it is an evil at the origins of human agency, and from which human agency flows."[2] Drawing on Augustine, the Christian theologian whose thought is most closely associated with the doctrine of original sin, Couenhoven argues that the doctrine of original sin has five parts and distinguishes between them as follows:

[1] See Mann 2001, 47.
[2] Couenhoven 2016, 193. See also the discussion in Mellema 2021.

(a) the primal sin;
(b) the participation of the rest of the human race (except Jesus and, in some traditions, Mary) in that sin because of their solidarity with those who committed the first human sin;
(c) involuntary and inherited common guilt that all humans (again except Jesus and, in some traditions, Mary) are subject to because of that solidarity;
(d) penalty to human nature because of the primal sin; and
(e) the transmission of inherited sin and its penalty.[3]

While some associate original sin with the first of these, the primal sin understood as the temporally first sin, one can think of the primal sin as the cause of or occasion of original sin; in what follows, I set aside any further discussion of the primal sin.[4] Part (b) of the doctrine of original sin is a claim about the scope of those who are affected by original sin; I address it in section 1 given that it puts constraints on how we should think about the remaining parts of the doctrine of original sin. I then turn, in the rest of the paper, to claim (d), what Oliver Crisp calls "an inherited corruption of nature"[5] and in particular the claim that part of the penalty to human nature due to the effects of original sin, acts of sin are inevitable.[6] While the focus in what follows is entirely on acts of sin, I'm inclined toward agreeing with Ian McFarland's claim that acts of sin "do not get at the heart of the phenomenon of sin"[7] and certainly don't exhaust the kinds of sin we need to think carefully about. But every article must have its focus, and mine here is on acts of sin. I call the claim that, given (d) humans are now unable, on some understanding of 'unable', to avoid sinning 'the inevitability thesis'. In section 2 I draw on David

[3] Couenhoven 2013, 46; for slightly different categorizations, see Blocher 1997 and Crisp 2019, chapter 7. For an excellent recent paper on potential mechanism for the transmission of original sin, see Green 2022. Green argues that perhaps the transition of sin can be understood as involving maturational naturalness: "It isn't carried by our genes in any deep sense. It is not communicated mysteriously, one immaterial soup to another in the womb. It certainly isn't placed in the soul by divine fiat. Rather, if one takes human nature together with what human environments have in common, the result is a human disposition to be alienated from God, self, others, and nature" (35). Such a disposition is triggered by a salience framework that is modeled for us by other human agents.

[4] For a range of treatments, see the materials discussed in Timpe 2014, chapter 3, and Timpe 2021.

[5] Crisp 2015, 264.

[6] In what follows, I assume that free will is required for acts of sin. That is, I take an agent's moral responsibility for their acts of sin to presuppose that they have free will. See Timpe 2014, chapter 1.

[7] McFarland 2016, 303.

Lewis's work on counterfactuals to give one specification of how to understand the inevitability thesis.

It is sometimes claimed that certain views of human freedom, namely libertarian non-Molinist views, cannot account for the inevitability thesis, and thus that such views of freedom should be rejected by those who want to affirm traditional Christian doctrines (among which I think the inevitability thesis is numbered). But it's not clear that such an objection to libertarian non-Molinism holds. Section 3 shows how views of divine providence that endorse theological determinism and theological compatibilism, can account for the inevitability thesis using Lewis's account of inevitability. Section 4 attempts to provide a Molinist defense of the same. What these two sections show is that sin's inevitability only follows if we restrict our focus to certain possible worlds that share certain features with the actual world. But once such a restriction is made explicit, in section 5 I develop a framework for how the libertarian non-Molinist can use a similar restriction to give an account of the inevitability of sin given original sin. Perhaps there are reasons to reject Lewis's account of inevitability and the underlying understanding of conditionals it builds on. But what the paper attempts to show is that the criticism of libertarian non-Molinist views from the inevitability thesis is more complicated than it first appears. And I don't yet think it has been shown that the libertarian non-Molinist cannot endorse the inevitability thesis without contradiction.

1. Exceptions that Question the Rule

Original sin is sometimes referred to as human beings possessing a sinful nature. But such a view leads quickly to Christological worries. As theologian John E. McKinley understatedly puts it, "Jesus Christ is a complicated person."[8] According to Christian doctrine, there's something unique about the Incarnate Christ: he is fully human but also, given the hypostatic union between his human nature and the divine nature of the second person of the Trinity, as the Council of Chalcedon put it he is "equal in all things to us except sin." There are two parts to this uniqueness. First, there's the claim that the Incarnate Christ not only did not sin but could not sin, which Timothy Pawl refers to as 'the impeccability thesis.[9] The impeccability thesis prevents us from understanding original sin as involving a sinful human nature. If, as the Christian tradition has held, the Second Person of the Trinity

[8] McKinley 2021, 119.

[9] Pawl 2021, 94; see also Pawl 2019.

becomes incarnate and assumes a human nature, then if human nature were somehow itself sinful, Christ would also be sinful in virtue of assuming a sinful human nature. But Christianity holds that the Incarnate Christ is fully human as well as fully divine, and yet without original sin.[10] So if the impeccability thesis is true, then we ought not understand original sin as entailing that human nature became sinful through the impacts of the primal sin. Human nature, even post Fall, isn't inherently sinful: "the exclusion from sin from the combination of the true humanity of Christ seems to imply that true humanity does not need to entail the ability to sin."[11] Instead original sin should be understood as a distortion in human nature Because such language about human nature itself becoming sinful can be misleading[12] and also potentially in conflict with the conviction that all things created by God are good, original sin is perhaps better described in terms of human nature's being distorted.

However, the impeccability thesis by itself doesn't establish the uniqueness of the Incarnation. For that, it must also be the case that all other humans do in fact sin. For if other humans were also able not to sin and actually avoided sinning, then the Incarnate Christ wouldn't be unique. So to secure the second half of the uniqueness claim, one also needs the claim that all other humans are *able to* and *do* sin. On some views of human origins, Adam and Eve (whether as historical figures or as representatives of the first humans[13]) were able to avoid sinning but didn't, and thus were able to sin given that they did. Furthermore, some branches of Christianity hold that Mary, the mother of Jesus and *theotokos*, was in a similar position. Second, to establish that Jesus was unique in not sinning, it must also be the case that all other humans not only *can* but *do* sin. Here, of course, one runs headlong into Marian doctrine, as the Catholic Church holds that because of divine grace Mary is free from the effects of original sin and, as a result, was "free of every personal sin her whole life long."[14] Pope Pius IX wrote that "The most Blessed Virgin Mary was, from the first moment of her conception, by a singular grace and privilege of almighty God and by virtue of the merits of Jesus Christ, Savior of the human race, *preserved*

[10] For discussions of the relationship between Christ's divine nature and the human nature assumed in the Incarnation, see Pawl 2019 and 2016.

[11] Dockter 2021, 75. Similarly, if the impeccability claim is true, then the ability to sin cannot be a necessary condition for freedom; see Pawl 2021, 96 and Timpe 2014. For a different view, see Sumner 2014.

[12] See Copan 2003, 523.

[13] See van Inwagen 2006, lecture 5 for a discussion of human creation and sin that doesn't involve Adam and Eve being historical individuals.

[14] *Catechism of the Catholic Church*, 493.

immune from all stain of original sin."[15] The sinlessness of Mary is official Catholic doctrine, with Vatican II declaring that "the immaculate Virgin, *preserved free from every stain of original sin*, when the course of her earthly life was completed, was taken up body and soul to heavenly glory, and exalted by the Lord as queen of all so that she might be more fully conformed to her son, the lord of lords . . . and victor over sin and death."[16]

The Orthodox tradition is more complicated, given that it tends to have a different understanding of original sin, where the doctrine is less dogmatically circumscribed than it tends to be in the Latin West. It's also harder to specify what Orthodox dogma commits one to beyond the decrees of the ecumenical councils. Many doctrines not explicitly binding by the ecumenical councils end up locally promulgated as a common theological commitment. In this way, many Orthodox Christians hold that while Mary was born into the same sinful world as the rest of us are and thus needed to be redeemed from sin and its effects in some sense, she did not commit any personal acts of sin. I think it's accurate to say that there is a general Orthodox consensus that Mary committed no acts of sin.[17]

Some Protestant confessions seem to at least potentially allow for Mary to be excluded from the scope of original sin insofar as their statements regarding the doctrine have implicit quantifier restrictions. Article 15 of the Belgic Confession, for instance, states that "by the disobedience of Adam original sin has been spread through the whole human race;"[18] but clearly the 'all' isn't intended to be an unrestricted quantifier given the Incarnation. Something similar can also be said of article 16's claim about "all Adam's descendants having thus fallen into perdition and ruin by the sin of Adam."[19] The Canons of Dort make clear that Jesus alone is excluded from the scope of the quantifier: the corruption of original sin "spread, by God's judgment, from Adam and Eve to all their descendants—except for Christ alone— . . . by way of the propagation of their perverted nature."[20] Belief in the

[15] Pius IX, *Ineffabilis Deus* (1854): DS 2803, as quoted in *Catechism of the Catholic Church,* 491.

[16] *Decrees of the Ecumenical Councils* volume 2, 894f.

[17] I thank Jon Jacobs, Jeremiah Carey, James Dominic Rooney, Nate Placencia, and Omar Fakhri for helpful conversations here.

[18] Belgic Confession article 15; in *Our Faith: Ecumenical Creeds, Reformed Confessions, and Other Resources*, 40.

[19] Belgic Confession article 16; in *Our Faith: Ecumenical Creeds, Reformed Confessions, and Other Resources*, 41.

[20] Canons of Dort, The Third and Fourth Main Points of Doctrine, article 2; in *Our Faith: Ecumenical Creeds, Reformed Confessions, and Other Resources*, 130.

immaculate conception and sinlessness of Mary is explicitly ruled out by other Protestant denominations as well. The Episcopal Church's article IX on original sin claims "it is the fault of corruption of the Nature of every [human], that naturally is engendered of the offspring of Adam; whereby [humanity] is very gone from original righteousness," and article XV explicitly claims that other than the Incarnate Christ, "all the rest [of humanity] . . . offend in many things."[21]

But having now discussed the exception(s) to the rule, I set them aside. In sections 2 through 5, all utterances of the sort that "all humans *such and thus*" should be interpreted as implicitly restricted in light of the discussion in this section.

2. The Inevitability Thesis

With the exception(s) established, I turn then to the rule (though the exception(s) should be kept in mind, even though I will no longer make them explicit). The 'rule' is found in Couenhoven's claim (c) of his treatment of original sin, which holds of humans in virtue of the solidarity mentioned in claim (b). He describes this third component as the conceptual center of the doctrine of original sin, with the other parts of the doctrine providing "a background that assists us in understanding the center of the doctrine."[22] He understands this conceptual center to itself be the conjunction of two smaller claims: the first being about constitutional fault, which he sometimes calls "original sin itself,"[23] and the second being that one is morally blameworthy or suffers from original guilt in virtue of that constitutional fault. While I take Couenhoven to be correct that much of the historical reflection on original sin involves both of these elements, the second of these claims has been widely rejected by many philosophers writing on original sin.[24] Fortunately, for present purposes, I can sidestep the issues of original guilt and focus on constitutional fault.

---

[21] *Book of Common Prayer*, 1979 version, articles IX and XV, 869f. Article VII of the Methodist Articles of Religions parallels the article IX of the Episcopal Church on this point though without the explicit exemption of Jesus that one finds in article XV.

[22] Couenhoven 2013, 47.

[23] Couenhoven 2013, 23.

[24] Thomas McCall refers to those views that reject original guilt as "corruption-only doctrines" (McCall 2019, 156). The denial of original guilt is found in the Catholic tradition (see *Catechism of the Catholic Church* 2003, part I, section 2, paragraph 7, 405) and many of the Orthodox traditions (see Louth 2020). Among contemporary philosophers, corruption-only views are endorsed by Crisp 2020, McFarland 2016, McFadyen 2016, Hudson 2014, Wyma 2004, Plantinga 2000, Quinn 1997, Morris 1992, and Swinburne 1989. William Wainwright, in contrast, holds that original guilt is essential to the doctrine of original sin (Wainright 1988, 31).

Couenhoven describes constitutional fault as a kind of moral "improper functioning *qua* human being" that involves "an inherited state of disordered desire and ignorance."[25] Acts of sin arise from this constitutional fault.[26] While some have held this disordering to only incline one to commit acts of sin,[27] it's much more common to understand it as implying that acts of sin are inevitable.[28] The Council of Trent, for instance, anathematized all who hold that Adam and Even's sin didn't defile the whole human race by causing the loss of original holiness and justice such that they will sin.[29] In a later session, the council made it even clearer that because of original sin humans are not just inclined toward sin but will commit acts of sin:

> If anyone says that a person, once justified, cannot sin any more or lose grace, and therefor that one who falls and sins has never been truly justified; or, on the other hand (apart from a special privilege from God such as the church holds in the case of the blessed Virgin), *that he can avoid all sins, even venial sins, throughout his life: let him be anathema*.[30]

The inevitability of committing acts of sin given original sin shouldn't be understood as the view that particular sins are themselves inevitable, but rather that some sin or other is inevitable. As Paul Copan puts it, "though we do not sin necessarily (that is, it is not assured that we must commit this or that particular sin), we sin inevitably (that is, in addition to our propensity to sin, given the vast array of opportunities to sin, we eventually do sin at some point)."[31] This is, as I've called it, the inevitability thesis.

But how exactly we should this claim that given the impact of original sin on the human condition, all humans are now unable to avoid sinning?[32] In an interesting paper, W. Paul Franks understands the inevitability thesis as follows:

---

[25] Couenhoven 2013, 12 and 30.

[26] This claim too needs to be understood as having an implicit restriction; it is not the case that the initial (primal) human sin arose from constitutional fault.

[27] see for instance Swinburne 1989, 138.

[28] See for instance Crisp2019, 150.

[29] Session 5 1st decree, paragraph 5-6; in *Decrees of the Ecumenical Councils* volume 2, 667.

[30] Session 6, in *Decrees of the Ecumenical Councils* volume 2, 680 (emphasis added).

[31] Copan 2003, 531.

[32] If you felt any inclination to argue with the 'all' in this sentence, reread the last sentence of the previous section.

> (1) Necessarily in a world tainted by original sin, (a) every human subsequent to Adam and Eve is born in a condition such that it is inevitable that she sin (given that she performs at least one morally significant action), but (b) it is not inevitable that she sin on any given occasion.[33]

(I return to Franks' evaluation of (1) in section 5 below, where I argue that it's too strong of a way to understand the inevitability claim.) Claim (1) involves what David Lewis calls 'a necessity operator':

> an operator that acts like a restricted universal quantifier over possible worlds. Necessity of a certain sort is truth at all possible worlds that satisfy a certain restriction. We call these words *accessible*, meaning thereby simply that they satisfy the restriction associated with the sort of necessity under consideration. Necessity is truth at all accessible worlds, and different sorts of necessity correspond to different accessibility restrictions. (Lewis 1973, 4–5)

Necessity operators involve an accessibility relation, where that relation serves to restrict quantification over possible worlds in giving the truth conditions for that operator. So, for a necessity operator $\Box$, any possible world i, and a proposition φ, the proposition $\Box\varphi$ is true at world i *iff* for every possible world j such that j is accessible from i, φ is true in j.

Lewis specifies an inevitability relationship as one kind of necessity operator, one involving a temporally restricted accessibility relation:

> Corresponding to a kind of time-dependent necessity we may call inevitability at time t, and its strict conditional, we assign to each world i as its sphere of accessibility the set of all worlds that are exactly like i at all times up to time t, so $\Box(\varphi \supset \psi)$ is true at i if and only if ψ is true at all $\Box\varphi$-worlds that are exactly like i up to t. (Lewis 1973, 7)

The type of inevitability that Lewis has in mind here is temporally inevitability: some proposition is inevitable in a world relative to a time *iff* all the worlds that share that history of the world up to that time are worlds in which the proposition

[33] Franks 2012, 358; citations omitted. Franks' discussion focuses on those view which affirm the inevitability of sin without original guilt. But it's worth exploring whether or not even those views which do accept original guilt as part of original sin are able to account for the inevitability of sin.

in question is true.[34] But the inevitability of sin isn't inevitable given the history of a world up unto some time *t*; rather, it's inevitable given certain facts about the human condition. The inevitability of a human sinning follows given that they suffer from original sin as that doctrine has historically been understood. Sin will be inevitable for a particular human, *H*, in a world *i* if its sphere of accessibility includes no worlds in which *H* freely avoids sinning.[35] We thus look to not only what *H does* in *i*, but what *H would do* in other nearby possible worlds. Inevitability is thus a counterfactual notion.

Lewis understands counterfactuals in terms of overall similarity of worlds. Suppose ~*p* is true of the actual world, α; under what conditions would it be true that either *q* or ~*q* would have been true if *p* had been true?

$$p \mathrel{\square\!\!\rightarrow} q$$

is true *iff* the set of all worlds closest to α in which *p* is true, $S_{(p, \alpha)}$, are worlds in which *q* is true. Similarly,

$$p \mathrel{\square\!\!\rightarrow} \sim q$$

true *iff* the set of all worlds closest to α in which *p* is true, $S_{(p, \alpha)}$, are worlds in which *q* is false.[36]

We can then understand the claim that it is inevitable that a particular human *P* sins given that they suffer the penalty to human nature because of the primal sin as follows:

[34] Lewis puts most of *Counterfactuals* in terms of sentences being true or false, though he also talks about propositions being true or false (Lewis 1973, 46-47). Following Plantinga (1974, chapter 4), I'll speak in terms of propositions rather than sentences.

[35] I put the point this way in terms of there being no accessible worlds in which *H* freely refrains from sinning rather than only worlds in which *H* sins for reasons that will become clearer as we go, but the basic point is this: which other considerations are allowed to help determine the accessibility relation?

[36] Here I am making the Limit Assumption which Lewis himself did not make, but that's for the sake of simplicity of presentation. I also do not want to suggest that the principle of conditional middle holds in general for Lewis's view of counterfactuals. Putting the point in this way is, again, done for the sake of simplicity of application to the inevitability of sin.

> It is inevitable that *P* sins in *i* given that they suffer from original sin *iff* the set of all worlds most like *i*, $S_{(oP,\ i)}$, are worlds in which *P* commits at least one sinful action if they commit free actions at all.[37]

Likewise, it will be inevitable that some particular human *P* sins in the actual world as follows.

> It is inevitable that *P* sins in α as a result of original sin *iff P* suffers from original sin (that is, *oP*) in α and the set of all worlds most like α, $S_{(oP,\ \alpha)}$, are worlds in which *P* commits at least one sinful action if they commit free actions at all.

We can then understand the claim that all humans in the actual world suffer from original sin then as follows:

> It is inevitable that every human in α sins as a result original sin *iff* every human in α suffers from original sin (that is, *oH*) in α and the set of all worlds most like α, $S_{(oH,\ \alpha)}$, are worlds in which the humans in those worlds commit at least one sinful action if they commit free actions at all.

So the accessibility relation involves holding fixed the people and original sin. But what else? Different views specific different thing. But the various views can be comparted in terms of what else needs to held fixed in comparing overall similarity of worlds, and what constrains the accessibility relation that the modal operator is understood in light of. (The accessibility relationship here cannot be one of logical necessity or physical necessity.)

Lewis admits that world similarity is vague, and that it depends on, among other things, context. But, he makes clear, "not anything goes . . . There is a rough consensus about the importance of respects of comparison, and hence about comparative similarity."[38] Let us refer to those factors that we hold consistent to evaluate worlds for their similarity 'the comparison base'. Given the vagueness of the similarity relationship (and thus the coordinated vagueness of counterfactuals), we may not be able to fully specify all and only the things that determine similarity are explicit. We can assume, for instance, that a discussion of the inevitability of sin ought to hold fixed the individuals in mind, as well as their suffering from original sin. So we will take these two factors to be part of the comparison base. What else

[37] Since I am assuming facts about *P* across all the worlds in $S_{(oP,\ i)}$, this is related to Lewis' discussion of centering. For more on centering, see Lewis 1979 and Liao 2012. Accepting centering means that the actual world is closest to itself, though that doesn't help us regarding counterfactuals.

[38] Lewis 1973, 93-4.

should be held fixed? That depends on the rest of one's theological views. We can compare theological views in terms of the different things that they take to be among the features that determine world similarity (that is, we can compare them in terms of their comparison base) for the purpose of understanding how they account for the inevitability thesis. The next three sections examine how three different theological views could understand the inevitability thesis, and what they'd take to be among the comparison base in order to account for sin's inevitability. Folks have suggested that theological views that include libertarianism about human freedom, and particularly those libertarian views that reject Molinist accounts of divine knowledge and providence, cannot provide a way of understanding the inevitability thesis. But I argue that it's not clear that they can't. More specifically, unless there is independent reason for ruling out certain ways of accounting for the accessibility relation, there is logical space for the libertarian non-Molinist to endorse the inevitability thesis.

3. Inevitability on Theological Determinism

This section explores how the theological determinist can account for the inevitability thesis, and what it holds to be crucial to the relevant accessibility relation. Assume that theological determinism is true. That is, assume that it is true that for every event, God's actively willing that particular event is both necessary and sufficient for the occurring of that event.[39] Assume also that theological compatibilism is true—that is, assume that the truth of theological determinism is compatible with human freedom and responsibility.[40] According to the first assumption, everything that happens happens exactly as God wills it to happen. Since God's willing is sufficient for the occurrence of what God wills, it is not possible for God's willing to be frustrated. And, according to the second assumption, if among the things that God wills is that God wills some human agents to act freely

[39] Derk Pereboom defines theological determinism as the view that "God is the sufficient active cause of everything in creation, whether directly or by way of secondary causes such as human agents" (Pereboom 2011, 262).

[40] Not all theological determinists are committed to theological compatibilism simply in virtue of endorsing theological determinism. But if theological determinism is true and theological compatibilism is false, then humans are neither free nor responsible (in the basic desert sense). If no humans are free or responsible, discussions of sin would have to be sufficiently revisionist that I won't consider them here (despite liking many a revisionist!). Pereboom explores Christianity without moral responsibility (see Pereboom 2005 and 2011), but I am not aware that he addresses the nature of sin on such a view.

and responsibly, then they do so. The conjunction of these assumptions establishes the claim that humans sin *iff* God wills for them to sin.

Proponents of this pair of views can account for the truth of the inevitability thesis by taking the following to be true (called 'TD1' for being the first claim regarding theological determinism):

TD1 God wills that each human person in world *i* commits at least one sin.

While this claim establishes that all humans *do* sin, it doesn't yet establish that all individuals *must* sin, as the inevitability thesis claims. Given the sufficiency of God's volition regarding events for their occurring, theological determinists are also committed to the following claim:

TD2 Necessarily, if God wills that each human person in world *i* commits at least one sin, then each human person in *i* commits at least one sin.

When TD1 and TD2 are combined, they entail the following:

TD3 Necessarily, given that God wills that each human person in world *i* commits at least one sin, then each human person in *i* commits at least one sin.

There is a sense in which TD3 expresses a claim about what humans *must* do in *i* — given the contents of God's will in TD1, then it is not the case that they can avoid sinning in that world. And this seems sufficient to establish the inevitability thesis.

One might object that this isn't enough for the inevitability thesis given that it is conditional upon God's volition. That is, hasn't yet been established for theological determinism is the following necessary claim (NC):

NC Necessarily, all humans must commit at least one sin.

Christianity has typically held that not only that *what* God chooses to create, including which if any humans are part of that creation, is contingent, but the fact that God chooses to create *at all* is itself contingent. So at the very least, NC needs to be restricted to those worlds in which God has chosen to create, and more specifically those worlds in which God has chosen to create free human creatures. Given divine freedom as it relates to *whether* and *what* to create, we can specify the restriction to such worlds as follows:

NC1 Necessarily, for every world *i* which contains free humans, all those humans commit at least one sin.

But NC1 is too strong, since on it sin becomes an essential property of every free human person. What is needed is a way of restricting the relevant worlds such that sin is inevitable for every human in that world without it being true that they sin in every world in which they exist. Earlier we said that what was needed to secure the inevitability thesis was for the set of accessible worlds to be worlds in which the humans in those worlds commit at least one sinful action if they commit free actions at all. Since, on this theological view, God's willing that the world be *such and so* is necessary and sufficient that it be *such and so*, we can restrict the accessible worlds to those which include the divine volition indicated in TD1. (That is, we can say that God's volition that each human person in that world commits at least one sin is part of the comparison base by which we determine which worlds are accessible to *i*.) Had God's volition indicated in TD1 been different, then it would not have been either actual or inevitable that each person in *i* would sin. So even on the strongest account of divine providence, the conjunction of theological determinism and compatibilism, the inevitability of sin is accounted for by restricting the comparison base in a particular way.

The theological determinist must give an account, of course, for why TD1 would be true given its implications for the problem of evil. Why would God will in such a way as TD1 claims? But the proponent of theological determinism is already committed to there being a reason for why God wills all the evils that occur,[41] and thus I don't see that this is a further difficulty for the view rather than another aspect of a difficulty that's already acknowledged.[42]

4. Inevitability and Molinism

I now turn toward exploring the inevitability thesis according to Molinism. Particularly since the publication of Alvin Plantinga's *God, Freedom, and Evil* and Thomas Flint's *Divine Providence: The Molinist Account*, Molinism has come to play a very prominent role in contemporary analytic philosophy of religion. Molinists defend a robust account of divine providence on the basis of what is referred to as

[41] For attempts to defang the 'author of sin' objection against theological determinism, see Welty 2016, White 2016, and Bignon 2018. It is not clear to me that any of these attempts are ultimately successful.

[42] See Pereboom 2005 and Mann 1988 for discussions.

God's 'middle knowledge.' Molinists differentiate different aspects of God's knowledge with an ordering holding between those different aspects.[43] First, God's natural knowledge is God's pre-volitional knowledge of necessary truths. For example, God knows that it is a necessary truth that all parts of creation depend upon God for their creation and conservation, even prior to or apart from the divine decision to create anything. God's free knowledge, on the other hand, is God's post-volitional knowledge of contingent truths; God can know that it is a contingent truth that Oliver the sassy sheepadoodle does in fact exist, but only logically posterior to His volition to bring creation, including canine crossbreeds, into existence and sustaining at least one such dog in existence. Most accounts of the nature of divine knowledge include both natural and free knowledge. What is unique to Molinism is its holding that in addition to natural and free knowledge, God also has middle knowledge—a kind of knowledge that is 'in the middle of' or between God's natural and God's free knowledge. Like God's natural knowledge, God's middle knowledge is pre-volitional. But, on the other hand, like free knowledge it is knowledge of contingent truths. In terms of providence, the most important objects of God's middle knowledge are what are called 'counterfactuals of creaturely freedom.'[44] A counterfactual of creaturely freedom (CCF for short) is a contingent proposition about how a creature would freely act in a particular situation. Such propositions have the following form:

> If agent *A* were in circumstances *C*, *A* would freely do *X*.

The 'freely' in CCFs should be understood as involving libertarianism; *A* freely does *X* only if *A*'s doing *X* is neither causally nor theologically determined. At the heart of the Molinist's view, God pre-volitionally knows what all possible created persons would freely do in every possible circumstance by knowing what CCFs are (contingently) true and which ones are (contingently) false. God can only actualize those possible worlds that are consistent with the set of true CCFs. Those worlds which are not consistent with the set of true CCFs are possible (since the set of true CCFs could have been different), but they are not, in Thomas Flint's terminology, feasible.[45]

Folks who endorse Molinism can also endorse the inevitability thesis by taking advantage of what Al Plantinga has labeled 'transworld depravity':

---

[43] This is a logical, not temporal, ordering. See Flint 1998 and Hasker 2009, 334.

[44] Flint 1998; since not every CCF is in fact a counterfactual, it is perhaps better to call them 'free will subjunctive conditionals' as Mann 1988 does, 50.

[45] Flint 1998, 51.

> A person *P suffers from transworld depravity* if and only if the following holds: for every [possible] world *W* such that *P* is significantly free in *W* and *P* does only what is right in *W*, there is an action *A* and a maximal world segment *S′* such that
>
> (1) *S′* includes *A*'s being morally significant for *P*
> (2) *S′* includes *P's* being free with respect to *A*
> (3) *S′* is included in *W* and includes neither *P*'s performing *A* nor *P*'s refraining from performing *A*
>
> and
>
> (4) If *S′* were actual, *P* would go wrong with respect to *A*.[46]

Bruce Langtry gives the following as a loose approximation of transworld depravity: person *P* suffers from transworld depravity if and only if, as a matter of contingent fact, if God were to create *P* and *P* is free, then whatever else God were to allow regarding creation, *P* would go wrong with at least one morally significant action.[47] That is, if a person, Paul, suffers from transworld depravity, then no matter which possible world God actualizes[48] such that Paul exists in that world and is free, there is a true CCF regarding Paul such that it is true that Paul freely commits at least one sin (e.g., endorses misogynistic expectations) in that possible world. No possible world in which Paul exists, is free, and there is no sin is a feasible world. If Paul suffers from transworld depravity, then there is a sense in which he will sin in that world. And if all possible humans suffer from transworld depravity, then it's the

---

[46] Plantinga 1977, 48. He restates TWD on 52 (and on Plantinga 1974, 188) in terms of essences rather than persons, where persons are instantiations of essences, which are understood as abstract objects. This modification need not concern us at present.

[47] Bruce Langtry 2010, 145. Langtry's actual formulation of TWD is about essences rather than persons since "Plantinga does not believe that there are non-actual persons, yet the Defence [i.e., Plantinga's free will defense] requires consideration of both whether actual persons could suffer from transworld depravity and also whether God could have created other people instead, who did not suffer from transworld depravity. Talk of essences provides a way of discussing both questions" (145). For simplicity's sake, I'll focus only on persons and not essences, though those who prefer not to admit the existence of non-actual persons as abstract objects are welcome to interpret what follows using essences instead. Both Langtry and Richard Otte have taken issue with Plantinga's formulation of transworld depravity; Plantinga agrees that Otte has shown his initial treatments of TWD were problematic in Plantinga, 2009. Nevertheless, in what follows I set these issues aside, as the central point I want to make would still follow with an updated account of transworld depravity. For a discussion of how Plantinga's view of transworld depravity relates to his '*o felix culpa*' theodicy, see Davis and Franks 2018.

[48] Unless the circumstances in which Paul freely acts (and sins) are causally isolated from earlier free choices, either by Paul or by other free creatures, the use of 'actualizes' here will be 'weakly actualizes'; see Plantinga 1974, 172.

case that if God creates a world containing free humans, then they will sin. That is, no world containing free humans and yet no sin would be among the set of feasible worlds. As Plantinga himself says, "The whole point of introducing TWD [is] to show how it could be that it wasn't within God's power to actualize a world containing free creatures who always do what is right; it is possible that the counterfactuals of freedom should fall out in such a way as to preclude God's doing that."[49]

Though not himself a Molinist, William Hasker argues that while the Molinist can account for the inevitability of sin using transworld depravity, that view would actually entail something stronger: "it seems that anyone who suffers from it [i.e., transworld depravity] must of necessity go wrong morally on *the very first occasion* on which that person makes a choice between moral good and evil."[50] We can understand Hasker as here endorsing the following necessary clam:

> NC2 Necessarily, if a human person suffers from transworld depravity, then they must freely choose evil on the very first occasion of making a choice between moral good and moral evil.

Why does Hasker think this follows from transworld depravity? His reasoning is as follows:

> Suppose this is not so [i.e., assuming that a human who suffers from transworld depravity does not sin on the very first occasion when they make a choice between moral good and evil]; suppose someone who suffers from transworld depravity makes her very first moral choice a choice for the good. This poses no problem, you might think; that person will have plenty of additional chances to go wrong. But suppose she doesn't have any more chances? It seems entirely possible that, having made one free choice for the good, the person might have no further opportunities at all: perhaps her existence comes to an end at that point or circumstances change in some way that she never again chooses between moral right and world. If this were to occur, then the person would have been given freedom to choose between good and evil, and yet would never have chosen to do evil—but that would mean that, contrary to our supposition, the person was not trans worldly depraved after all. Indeed, if this scenario is even *possible*, it follows that she is not transwordly depraved, for transworld depravity means that *in any possible set of circumstances* in which the person is free to choose between good and evil (*including* the scenario in which the person makes only one morally significant choice in her lifetime), she

[49] Plantinga 2009, 182.

[50] Hasker 2008, 61.

> chooses evil at least once. So if the person is in fact transwordly depraved, she will choose evil on the very first opportunity that presents itself to her.[51]

This, however, doesn't follow. The Molinist's defense of the inevitability of sin claims only that in each feasible world in which a human person exists, that human finds themselves in at least one situation in which they freely sin. But what is feasible depends on the set of true CCFs. Molinism does not commit one to the claim that every world in which the agent is in that same situation is one in which the CCFs are the same. The possible world in which the agent in question dies immediately after not sinning at their first possible opportunity and the possible world in which they never do are distinct, and thus may have different true counterfactuals of creaturely freedom. At the very least, the Molinist would need to give an account of why the counterfactuals of creaturely freedom align in such as a way as Hasker presupposes.

As we saw in the previous section, saying that all humans will sin doesn't yet establish the inevitability thesis which says that they *must* sin. But, as we also saw, it's asking too much to establish the necessity of all humans sinning. The inevitability of sin only requires that the comparison base for the similarity between worlds contains something such that all the accessible worlds share that feature. But assume that it is contingently true in i that all possible human persons suffer from TWD, and that this truth is part of the comparison base. It would follow that not only every feasible world containing free humans but also every accessible world in $S_{(oH,\ i)}$ is a world in which each human commits at least one sin in that world. We've already seen that it is appropriate, when seeking to account for the inevitability of sin in i, to restrict accessibility to the set of worlds $S_{(p,\ i)}$ that share a contingent truth p (here, the conjunction of the relevant counterfactuals that go into establishing TWD). On this view too then, like with theological determinism, Molinism can establish the inevitability of sin with the additional assumption of transworld depravity by such an account of the accessibility relation.[52]

The Molinist's defense of the inevitability thesis will only work if both of its driving assumptions, namely the truth of Molinism and the thesis of transworld depravity, are true. Unfortunately, I don't think that either is. I've indicated my two primary objections to Molinism elsewhere.[53] First, I think that a version of the grounding objection against counterfactuals of creaturely freedom works. God is not

---

[51] Hasker 2008, 61

[52] See also the discussions in Rea 2007, Anderson 2021, and Loke 2022 chapter 6 for other Molinist-based proposals.

[53] Timpe 2018.

the ground of the true counterfactuals of creaturely freedom, nor are the agents that the counterfactuals are about, since their truth is independent of the agents' existence. As Richard Gale puts it regarding what makes the CCFs true: "As they used to say in the Bronx, 'Don't ask!' Here's where the regress of explanations hits the brick wall of brute, unexplainable contingency. There are no further elephants or tortoises upon whose back this contingency rests."[54] I think brute counterfactuals of creaturely freedom cannot bear the weight needed for Molinism. But second, even if they could, one can run an argument parallel to the Consequence Argument for incompatibilism against Molinism.[55] And if this is right, then Molinism can't secure the high level of providential control that it seeks to while still maintaining libertarianism. Even if I were inclined to think that Molinism was true, I'd agree with Josh Rasmussen that the probability of trans-world depravity is sufficiently low that we ought to reject it—or at least look for accounts of theological doctrines that don't require TWD for their plausibility.[56]

5. Libertarians on Inevitability

But can a libertarian who rejects Molinism still endorse the inevitability thesis? W. Paul Franks suggests the answer is 'no'. As mentioned in section 2, Franks understands the inevitability thesis as follows:

> (1) Necessarily in a world tainted by original sin, (a) every human subsequent to Adam and Eve is born in a condition such that it is inevitable that she sin (given that she performs at least one morally significant action), but (b) it is not inevitable that she sin on any given occasion.[57]

Franks raises (1) as an objection to those views that endorse only what he calls original inclination and not original guilt, where the latter is "one's being born guilty" and the former is "one's being born in a condition that inevitably leads to

[54] Gale 2007, 55. See also McCann 2011 and Craig 2001.

[55] For a further development of this concern, see Climenhaga and Rubio 2021.

[56] Rasmussen 2004. More specifically, he thinks that the probability of there being no feasible evil-free worlds is 0. One need not think it's that low to think that it's sufficiently low that we ought not endorse it. For another criticism of TWD and its role in Plantinga's free will defense, see Howard-Snyder and O'Leary-Hawthorn 1998.

[57] Franks 2012, 358; citations omitted. Franks' discussion focuses on those view which affirm the inevitability of sin without original guilt. But it's worth exploring whether or not even those views which do accept original guilt as part of original sin are able to account for the inevitability of sin.

sin."[58] Franks is a libertarian, and as a result thinks that humans are free and thus cannot be determined to always act as they do. But the doctrine of original guilt, Franks claims, is difficult to reconcile with libertarianism, since it seems to require one to be blameworthy for things over which one has no control.[59] He thus restricts his attention to original inclination and argues that (1) false. To see why, he asks us to consider a possible world (much like the worlds that Hasker considered, but without assuming the truth of Molinism) in which a human performs only one morally significant action in their lifetime. The existence of such a world would mean that:

(2) Possibly, some human performs only one morally significant action in her lifetime.[60]

Franks thinks that the truth of (2) "seems assured" and takes it to be true that there is "no reason one could give for taking it to be false."[61] However, according to (1a) above, it is true of that individual that she inevitably sins in that particular action. That is, it follows from (1a) and (2) that

(3) If some human performs only one morally significant action in her lifetime, then that action is inevitably sinful.[62]

However, from (1b) it follows that:

(4) If some human performs only one morally significant action in her lifetime, then that action is not inevitably sinful.[63]

(3) and (4) together entail a contradiction. Thus Franks thinks that one can't be a libertarian who endorses original inclination without original guilt and endorse (1a). He rejects that the constitutional fault of original sin ensures sinful actions: "we are

[58] Franks 2012, 355.

[59] Loke agrees: "the view that 'All human beings (except, at most, four) are guilty from birth in the eyes of God, and this guilt is a consequence of the first sin of the first [human]' (Original Guilt) is unwarranted and is contrary to the principle of justice that we should not be judged for what we are not responsible for" (Loke 2022, 132). It's not clear that libertarians cannot give an account of participation in the first human sin; see Rea 2007 and the discussion in Crisp 2019.

[60] Franks 2012, 359.

[61] Franks 2012, 359 and 363.

[62] Franks 2012, 359.

[63] Franks 2012, 359.

influenced by a fallen world to sin, but are free at any point to refrain from sinning."[64] Sin thus is perhaps very likely, but not strictly speaking inevitable. And Franks is not the only one to think that the libertarian non-Molinist cannot account for the inevitability of sin. Richard Swinburne also rejects that, as a result of original sin, sin is inevitability given his commitment to a libertarian view, which he sees as requiring a commitment to the view that a free creature cannot be responsible for an unavoidable or inevitable action:

> Adam's responsibility for our sinfulness is confined to a responsibility for beginning the social transmission of morality (as such a good thing) which made sin possible, but a morality which, as a result of his own sinful example and perhaps false moral beliefs, was no doubt a corrupt morality and so made it easier for our genetically inherited proneness to sin to work in Adam's successors.[65]

Swinburne thus thinks that sinful acts are "almost unavoidable"[66] but not inevitable, as expressed by Franks' (1). And theologian McFarland endorses a compatibilist view of human freedom in part because he doesn't think that the libertarian (presumably who also rejects Molinism) can account for the inevitability of sin.[67]

There are two ways that one could argue that libertarian non-Molinism could go to avoid the claimed inconsistency. The first way is to recognize that not all libertarians are committed to (1b) insofar as they think that one can be free and responsible for an action even if it's inevitable so long as they played the right role in it's becoming inevitable. Source libertarians are not committed to the claim that every action for which one is morally responsible is one for which the agent must have alternative possibilities. Consider, for instance, Dean Zimmerman's 'virtue libertarianism':

> The highest such good [that libertarian freedom is necessary for] is the very possibility of creatures capable of displaying moral virtues—hard-won habits due, at least in part, to a lifetime of free choices. Choices made because of a genuinely moral virtue (as opposed to a merely excellent disposition) redound to the credit of the agent even when the virtue is so ingrained to make the choice, now, inevitable. Likewise, *mutatis mutandis*, for choices rendered inevitable by vices. Strictly speaking, such choices and action are not freely taken—i.e. they are not examples of the base case of indeterministic, free choosing. Still, if an agent is an uncoerced

[64] Franks 2012, 370.
[65] Swinburne 1989, 143.
[66] Swinburne 1989, 146.
[67] McFarland 2016, 311-317.

> expression of character traits for which one is responsible, the action is one for which we should hold a person responsible. We might even want to say that such actions and choices are "freely undertaken" in the sense that they are the expression of a character formed by a history of freely chosen action, despite the fact that the agents no longer have a choice about their behavior in these circumstances. Still, libertarians will think actions free in this broader sense could not occur without, somewhere down the line, free choices in the narrower sense—instances of the base case of freedom, requiring indeterministic circumstances.[68]

How might an agent be responsible for the first human sin that makes inevitable for them further acts of sin? Realists, as opposed to federalists, about original sin hold that there is a real unity between all who are justly punished for original guilt and the first human sinner. As Augustine puts it,

> The first human beings . . . having become the first sinners, were then punished by death in such a way that whatsoever sprang from their stock should also be subject to the same penalty. For nothing could be born of them which was not what they themselves had been . . . so that what arose as a punishment in the first human beings who sinned also follows as a natural consequence in the rest who are born of them.[69]

The best-known elaboration of realism comes from Jonathan Edwards, who held that all humans are one simply because as God declares us to be: "there is no identity or oneness [that does not] depend on the *arbitrary* constitution of the Creator . . . *Divine constitution* [God's treating all humans as one entity] is the thing which *makes truth*."[70] Influenced by Edwards, a number of philosophers have considered the possibility of perdurantism being used to provide a basis for realism.[71]

---

[68] Zimmerman 2012, 176-7. On my preferred view, we can think of those actions that are the inevitable results of the moral character an agent has freely formed as both free *and* responsible actions; but I don't think that anything of substance here hangs on this difference between Zimmerman's and my view.

[69] Augustine *City of God*, XIII.iii.

[70] Edwards 1758: part 4, ch. III [1970: 404].

[71] Perdurantism is, roughly, the view that ordinary objects like humans persist through time in virtue of having temporal parts. See the discussions in Wainwright 1988, Wyma 2004, Crisp 2005 and 2009, Rea 2007, and Hudson 2014. Anderson thinks that the "creative metaphysical wrangling" needed to get participation of all humans in the first human sin would be a "hard sell" (Anderson 2021, 8 and 9). It may be, however, that such humans on such a view might not satisfy the epistemic condition on moral responsibility, even if it can be shown that they satisfy the control condition (that is, even if they have free will). See Timpe 2011 for a discussion of the difficulties spelling out the epistemic condition even apart from considerations of original sin.

Rather than going into the details of how such a view would look, notice from the quotation from Edwards that realism is dependent upon divine willing. If God had made a different decision regarding identity, then not all humans would be responsible for the first sin that formed their character in such a way that later sins were inevitable. So, as was the case with theological determinism in section three above, God's willing that the world be a certain way (here, that the relevant personal identity relation holds) restricts the accessible worlds as part of the relevant comparison base. That is, we can say that God's volition that Sinner Sal in *i* be identical with the person who committed the first sin in *i* is part of the comparison base by which we determine which worlds are accessible to *i*. Had God's volition indicated in *i* been different, then it would not have been either actual or inevitable that Sinner Sal would sin, since then the set of all worlds most like *i* would have been different. The libertarian non-Molinist is no more committed than the theological determinism to the claim that a person must sin in every possible world in which they suffer from original sin, only that they sin in all the accessible worlds. That is, Franks' (1a) need not be interpreted as making a claim about all possible world's in which humans suffer from original sin, but only those worlds that are most like *i*. Sin is inevitable for a particular human in a world *i* if that worlds' accessibility relation includes no worlds in which they freely avoid sinning.

This then leads us to the second possibility for avoiding Franks' claimed inconsistency for the libertarian non-Molinist. I've suggested that for the theological determinist, the Molinist, and the libertarian non-Molinist who endorses realism, the set of accessible worlds relevant to the claim that it is inevitable that a human sin can be restricted by what I called the comparison base. For the theological determinist, the comparison base includes the divine volitions regarding the human in question. For the Molinist, it includes the set of true CCFs that entail that the human suffers from TWD. For the libertarian non-Molinist who endorses realism, it includes divine volitions regarding personal identity. It would thus be inappropriate to require of the libertarian non-Molinist more generally to establish that some human *P* inevitably sins in α as a result of original sin *iff P* suffers from original sin in α and all the worlds in which *P* exists and suffers from original sin are worlds in which *P* commits at least one sinful action if they commit free actions at all. It only requires that the set of all worlds most like α, *S (P,* α), are worlds in which *P* commits at least one sinful action if they commit free actions at all. So we do not yet have reason to think that the libertarian non-Molinist cannot give an account of inevitability involving a different comparison base.

Of course, to develop this possibility beyond a mere framework, the libertarian non-Molinist would need to give an account of just what features ought to be in the

comparison base that would restrict the accessibility relation in this way. Why are the possible worlds in which the agent does sin more relevant for evaluating the counterfactuals than worlds in which they don't? But, as the discussions of theological determinism and Molinism make clear, there might be reasons to think that other elements of one's theological package should be part of the relevant comparison base. To avoid being *ad hoc*, the comparison base would need to be plausible. But, as I discuss elsewhere, we need to make judgements about the philosophical commitments of our theological packages in a wholistic way, and judgements about plausibility are in part a function of other positions that we're antecedently committed to.[72] For those Christian philosophers that take Church tradition to be a source of evidence regarding theological and philosophical views, the Church's historical commitment to the inevitability thesis carries at least some, and perhaps significant, epistemic weight. It may even give us theological reasons to favor certain ways of determining the accessibility relation even if there are not decisive philosophical arguments on offer.

Conclusion

In the previous pages, I've tried to give an account of the historical Christian theological claim that given the impact of original sin, acts of sin are unavoidable for human agents. I've shown how a number of theological traditions can understand that inevitability thesis. I've also sought to show that despite the claim that libertarian non-Molinist views cannot account for sin's inevitability, there are philosophical resources that can be brought to bear in giving a framework for how they can. Ian McFarland writes that "original sin is a derivative doctrine: it is deduced from the more fundamental Christian claims that Christ is the Saviour of all, and that that all need to be saved."[73] A commitment to this fundamental Christian claim gives us reasons for further exploring the possibilities regarding how best to understand the inevitability of sin.[74]

[72] See the discussion in Timpe forthcoming, particularly section 3, for considerations relevant to this paragraph.

[73] McFarland 2016, 311.

[74] Thanks to Christa McKirland, Chris Menzel, and two anonymous reviewers for *TheoLogica* for their helpful comments on previous drafts of this paper.

Bibliography

Anderson, Mark B. 2021. "On Responsibility and Original Sin: A Molinist Suggestion." *Faith and Philosophy* 38.1: 5–25. https://doi.org/10.37977/faithphil.2021.38.1.2.

Augustine. 1998. *The City of God Against the Pagans* (*De civitate Dei*). Translated and edited by R.W. Dyson. Cambridge: Cambridge University Press, 1998. https://doi.org/10.1017/CBO9780511802300.

Bignon, Guillaume. 2018. *Excusing Sinners and Blaming God*. Eugene, OR: Pickwick Publications.

Blocher, Henri. 1997. *Original Sin: Illuminating the Riddle*. Downer's Grove, IL: InterVarsity Press.

Book of Common Prayer. 1979. New York, Church Publishing.

Catechism of the Catholic Church, 2nd ed. 2003. New York: Doubleday.

Climenhaga, Nevin and Daniel Rubio. 2021. "Molinism: Explaining our Freedom Away," *Nous*. https://doi.org/10.1093/mind/fzab042.

Copan, Paul. 2003. "Original Sin and Christian Philosophy." *Philosophia Christi* 5(2): 519–541. https://doi.org/10.5840/pc20035252.

Couenhoven, Jesse. 2013. *Stricken by Sin, Cured by Christ: Agency, Necessity, and Culpability in Augustinian Theology*, Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199948697.001.0001.

Couenhoven, Jesse. 2016. "Augustine", in *T&T Clark Companion to the Doctrine of Sin*. Eds. Johnson, Keith L. and David Lauber. New York: T&T Clark Ltd., 181–198.

Craig, William Lane. 2001. "Middle Knowledge, Truth–Makers, and the "Grounding Objection." *Faith and Philosophy* 18(3): 337–352. https://doi.org/10.5840/faithphil200118329.

Crisp, Oliver D. 2005. *Jonathan Edwards and the Metaphysics of Sin*. New York: Routledge. https://doi.org/10.4324/9781315251387.

Crisp, Oliver D. 2009. "Original Sin and Atonement." In *The Oxford Handbook of Philosophical Theology*. Eds. Thomas P. Flint and Michael C. Rea. Oxford: Oxford University Press, 430–451. https://doi.org/10.1093/oxfordhb/9780199596539.013.0020.

Crisp, Oliver D. 2015. "On Original Sin." *International Journal of Systematic Theology*, 17(3): 252–66. https://doi.org/10.1111/ijst.12107.

Crisp, Oliver D. 2019. *Analyzing Doctrine: Toward a Systematic Theology*. Waco, TX: Baylor University Press.

Crisp, Oliver D. 2020. "A Moderate Reformed View." In *Original Sin and the Fall: Five Views*. Eds. J.B. Stump and Chad Meister. Downer's Grove, IL: IVP Academic, 35–54.

Davis, Richard Brian and W. Paul Franks. 2018. "Plantinga's Defence and His Theodicy are Incompatible." In *Does God Matter? Essays on the Axiological Consequences of Theism*, ed. Klaas Kraay, New York: Routledge, 203–223. https://doi.org/10.4324/9781315210995-11.

Decrees of the Ecumenical Councils. 1990. Ed. Norman Tanner. Washington, DC: Georgetown University Press.

Dockter, Cornelia. 2021. "God's Work and Human's Contribution." In *Impeccability and Temptation*. Ed. Johannes Grössl and Klaus von Stosch. London: Routledge, 75–93. https://doi.org/10.4324/9780367618933-6.

Edwards, Jonathan. 1758 [1970]. *The Works of Jonathan Edwards*. Ed. Clyde A. Holbrook, New Haven, CT: Yale University Press.

Flint, Thomas P. 1998. *Divine Providence: The Molinist Account*. Ithaca, NY: Cornell University Press.

Franks, W. Paul. 2012. "Original Sin and Broad Free-Will Defense." *Philosophia Christi* 14(2): 353–371. https://doi.org/10.5840/pc201214230.

Gale, Richard M. 2007. "Evil and Alvin Plantinga." In *Alvin Plantinga*. Ed. Deane-Peter Baker. Cambridge University Press: New York, 48–70. https://doi.org/10.1017/CBO9780511611247.003.

Green, Adam. 2022. "The Maturational Naturalness of Original Sin." *TheoLogica* 6(1): 20–43. https://doi.org/10.14428/thl.v6i1.61273.

Hasker, William. 2008. *The Triumph of God over Evil: Theodicy for a World of Suffering*. Downer's Grove, IL: InterVarsity Press.

Hasker, William. 2009. "The Openness of God." In *Arguing about Religion*. Ed. Kevin Timpe. New York: Routlege, 349–61.

Howard-Snyder, Daniel and John O-Leary-Hawthorne. 1998. "Transworld Sanctity and Plantinga's Free Will Defense." *International Journal for Philosophy of Religion* 44: 1–21. https://doi.org/10.1023/A:1003210017171.

Hudson, Hud. 2014. *The Fall and Hypertime*, Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198712695.001.0001.

Langtry, Bruce. 2010. "The Prospects for The Free Will Defence." *Faith and Philosophy* 27(2): 142–152. https://doi.org/10.5840/faithphil201027214.

Lewis, David K. 1973. *Counterfactuals*. Malden, MA: Blackwell.

Lewis, David K. 1979. "Attitudes *De Dicto* and *De Se*." *The Philosophical Review* 88: 513–543. https://doi.org/10.2307/2184843.

Liao, Shen-yi. 2012. "What are Centered Worlds?" *The Philosophical Quarterly* 62(247): 294–316. https://doi.org/10.1111/j.1467-9213.2011.00042.x.

Loke, Andrew Ter Ern. 2022. *Evil, Sin, and Christian Theism*. New York: Routledge. https://doi.org/10.4324/9781003265689.

Louth, Andrew. 2020. "An Eastern Orthodox View." In *Original Sin and the Fall: Five Views*. Eds. J.B. Stump and Chad Meister. Downer's Grove, IL: IVP Academic, 78–100.

Mann, William. 1988. "God's Freedom, Human Freedom, and God's Responsibility for Sin." In *Divine and Human Actions*. Ed. Thomas Morris. Ithaca, NY: Cornell University Press, 182–210. https://doi.org/10.7591/9781501746123-007.

Mann, William. 2001. "Augustine on Evil and Original Sin." In *The Cambridge Companion to Augustine*. Eds. Eleonore Stump and Norman Kretzmann. Cambridge: Cambridge University Press, 40–48. https://doi.org/10.1017/CCOL0521650186.004.

McCall, Thomas H. 2019. *Against God and Nature: The Doctrine of Sin*. Wheaton, IL: Crossway Books.

McCann, Hugh J. 2011. "The Free Will Defense." In *Molinism: The Contemporary Debate*. Ed. Ken Perszyk. Oxford: Oxford University Press, 239–61.

McFarland, Ian A. 2018. "Original Sin." In *T&T Clark Companion to the Doctrine of Sin*. Eds. Keith K. Johnson and David Lauber. London: T&T Clark, 303–318.

McKinley, John E. 2021. "Seven Questions Ingredient to Jesus Christ's Temptation." In *Impeccability and Temptation*. Ed. Johannes Grössl and Klaus von Stosch. London: Routledge, 119–42. https://doi.org/10.4324/9780367618933-9.

Mellema, Gregory. 2021. *Sin*. Notre Dame: University of Notre Dame Press.

Morris, Thomas V. 1992. *Making Sense of It All: Pascal and the Meaning of Life*. Grand Rapids, MI: William B. Eerdmans Publishing Company.

Otte, Richard. 2009. "Transworld Depravity and Unobtainable Worlds." *Philosophy and Phenomenological Research* 78: 165–77. https://doi.org/10.1111/j.1933-1592.2008.00236.x.

Our Faith: Ecumenical Creeds, Reformed Confessions, and Other Resources. 2013. Grand Rapids, MI: Faith Alive.

Pawl, Timothy. 2016. *In Defense of Conciliar Christology: A Philosophical Essay*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198765929.001.0001.

Pawl, Timothy. 2019, *In Defense of Extended Conciliar Christology: A Philosophical Essay*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198834144.001.0001.

Pawl, Timothy. 2021. "Conciliar Christology, Impeccability, and Temptation." In *Impeccability and Temptation*. Ed. Johannes Grössl and Klaus von Stosch. London: Routledge, 94–115. https://doi.org/10.4324/9780367618933-7.

Pereboom, Derk. 2005. "Free Will, Evil, and Divine Providence." In *God and the Ethics of Belief: New Essays in Philosophy of Religion*. Eds. Andrew Dole and Andrew Chignell. Cambridge: Cambridge University Press, 77–98. https://doi.org/10.1017/CBO9780511499166.004.

Pereboom, Derk. 2011. "Theological Determinism and Divine Providence." In *Molinism: The Contemporary Debate*. Ed. Ken Perszyk. Oxford: Oxford University Press, 262–280. https://doi.org/10.1093/acprof:oso/9780199590629.003.0017.

Plantinga, Alvin. 1974. *The Nature of Necessity*, Oxford: Oxford University Press. https://doi.org/10.1093/0198244142.001.0001.

Plantinga, Alvin. 1977, *God, Freedom, and Evil*. Grand Rapids, MI: Eerdmans.

Plantinga, Alvin. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press. https://doi.org/10.1093/0195131932.001.0001.

Plantinga, Alvin. 2009. "Transworld Depravity, Transworld Sanctity, & Uncooperative Essences" *Philosophy and Phenomenological Research* Vol. 78.1: 178–191. https://doi.org/10.1111/j.1933-1592.2008.00237.x.

Quinn, Philip L. 1997. "Sin and Original Sin." In *A Companion to Philosophy of Religion*. Ed. Philip L. Quinn, and Charles Taliaferro. Malden, MA: Blackwell, 541–548. https://doi.org/10.1111/b.9780631213284.1999.00071.x.

Rasmussen, Josh. 2004. "On Creating Worlds without Evil-Given Divine Counterfactual Knowledge." *Religious Studies* 40: 457–470. https://doi.org/10.1017/S0034412504007255.

Rea, Michael C. 2007. "The Metaphysics of Original Sin." In *Persons: Human and Divine*. Eds. Peter van Inwagen and Dean Zimmerman. Oxford: Oxford University Press, 319–356.

Sumner, D.O. 2014. "Fallenness and Anhypostatsis: A Way Forward in the Debate Over Christ's Humanity." *Scottish Journal of Philsoophy* 67(2): 195–212. https://doi.org/10.1017/S0036930614000064.

Swinburne, Richard. 1989. *Responsibility and Atonement*, Oxford: Clarendon Press. https://doi.org/10.1093/0198248490.001.0001.

Timpe, Kevin. 2011. "Tracing and the Epistemic Condition on Moral Responsibility." *The Modern Schoolman* 88(1/2): 5–28. https://doi.org/10.5840/schoolman2011881/22.

Timpe, Kevin. 2014. *Free Will in Philosophical Theology*. London: Bloomsbury.

Timpe, Kevin. 2018. "'Upright, Whole, and Free'—Eschatological Union with God." *TheoLogica* 2(2): 1–16. https://doi.org/10.14428/thl.v2i2.2123.

Timpe, Kevin. 2021. "Sin in Christian Thought." *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition). Ed. Edward N. Zalta. https://plato.stanford.edu/archives/win2021/entries/sin-christian/.

Timpe, Kevin. Forthcoming. "Weighing Compatibilism and Libertarianism in Analytic Theology." In *Theological Perspectives on Free Will: Compatibility, Community, Christology*. Eds. Olli-Pekka Vainio and Aku Visala New York: Routledge.

van Inwagen, Peter. 2006. *The Problem of Evil*. Oxford: Oxford University Press. 2006. https://doi.org/10.1093/acprof:oso/9780199245604.001.0001.

Wainwright, William J. 1988. "Original Sin." In *Philosophy and the Christian Faith*. Ed. Thomas V. Morris. Notre Dame, IN: University of Notre Dame Press, 31–60.

Welty, Greg. 2016. "Molinist Gunslingers: God and the Authorship of Sin." In David E. Alexander and Daniel M. Johnson, *Calvinism and the Problem of Evil*. Eugene, OR: Pickwick Publications, 56–77.

White, Heath. 2016. "Theological Determinism and the 'Authoring Sin' Objection." In *Calvinism and the Problem of Evil*. Eds. David E. Alexander and Daniel M. Johnson. Eugene, OR: Pickwick Publications, 78–95.

Wyma, Keith D. 2004. "Innocent Sinfulness, Guilty Sin: Original Sin and Divine Justice." In *Christian Faith and the Problem of Evil*. Ed. Peter van Inwagen 2004: Grand Rapids, MI: William B. Eerdmans, 263–276.

Zimmerman, Dean. 2012. "An Anti-Molinist Replies." In *Molinism: The Contemporary Debate*. Ed. Ken Perszyk. New York: Oxford University Press, 163–86. https://doi.org/10.1093/acprof:oso/9780199590629.003.0012.

Published Online First: November 30, 2022

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.65803

On the Privation Theory of Evil: A Reflection on Pain and the Goodness of God's Creation

PARKER HARATINE
Logos Institute, University of St Andrews
tph2@st-andrews.ac.uk

Abstract: Augustine's privation theory of evil maintains that something is evil in virtue of a privation, a lack of something which ought to be present in a particular nature. While it is not evil for a human to lack wings, it is indeed evil for a human to lack rationality according to the end of a rational nature. Much of the literature on the privation theory focuses on whether it can successfully defend against counterexamples of positive evils, such as pain. This focus of the discussion is not surprising, given that the privation theory is a theory about the nature of evil. But it is also a theory that protects venerable theological concerns, namely, that God is the good creator of everything, and that everything is good. It is the purpose of this article to further this discussion on both fronts. I argue that the counterexample of pain still defeats the privation theory despite the most recent defense. What is more, I suggest, this is not theologically disastrous. The individual who rejects the privation theory is not obligated to reject the theological theses which motivate it. To show how a rejection of the privation theory is a live option, I offer an alternative view of evil that also maintains these theological theses and encompasses both privative and positive evils.

Keywords: Positive Evil, Good, Privation, Pain, Creation, Being, Opposition

1. Introduction

Augustine's privation theory of evil holds sway as the orthodox account of the nature of evil in Christian theology (Cf. Augustine 2019, book 7). According to this theory, something is evil in virtue of a privation, a lack of something which ought to be present in a particular nature. While it is not evil for a human to lack wings, it is indeed evil for a human to lack rationality according to the end of a rational nature.

Much of the literature on the privation theory focuses on whether it can successfully defend against counterexamples of positive evils, such as pain (e.g., Calder 2007; Lee 2007; Samet 2012). Pain seems to be evil, though it does not seem to be a privation of something. This focus of the discussion is not surprising, given that the privation theory is a theory about the nature of evil. But it is also a theory that protects venerable theological concerns. Not much of the literature discusses whether the privation theory is required by these certain theological concerns, namely, that God is the good creator of everything, and that everything which exists is good. Indeed, adherents normally assume that these theological theses entail the privation theory of evil.

It is the purpose of this article to further this discussion on both fronts. I argue that the counterexample of pain still defeats the privation theory despite the most recent defense. What is more, I suggest, this is not theologically disastrous. Rather, the individual who rejects the privation theory is not obligated to reject the theological theses which motivate it; these two theological theses do not entail the privation theory of evil. To show how a rejection of the privation theory is a live option, I offer an alternative view of evil that also maintains these theological theses and encompasses both privative and positive evils.

I shall proceed by discussing the motivation for and mechanics of the privation theory. I shall then argue that the most prominent and recent defense of the privation theory does not successfully defend against the counterexample of pain. I shall then suggest this conclusion is not theologically disastrous, for the two theological theses do not entail the privation theory. I then close by briefly offering an alternative view of evil that encompasses both privative and positive evils.

2. The Privation Theory of Evil: Its Conditions & Motivations

For present purposes, I use 'evil' interchangeably with 'bad.' Evil will denote something necessarily evil for a relevant object. It is necessarily evil in that it is evil across all possible worlds for an object, not occasionally or in some possible worlds. It is evil as a kind of phenomenon, regardless of the degree it manifests—I do not limit myself to the discussion of horror or trauma but include the whole gamut of evils, including minor injuries and pains. The privation theory of evil addresses the evils which are necessarily bad for the object in whatever degree they manifest. As will become clear below, the privation theorist's main concern is limited to intrinsically evil things.

There are two main reasons to believe that evil is a privation. The first involves a generally Aristotelian metaphysic of human nature which involves natures and

ends. The second involves explicitly theological reasons about God's being the good creator of everything which exists (Cf. Oderberg 2019; MacDonald 1990). While these two overlap to a degree, I will mainly address the latter theological concerns.

The theological motivation splits into two distinct theses. The first thesis is that God is the creator of everything aside from Godself, or *Deus Creator Omnium* (henceforth, DCO). The second thesis is that goodness and being are interconvertible in reality—everything which exists is good; and everything good exists (henceforth, BG). Different versions of each thesis are almost universally accepted by Christian theologians, including those from Orthodox, Protestant, and Catholic traditions.

Before I detail the mechanics of such claims, consider how thinkers from each of the major Christian traditions subscribe to both theses. Representing an Orthodox position, David Bentley Hart writes that evil is "a privation of the good, a purely parasitic corruption of created reality, possessing no essence or nature of its own . . . [God] is the source of all things, the fountainhead of being, everything that exists partakes in his goodness and is therefore, in its essence, entirely good." Further on in the same text, Hart writes, "This is not to say that evil is then somehow illusory; it is only to say that evil, rather than being a discrete substance, is instead a kind of ontological wasting disease. Born of nothingness, seated in the rational will that unites material and spiritual creation, it breeds a contagion of nothingness throughout the created order" (Hart 2005, 73).

Similarly, Ian McFarland represents a Protestant position. He writes,

> Within the context of creation from nothing, because anything that exists other than God is by definition a product of divine willing and therefore good, it follows that evil, as that which God does *not* will, is a lack (or privation) of being. Evil can therefore be said to "exist" only in an improper sense: it has no genuine being of its own but is instead parasitic upon that which does exist, in the way that the evil of rot is dependent on the goodness of an apple. (McFarland 2014, 114)

McFarland qualifies and nuances this claim: evil arises out of the "inexplicable creaturely rejection of God and thus lacks any ontological ground" (McFarland 2014, 120). Moreover, "it can only be described as a failure of being: the creature (whether human or angelic) failing to be what it properly is" (McFarland 2014, 200). While McFarland is overtly Augustinian in his emphasis that the choice of evil is 'inexplicable', McFarland and Hart are in essential agreement with respect to DCO and BG.

Regarding the Catholic position, the Catechism and Aquinas each maintain that God is the Creator of everything which exists, and everything is good (Catechism

1993, 299, 386). Aquinas explicitly draws the conclusion that evil as such is a privation (Aquinas 2003, 58). A contemporary Catholic Philosopher, Patrick Lee, succinctly writes the following:

> The position that evil as such is privation is entailed by the theistic position that all positive reality is God and what he creates. According to theism, God is immediately operative in every effect, and is thereby omnipresent, since a creature, not having existence as part of its nature, cannot cause new existence by itself. If evil were something positive, then one would have to say either that this evil is immediately caused by God, in which case God is in some way evil (since the effect reflects to some degree the nature of the cause), or that there is some being in the universe which is not immediately caused by God, in which case there is some creator other than the one God. If evil is not privation, then theism is incoherent. (Lee 2007, 470)

Despite minor discrepancies, there is profound agreement among adherents across traditions over the two theses DCO and BG. These theses in turn support the claim that evil is nothing but a privation of the good. Indeed, the above thinkers maintain that these theses *entail* the thesis that evil is a privation.[1] If this is the case, a denial of the privation theory would require a denial of DCO or BG (more below).

Let us turn to analyze the mechanics of these claims. Consider the first (DCO), that God is the Creator of everything aside from Godself. This thesis maintains that nothing can exist without God's creative activity, and only God exists absolutely and independently. All other things, anything with 'positive ontological status,' is created by God, exists dependently, and participates in His being. God is both radically sovereign over and ontologically independent of His creation, and all which exists is, asymmetrically, radically dependent upon and created by Him.[2]

This thesis comes in different versions, though the strongest maintains that God is the proximate and immediate cause of everything which exists.[3] God is not merely the remote cause of things, such as the builder of a window who causes sunlight to

[1] While many thinkers assume this achieves an entailment relation, Lee provides a generic reason: that 'the effect reflects to some degree the nature of the cause.' However, giving such a reason seems to be the exception rather than the norm.

[2] I must presently set aside the issue of abstract objects and necessary truths such as 2+2=4. Aside from these, it seems relatively uncontroversial for the Christian theologian and philosopher to accept the dependency of creation and objects aside from God.

[3] Notice that here and in the forthcoming, 'cause' is not necessarily limited to efficient causation in the traditional sense of the term; it is equally applicable to cause as determination or ground of being, in which there are different levels of 'causality' (such as the standard Thomistic narrative of primary and secondary causality). For present purposes, I will limit myself to the terminology of remote vs proximate cause, where proximate cause can be understood as efficient or not.

come into the room, but an unmediated and proximate cause of everything. In addition to Patrick Lee's statement that God is "immediately operative in every effect," consider Hugh McCann's position. McCann maintains that God is like the author of a novel who creates characters in their very being and actions, directly causing them to be and act the way that they are. Creatures do not add to the sum of things in the world when they act; they do not act in a way that is existence conferring. Rather, God is sovereign over and acting with and through creatures but not 'upon' them in a way that, so McCann urges, does not violate creaturely freedom (McCann 2012, chaps. 2 & 5).[4] So too for the remainder of creation.

Notice that this position can be understood to mean that God is the necessary cause of all things, the *sine qua non* of everything's existence; it could also be understood to mean that God is the sufficient and unmediated cause of everything, a position which seems to leave no room for undetermined causal input on the side of creation. On either understanding, however, the creative and sustaining activity is proximate and immediate, not remote. Given that most defenders of the privation theory maintain God is the immediate and proximate cause of everything,[5] and given that sustenance and creation are treated as logically equivalent (at least),[6] I will assume this 'thick' version of DCO is the thesis that motivates the privation theory.

The second thesis (BG) is that goodness and being are interconvertible in reality (though not in sense).[7] Everything which exists is good, and everything good exists. BG is motivated most acutely by the concern that God is a *good* creator, so everything which comes from Him must also be good. It would be a categorical error to maintain that God is the creator of evil. For evil (the arguments go) is the opposite of the good. Whatever it might mean for something to be an opposite, it is clear that evil cannot comprise or be comprised by good—it is in some significant sense a negation of, a privation of, or contrary to the good.[8]

---

[4] Another instance of this view is that of (Grant 2016, 231). Grant defends an 'extrinsic model' of divine activity in which God's act is not prior to but concurrent with the human act and is importantly *not* logically sufficient for the human act to obtain. Grant's view seems to be similar to that of Brian Leftow, who maintains what he calls an 'Immediate Late Creation' view in (Leftow 2012, 14-22). There, Leftow claims that God's immediate and late creation is sufficient in most cases, but only necessary in some cases. Thus, it does not result in overdetermination.

[5] In addition to the above footnote, see (Grant 2009).

[6] There is a question whether creation and sustaining are in fact different types of activity, or ever separate in practice. McCann thinks they are just the same, while Leftow thinks they different kinds though are never separate.

[7] For an example of the distinction between sense and referent, consider how water and H2O are interconvertible in reality, though not in sense.

[8] The opposite of correlative, such as 'father of' and 'son of', does not adequately describe evil.

Patrick Lee grounds the above intuition, that God can only create good, with the following principle. He maintains, "the effect reflects to some degree the nature of the cause" (Lee 2007, 470). One would do well to strengthen this claim and add, '*necessarily*, the effect reflects to some degree the nature of the cause.' Otherwise, it is not clear to which cases this is applicable and whether it is also applicable to the case of God's creative activity. Call this the principle of like-effect. Given DCO, if evil were something with positive ontological status, then it would be something God creates. Thus, we can conclude either that God is evil in that he has caused something evil (by the principle of like-effect), or evil is ontologically nothing. Notice this argument holds a lot of intuitive force in the thick interpretation of DCO. In contrast, if God were the remote cause of some objects which exist, it is not as clear that these objects must be good because it is not clear that the principle of like-effect applies to remotely caused objects. So, I will understand this principle to address those things which are immediate effects of a cause.[9]

This second thesis, together with DCO, neatly generates two theses to which the privation theorist subscribes:

1. Everything which exists, insofar as it exists, is good;
2. Everything which is good, insofar as it is good, exists.

From this and the general notion of evil being good's opposite, we derive:

3. Nothing which exists, insofar as it exists, is evil;

[9] Some have suggested that, because evil is ontologically nothing, God cannot be responsible for it. For example, Peter Furlong remarks, "Since this omission [viz. privation of following the moral law] is not itself an entity, it is not obviously the case that God is responsible for it." (Furlong 2014, 426) Admittedly, Furlong moves quickly on from this claim and has since updated/ expanded his view in (Furlong 2019, chap. 4) Nonetheless, it is worth noting why the original claim will not do. If God cannot be said to cause evil because of its lack of positive ontological status, it is not for any lack of ability on God's part, the creator who can and has brought something from nothing. Rather, it is because 'nothing' is precisely the total lack of positive ontological status such that it cannot be the proper object or recipient of causation and cannot be an effect of a cause in any sense.

If 'nothing' cannot be the proper recipient or effect of a cause, it follows that neither divine nor human agents can be said to cause evil. While this response may succeed in 'getting God off the hook' of moral and/or causal responsibility, this is so because it gets everyone off the hook. However, this is too high a price to pay. The causal issue is better addressed by the interconvertibility of goodness and being and doctrine of creation—these more obviously require that everything is good because it originates from a good God.

4. Nothing which is good, insofar as it is good, is evil.[10]

For present purposes, notice that these two theses provide powerful motivation to maintain that evil does not exist. More articulately, nothing is evil in virtue of its existence. Notice this claim is exhaustive, meaning, for the domain of existing things, there is no x such that x is evil in virtue of its existence.

What does it specifically mean to claim that nothing is evil (or everything is good) insofar as it exists? It depends on what 'existence' or 'exists' means. There are two relevant ways to understand this: absolute existence and kind-relative existence (and thus, absolute and kind-relative goodness).[11] Absolute existence is an all-or-nothing affair. It is a binary that does not admit of degrees—either x exists, or x does not. Consequently, something is absolutely good in virtue of its absolute existence. Kind-relative existence is not a binary and admits of degrees. Something is good relative to its kind just in case it exemplifies its nature and fulfills the end of its nature. For instance, both a human being and a horse are absolutely good just because they exist and are created by God. But a horse and a human being will be good very differently concerning their kind. To exemplify its kind-nature, a horse will have four legs and eat grass; a human, in contrast, will be rational and morally upright. According to the former, goodness is interconvertible with absolute existence—it is binary and does not admit of degrees. Either the horse exists and is good or not, regardless of it being a horse. According to the latter, goodness is interconvertible with kind-relative existence and *does* admit of degrees. A horse could have three legs and begin to eat rocks. This odd horse does not cease to be a horse, but simply fails to be horse-like and thus exemplifies 'horse nature' to a lesser extent.[12]

---

[10] For a very helpful overview of these thesis and additional relevant theses, cf. (Gracia 1990).

[11] While certain Thomists and Aquinas may dispute this distinction, it is not without historical precedence and good reason. For instance, Anselm of Canterbury subscribes to this distinction. Cf *Monologion* 1-3, and *De Casu Diaboli* in his discussion of justice as a kind-relative goodness. Moreover, consider how kind-relative goodness is understood best in terms of an entity having a potency to be a certain thing. Now consider how a lightbulb may have the potency to exist to some degree as a lightbulb; it may shine brightly and be a good lightbulb, or it may shine dimly and be a bad lightbulb. While this may be the case, a lightbulb does not have a potency to exist absolutely. For potencies refer to abilities (whether active or passive) of already existing entities. So, a lightbulb's mere existence is not reducible to a potency, but a modal or metaphysical possibility. And because a kind-relative existence is indeed best understood in terms of potency, it follows that absolute existence is different than kind-relative existence.

[12] While it may be odd to think of existence in terms of degrees, this is something (as will become clear) to which the privation theorist must subscribe. While I do not presently defend this as my own view, the privation theorist could maintain that an intuitive way to understand existence in degrees

So, when it comes to evil, we can interpret claims 1-4 in either of these ways. It seems that the Christian theologian and philosopher will at least maintain that everything is absolutely good insofar as God is the creator of everything; from this, it follows that nothing is evil in virtue of its absolute existence.

Many go further and posit that nothing is evil in virtue of its kind-existence—the fulfillment of all created things is good, and only good, with respect to their kind. Notice that this can be understood in two ways. First, that everything which exists is good only insofar as it fulfills its kind-nature in the sense of reaching its proper end and *telos* (the axiological sense). Second, that all things with positive ontological reality exemplify a kind-nature in the sense of being a defined entity (the metaphysical sense.)

Given my theological focus, I will only presently assume there is a necessary connection between the former view, between absolute goodness and being. Of the two kinds of existence under consideration, absolute existence is more conceptually minimal and follows from DCO and BG since any other type of existence must include absolute existence. In contrast, kind-relative existence requires additional explanation and premises to follow from DCO and BG. Even when an object exemplifies a kind-nature, and even when the object is good to the degree that it exemplifies this nature, it is not merely good because of this (viz., its goodness is not reducible to or exhaustively explained by this). God's creative activity is still the ultimate (if not proximate) explanation for the necessary connection between being and goodness, even when an object exemplifies a kind-nature. That is, while these explanations might not be exclusive or 'competitive', God's creative activity is still required for a full explanation because the privation theorist's original claim is that DCO grounds BG. Moreover, I need not assume that all things with positive ontological status exemplify a kind-nature in the sense of having a proper end and

is that a created object might exemplify its kind-nature, the exemplar of what it is to be that kind of thing, more or less. So, if the exemplar/kind/type horse has four legs, then the token three-legged horse does not 'match' the kind as well as it could, and thus exists a little less with respect to exemplifying the kind. One can also consider this in terms of set theory. Consider the set of 'winners'. Being in the set is a binary: either someone is in or out of the set, and only those who win are in the set. Now consider two individuals who race separate 500-meter foot races. One runner might complete (and win) this in 57 seconds and cheerily congratulate his opponents; another runner might complete (and win) his own race in 60 seconds and taunt his opponents for losing. Both of these runners are in the set of winners, though the first is a *better* winner on two counts: his race-time and sportsmanship.

telos as defined by the nature in question. DCO and BG do not require this, and it takes us too far afield.[13]

Within the above framework, the privation theorist maintains that evil does not exist with any positive ontological status. As should be clear, this claim does not mean evil is equivalent to nothing, an unfortunate and common misunderstanding (Stump 2009). This is rather to say that evil is *ontologically* dependent for its existence. Like a hole in a donut or the rot on the apple, evil does not exist on its own. When the donut goes away, the donut-hole also goes away. In this sense, evil is entirely dependent upon that which exists, and does not have a *subsistent* existence of its own—it is a privation of existence.[14]

What exactly, then, is a privation? According to the Augustinian account, consider the following definition: "There is a privation of x if and only if something y lacks or loses x, and the nature of y is such that it ought to have x" (Swenson 2014, 142).[15] Notice that the end of the nature prescribes what counts as a relevant lack. While I may lack wings as a human being, this is not an evil—my nature does not require me to have wings. My lack of rationality, however, would be an evil according to the end of the nature 'rational animal.' While the donut-hole example is an example of a purely metaphysical lack, a privation is not value-neutral. A privation is both a metaphysical and axiological lack.

Notice that the privation and its opposite of possession must apply to the numerically same object. For example, if person A is blind, it is only a privation for person A's eyes, not person B's eyes. This is required for all evils to be considered *intrinsically* evil with respect to the entity in question, be it a person, action, or state of affairs.[16]

---

[13] As is evident, the privation theorist requires a version of kind-existence in order to make sense of privations: absolute existence simply doesn't admit of degrees and does noy dictate which lacks are bad as privations.

[14] In this context, I use the term subsistence to mean what the ontologist would call existence. Where individuals may normally say things such as cracks and holes 'exist', this is not technically correct. To differentiate between the common language and the technical sense, I use subsistence as a way to denote something with positive ontological status.

[15] Swenson cites Aquinas, *ST* I, Q 48, A3 & A5.

[16] The reader may question the notion that actions and states of affairs are entities. This is not to say that actions and such are substantial entities, such as animals. Yet, as W. Matthews Grant notes, actions being entities, or having positive ontological reality, is simply an assumption that the reader must grant from the outset. It is also not one without precedence. Aquinas, for instance, considered actions to be non-substantial entities (*ST* I-II Q 79, A2). Cf. (Grant 2016, 224; Grant 2015, 272). It is important for the privation theorist for actions to be things, for privations pertain to the numerically same object as the thing which is under the obligation to possess a property. If actions are not things, then there are not morally evil (or good) acts on the privation theorists' metaphysic.

Some authors, such as Adam Swenson, have argued that the privation need not inhere in the numerically same object. That is, object A can be evil in virtue of another privation that belongs to object B (and A is not B) (Swenson 2009, 144).[17] While Swenson mainly outlines this as a logical possibility, I do not see any reason to entertain this as a legitimate position. First, the canonical version of the privation theory addresses intrinsic evils. These are evils that are of the numerically same object, and not extrinsically based on another object or a relation that A holds to B. This is what Augustine, Aquinas, and Suárez addressed as the privation theory, and an expansion of this is not what traditional defenders have in mind.

Moreover, I think we have two good reasons to limit ourselves. On the one hand, it would not be a problem for DCO and BG to admit that something is extrinsically evil. BG pertains to the intrinsic value of all objects. On the other hand, to admit that the privation theory encompasses such privations as Swenson outlines would be to say that not all evils are necessarily bad. But this is counterintuitive: when *something* is evil, we do not say it is accidentally or occasionally evil or evil in its effect upon something. Consider how an attempted act of murder is bad even if not effectively carried out. Consider also how states of pain are bad even if they do not cause a lack of happiness or health. Unless there is a necessary connection between the two supposed states, this requires that some things are accidentally bad. But this amounts to arguing that some things are accidentally evil, perhaps based on their context or the relation in which they stand. Rather than focus on things evils that are accidentally evil, the privation theory focuses on evils that are necessarily or intrinsically evil, across all possible worlds. There is good reason, then, to adhere to the view that a privation is evil for an object only if it inheres in the same object.

With this framework, a privation is an evil because it is a lack of something which ought to be present in a nature. A privation does not necessarily subsist and have positive ontological reality but exists dependently upon the good and the existing.[18] Notice again that this is exhaustive: for all evils, there is no x such that x is not a privation. In the above manner, the privation theory is both motivated by and guards the relevant concerns that God is the creator of everything, and that being and goodness are interconvertible in reality.

[17] This could be roughly understood as either T3 or T5 on Swenson's taxonomy, depending on whether object A also has a privation.

[18] Technically, we also need to say that a *diminution* of good is bad in itself to fully say why a privation is bad. On this, cf. (Swenson 2009).

3. Pain is Still a Problem for the Privation Theory of Evil

Thus far I have assumed that the foregoing theological theses hold water. One could certainly question this assumption, but my present aim is more constructive. I will argue that the privation theory of evil has yet to adequately address the counterexample of pain. I will then address the importance of this failure and suggest that, even if some pains are not privative evils, one is not required to reject these above theological theses wholesale.

In laying out this framework, I have intentionally used the language of 'motivation.' The two theological theses *motivate* the privation theory. Yet, as should be clear from the three representatives above, a much stronger claim is normally made. The theological theses are normally understood to *entail* the privation theory. This is a much stronger claim. Thus, by modus tollens, the denial of the privation theory would require the denial of one if not both of the above understandings of the theological theses.

In terms of logical structure, this amounts to the following: let DCO stands for the thick doctrine that God is the proximate and efficient Creator of everything; let BG stand for the interconvertibility of being and goodness; let P stand for the privation theory. Thus:

1. $(DCO \wedge BG) \rightarrow P$
2. $\sim P$
3. So, $\sim(DCO \wedge BG)$ (Modus Tollens)

Pace the above authors, I shall argue that DCO and BG do not entail P. Thus, my argument is that a denial of the privation theory does not require a denial of DCO or BG.

Why is pain a strong counterexample to the privation theory? Pain is an example of a positive evil. Pain seems to be evil, and it is not apparent that pain is always accounted for in terms of a privation of happiness, pleasure, or more generally, well-being. And if it is the case that pain is evil and not a privation, there is a non-privative evil. There are several potentially compelling responses the privation theorist can respond with to this issue, each of which I will address. I will focus on the work of David S. Oderberg, as his work is the latest and most thorough defense of the privation theory of evil. As will become clear through addressing the potentially compelling responses and Oderberg's work, 'painfulness' is still generally problematic for the privation theory, and *useless* pain is the most problematic for the privation theory.

One possible response to the issue of positive evils is to deflate the counterexample. Call this the *deflationary* response; this response deflates the counterexample by denying either that the evil is real or is evil. The first of the two deflationary responses maintains that a positive evil is evil but has no being. In this case, the pain is just a mental state, for example; it is but a perception of something.[19] For example, Irit Samet argues that the brain processes pain in two different ways: somatically and affectively. The somatic processing pertains to the duration, intensity, location, etc., and corresponds to A-*delta* fibers; the affective response of the brain responds to the C-fibers and correlates to the negative evaluation of such a state. This is why, for instance, patients given opioids to reduce the pain can attest to the location, intensity, and qualities of the pain (e.g., prickly, hot, sore) without reporting a negative affective attitude towards it. Accordingly, Samet argues that the pain is a phenomenological response that likely *supervenes* upon the C-fibers firing off in the brain. If this is the case, Samet argues, it is more difficult to conclude that pain is *real* and not merely the affective evaluation of the state in which one finds oneself (Samet 2012, 25-26).[20] If this were the case, pain's lack of being is not problematic for DCO and BG.

This response is illuminating in that it relies upon a standard distinction in the literature between pain and painfulness. Pain is the physical phenomenon that includes the intensity, duration, location, quality (e.g., prickly), and such; painfulness is the affective evaluation of this, that it is indeed bad, negative, and to be avoided (Oderberg 2019, chap. 5; Swenson 2009, 141). This distinction is similar to the standard distinction between pain and suffering. For example, while a fish might experience pain when it bites on the hook, it is not apparent that it suffers (or experiences painfulness). While this first distinction is standard, for simplicity I will generally use 'pain' to denote the negative aspect of pain and assume 'pain' correlates (though does not reduce) to different types of physical phenomena.

Though Samet's response is illuminating in this respect, this response is not successful. While Samet is clear that the affective evaluation is not reducible to the C-fibers, Samet must further deny that the affective and negative evaluation of pain which correspond to the neurons is also not something real. But a mental state or

[19] This approach appears similar to the contemporary criticism that all evil is a function of one's perception. The latter criticism relies upon value relativism, which would maintain that something is merely perceived as bad though not actually bad (for nothing is *actually* bad). On this approach and the different assumptions of contemporary approaches to the problem of evil, cf. (Gavrilyuk 2020, 66).

[20] For more on this general approach, and whether the affective representation of the content can be construed as a privation, Cf. (Oderberg 2019, 131-132).

perception is something and not merely a part or modification of the mind, reducible to something else. So, even if the pain were fictitious and not extra-mental, as in the case of a phantom limb, it does not follow that the pain is not real. The response simply relocates what about the pain is real.

The second route of the deflationary response is more promising. On the second route, the defender of the privation theory might rejoin that, while the pain is real and has being, it is not evil. This can be because the pain is a good of utility or merely a product of the organic function of the organism. Regarding the former, the pain could be a warning. When an individual stubs his toe or burns his hand, there is pain to warn against and deter the individual from such actions. As Oderberg argues, such pain simply *should* accompany such warnings, for otherwise, the warnings would be ineffective (Oderberg 2019, 130, 132).

While helpful in some cases, it is not clear that all cases of pain are goods of utility. For instance, the pain from a phantom limb does not apparently serve the further purpose of warning the individual.

According to the organic function response, however, some pain can be construed as non-instrumental and good. For example, consider a child who experiences growing pains. This is an achievement of the organism's operation and is good as such. This would account for the child's growing pains as well as potentially other cases such as chronic nerve damage. In the case of chronic nerve damage, there is a failure of the nervous system to reach its proper end. This failure could be a privation. Oderberg describes this pain *doing* vs. pain *achieving* (Oderberg 2019, 130). The growing pains are indicative of the organism or body part *achieving* a proper end, while the chronic nerve damage is a result of the nervous system *doing* but not achieving its proper end. The growing pains are good as a natural operation, and the nerve damage is a malfunction and thus accountable in terms of a privation.

While the above response is plausible in some cases, there are still counterexamples of 'useless' pain, such as phantom limb syndrome or inexplicable throbs and aches. Useless pain is not obviously construable as having an appointed end in light of which it is a privation. Useless pains are also not able to be construed as goods of utility, however they might be appropriated to develop one's character.[21] Useless pain may very well correspond to a disorder, such as the case of chronic nerve damage. Yet it is useless either because there is no underlying disorder, or because the alert it generates to the underlying disorder serves no purpose (Oderberg 2019, 133).

[21] So, for instance, though chronic nerve damage might not be a good of utility like the pain of burning one's hand, one could still choose to develop their character from the difficulty. But this does not mean the chronic nerve damage itself is a good of utility.

Oderberg addresses this issue of useless pain and considers it to be the central issue for the privation theory of evil. Regarding the case of the phantom limb, the pain cannot be construed as a disorder. The *painfulness* is not the physical neurons failing to achieve their end (which would be the *pain*), but rather the affective evaluation of the state (Oderberg 2019, 130-133). It follows that the painfulness itself is not construable as a disorder because it correlates but is not reducible to the physical disorder. And the painfulness is that which is bad, not the pain. It is thus 'useless' in that, if there even is an underlying disorder, it cannot effectively compel the agent to respond (an impossibility, one might say), and because it is not reducible to the disorder of the pain. So, while the deflationary issues are illuminating in certain respects, they do not capture and deflate the issue of useless pain.

This above point is applicable to the pains of utility and organic functioning. Even if there are goods of utility and organic functioning, this does not entail the pains which are goods of utility and organic functioning are intrinsically good. An object, in other words, can have a different intrinsic value from its extrinsic value. For example, Calder correctly maintains that even if something is a good of utility, it does not necessarily follow that it is intrinsically good. Consider money: money is a good of utility, though money's intrinsic value is neutral (Calder 2007, 374). So, even where pain may serve a good purpose, there remains the question of its intrinsic value. Since the painfulness is not reducible to the pain, there is a clear way to offer a different valuation. This point, along with the issue of useless pain, is still problematic for the privation theorist.

Let us turn to examine the privation theorist's final response to the issue of useless pain and Oderberg's construal of useless pain in particular. The privation theorists, rather than construe pain as a good of functioning or utility, can take the standard approach and argue that pain is intrinsically evil insofar as it is a lack of something. Rather than deflate the counterexample as above, this modifies why pain is bad. We can consequently call this the *modification* approach.

This is the classic *privatio boni,* where something is evil in itself just in case it lacks a perfection that it ought to have as prescribed by its nature. One might argue according to this line that pain is evil yet is not a total lack of being. Rather, one might simply say that pain is a degreed lack of something, such as pleasure, internal equilibrium, or happiness. On this approach, pain is something in virtue of its existence through pleasure, yet evil with respect to the lack of pleasure.

The problem is that pain is not always a lack of pleasure or internal equilibrium. While an individual might be in mental anguish over the loss (the painfulness) of a child (the pleasure), people can also experience pain without this being construed as a loss of something. Consider again the case of the phantom limb. If the individual

is in pain because of the phantom limb, the pain itself is not obviously a loss of pleasure, though the pain may further cause a loss of pleasure.

Oderberg instead argues that useless pain does not merely cause a lack of pleasure but causes a lack of ability to function well as a human being. Oderberg is careful to note a lack of mental equilibrium is bad because of useless pain, but it is not equivalent or reducible to useless pain. Rather, useless pain is bad because it prohibits functioning well as a human being. In making this point, Oderberg does not, unfortunately, use an example of useless pain but the example of the pain which accompanies a sprained ankle (pain of an organic function or utility, perhaps). In the case of the sprained ankle,

> There will be the same functionality as we find in the standard case – direction to damage – where the loss of equilibrium is still no privation. But there will also be other things happening, for instance when the pain of a sprained ankle causes me to ignore a red light at a pedestrian crossing, or some such. This specific loss of equilibrium is a privation: good functioning requires me not to ignore dangers to my well-being. So there is both privation and mere non-privative absence in such cases. (Oderberg 2019, 135)

Thus, the ankle-sprain is painful in terms of organic functioning and is a direction of one's attention to pain (a good of utility). It is therefore good. Yet the sprained ankle is also a 'proper' privation in that it causes failure to function well as a human being.

According to Oderberg, it would be problematic for two reasons if the useless pain and pains of organic functioning/ goods of utility were the same thing. First, useless pain is that pain which is not construable as a malfunction or pain of utility. And Oderberg, like any privation theorist, cannot say that the pain *just is* the privation of mental equilibrium; this proposition would entail that organic function and pain as a good of utility are just the same thing. So too for useless pain. If useless pain just is the privation of mental equilibrium, it is equivalent to these other types of pain, causing a total collapse between all types of painfulness. Thus, Oderberg posits that useless pain is pain that further causes a failure to function well.

Second, if the sprained ankle were only one phenomenon of pain, this would entail opposite predicates of the one phenomenon in the same respect. The pain would be good in that it alerts, and bad in that it causes failure of well-being. But, to ascribe contradictory predicates of a single phenomenon is a theoretically unacceptable conclusion. Yet the absences of the equilibrium are different in some respects. Oderberg writes:

> Does this mean that one and the same absence is both privative and non-privative? . . . The part of the overall feeling of pain that makes you look for damage is not the same as the part which causes you to ignore a red light. There are two disturbances, two reactions, and it is of little concern whether we say these are two parts of one overall reaction or not. Note that if you ignore the red light at the very same time as you are directed to look for damage, this still does not mean that there is a single disturbance that is both privative and non-privative. In such a case you ignore the red light because you are looking for damage, not because of the very same disequilibrium in virtue of which you are looking for damage. Such is the situation when you turn your attention from the light as you reach for your ankle. (Oderberg 2019, 135)

In one sense, pain is an absence of mental equilibrium and is not bad because the absence is due to the organic functioning of the body or a good of utility. In another sense, pain is bad insofar as it causes a privation of one's functioning well as a human being, such as the ability to be alert when crossing the street. If the latter disequilibrium accompanies the sprained ankle and thus the organic functioning pain, then it is two different aspects of the same state of affairs. And because it is two different aspects, it is not contradictory to predicate this of the disequilibrium.

I have detailed this last part because it is crucial to Oderberg's argument, and any argument, that is to avoid a collapse of types of pain into each other. Useless pain cannot merely be a lack of equilibrium as this would make it indistinguishable from the pain of utility or organic functioning. And by the law of identity of indiscernibles, these would just be the same.[22]

There must be a relevant distinction to avoid the collapse of useless pain into pain from organic functioning. To avoid this collapse, Oderberg posits this *causal* relation and maintains that the painfulness of useless pain is not merely the lack of equilibrium but causes a failure to function well.

Consequently, this means that the useless pain is bad because it *causes* a state of not functioning well. That is, Oderberg is consequently committed to saying that this is extrinsically bad, not intrinsically bad.[23] Thus, Oderberg is also committed to saying (as he does in the example) that useless pain is bad insofar as it causes failure to function well as a human being.

---

[22] The law of identity of indiscernibles is that, if two things share the exact same properties, then they are in fact identical (i.e., not two numerically different objects).

[23] If the disequilibrium does *not* accompany an organic function or pain of utility, then Oderberg is still reticent that one cannot say this disequilibrium *just is* pain. For there are certainly other different types of pains of disequilibrium that are not captured by useless pain.

This is highly problematic for several reasons. Oderberg has prohibited this move before—he clearly maintained the canonical version of the privation theory of intrinsic evils and took other authors to task for conflating intrinsic privation and extrinsic privations (Oderberg 2019, 131, 134). On Oderberg's own terms, this move is impermissible.

Aside from the authorial inconsistency, the available options are not promising. The defender of the privation theory such as Oderberg could bite the bullet and insist that an extrinsic privation is as much a privation as is an intrinsic privation. In other words, it is acceptable for something A to be bad insofar as it is bad in virtue of a privation of another object B, a lack which A does not exhibit. But as already noted, this fails to be a canonical form of the privation theory.

Even if Oderberg pursued this route, useless pain does not always cause a failure to function well. For instance, the soldier's pain from the phantom limb does not necessarily cause him to fail to look at the stoplight (inasmuch as the sprained ankle doesn't necessarily cause someone to do the same). Indeed, the pain of the phantom limb might cause the soldier to be *more* cautious when crossing the street and thus causally contribute to his functioning well as a human being. It is not clear that Oderberg's account details why useless pain is bad even when it does not cause a failure of functioning well. On his account, if the useless pain were to not cause a privation of functioning well in some instances, then it would not be bad in those instances.

But the problem of useless pain is that it is always bad for the individual it affects, not merely when it exhibits this causal relation. The soldier who experiences phantom limb pain is in pain regardless of the causal relation. To see this point, consider, for example, two individuals Sylvia and Tyron. Sylvia and Tyron have experienced a car accident and are both paralyzed from the waist down. Both, we might say, are inhibited from using their legs and thus functioning well. Imagine further that Sylvia experiences pain in her legs, whereas Tyron does not. Does the pain in Sylvia's legs inhibit her from functioning well? Not obviously. Rather, this pain seems to be both bad and useless without respect to the causal contribution it has on Sylvia. Indeed, as Calder argues about money's two values, it seems that useless pain is intrinsically bad, in addition to (or regardless of) frequently exhibiting this causal relation (Calder 2007, 374). Thus, the problem of useless pain (in addition to painfulness more generally) is still a problem for the privation theorist.

4. The Theological Upshot

Given my arguments about useless pain and painfulness more generally, it follows that there is a positive evil. The privation theory of evil is not exhaustive, then, even if many evils are still privative. If the above theological theses *entail* the privation theory, the conclusion requires the denial of one if not both of the antecedents. Rather than deny one of the antecedents, however, I would like to question whether it is an entailment relation.

To see how it is not an entailment relation, it will behoove us to consider some of the argumentative methods available at this point.[24] When an interlocuter claims an entailment relation, *necessarily*, *if p then q*, the interlocuter is claiming the antecedent of an entailment relation is, necessarily, a sufficient condition for the consequent. There is no case in which the antecedent obtains and the consequent does not obtain. So, how do I go about showing that this is not the case? First, one could show that *p* and not-*q* are compatible in some fashion (i.e., possibly, (*p and not-q*)). For example, I could show that the denial of the privation theory is logically compatible with the DCO and BG. Second, one could show that it is not obvious that *q* follows from *p*. Rather than demonstrate that *q* does not follow from *p*, this would be equivalent to saying my interlocutors have not yet shown decisively that the privation theory follows from the antecedent theses; the burden of proof has not yet been met.

Regarding the first sort of argument, the negation of the privation theory is compatible with the thick version of the doctrine that God is the creator of everything and the interconvertibility of being and goodness. To see this, simply consider that DCO and BG are compatible with the claim that no evils exist. If no evils exist, then no evils are privative evils. This claim is logically consistent with and DCO and BG (i.e., possibly (*p and not-q*)). Strictly speaking, DCO and BG only entail that evil does not exist, or more precisely, that nothing is evil in virtue of its existence.

Regarding the second sort of argument, the above authors have not provided evidence to even suppose the privation theory follows is the case. Recall from Section 2 that BG entails 'nothing which exists, insofar as it exists, is evil.' Because a lack of ontological status is not equivalent to a privation, it is a further jump in one's argument to say that evil is a privation of a proper good. Indeed, the only argument given from the above authors is the one from Lee, who invokes the principle of life-

[24] The following sketch is my own, though the thought to provide such a sketch was inspired by (Williams 2005, 581).

effect. Yet this principle only provides grounds for the interconvertibility of being and goodness. It does not ground the claim that evil is a privation.

The above is enough to show that the entailment claim does not hold. It might be objected here that such a claim does not disprove the privation theory. For the privation theory might be construed to claim that, necessarily, for any x, if x is evil, then x is a privation of a proper good. Thus, to claim that evil does *not* exist is not an issue for the privation theorist—the privation theory makes no such existential claim! While this may be the case, the objection misses the point. The present claim under consideration is not whether there are any evils, but whether DCO and BG entail the privation theory. And this they do not.

While DCO and BG do not entail the privation theory, I imagine that the reader is dissatisfied with the response that it is logically compatible with the claim that evil is nothing. Not many readers, I suppose, will deny the existence of evil as a live option. But I must press the point—even if DCO and BG are compatible with the denial of the privation theory, it is a very different question whether this compatibility is a *live* option. The one who subscribes to DCO and BG and also denies the privation theory may well have to deny that evils exist if there are no other options. Of course, if denying the existence of evils is not a live option, it may simply bring the theist back to the privation theory as the only obviously live option. This issue is particularly pressing because useless pain is an example of a positive evil. And since there is a positive evil, we have a case of something which exists and is evil. So, one might worry that DCO and BG must still be denied unless one can account for the positive evil in a way that is compatible with these theses.

So here is an alternative (very brief) understanding of non-privative evil that is compatible with the two theological theses. Call it the *Opposition View of Evil*. In addition to privative evils, some evils stand in opposition to the good. Anselm of Canterbury and Francisco Suárez maintain versions of this view.[25] While the evil effect or event (say, useless pain) might be good in virtue of its absolute existence, it need not be good in virtue of its kind-existence or even be something with a kind-existence in the sense of having a proper end and *telos*. Recall that only absolute goodness is required to explain BG, and thus, I am not presently assuming that all entities are good in virtue of their kind existence. Hence, nothing prohibits us from

[25] For example, cf. (Anselm 2007, *On the Fall of the Devil*, 26; Anselm 2007, *On the Harmony*, I.7; Suárez 1989, XI.8). My view is also similar to the view of (Pruss 2022). While Pruss also maintains that positive evils are only 'evil' in relation to other objects, he further argues that the complex entity made up of the evil (e.g., pain, act of murder) and the relation and other object does not exist. Mine is different from Pruss's view because I additionally maintain that a positive evil is intrinsically or necessarily evil while absolutely good.

maintaining that useless pain and painfulness are entities that are opposite to well-being. The pain is a positive phenomenon that might very well cause a lack of well-being, though is simply bad to experience for sentient creatures that can suffer. Pain would be a kind of phenomenon yet need not have a proper end and telos that is good for it to achieve.[26] More generally, positive evils would be described as those things standing in opposition to the good. There is something about these positive evils that, when in relation to another object, they are bad in relation to this object and in opposition to a good. On this understanding, something is necessarily evil in that it is evil in relation to another object (and thus, evil *for* sentient creatures) across all possible worlds in which that later object exists, though good in virtue of its absolute existence. We can thus say some entities are either intrinsically or necessarily evil while maintaining that things are good in terms of absolute existence.[27]

Apply this more fully to the above example of painfulness and useless pain. Recall the thought experiment involving Sylvia and Tyron who are paralyzed from the waist down. While Sylvia experiences useless pain in her leg, this pain is bad regardless of whether it causes a failure to function well. It is bad in relation to the person, Sylvia, and bad in that it is opposed to well-being. Pain is bad only in relation to sentient creatures that can suffer, for example, and not in relation to rocks. Thus, pain is bad only in the possible worlds in which sentient creatures that can suffer exist; pain is not bad in some possible world in which the only created entities are rocks. Now, the reader will notice that one can explain this pain in context of well-being. But the fact that well-being could be (or even needs to be) involved in the explanation of pain does not require that pain causes a privation of well-being. For example, the hedonist does not need to maintain that a certain pain is a privation of happiness in order to explain that pain is opposed to happiness (Cf. Calder 2007, 378-9). Likewise, pain is a positive and quality on its own that can be explained as opposite to well-being.

The merits to the opposition view of evil are thus: the opposition view of evil maintains DCO and BG; it accounts for both privative and non-privative evils; it

[26] The idea that pain has no end at all is similar to Marilyn McCord Adams' view of horrors, which are *dys*teleological. This is compatible with the view that God can redeem these evils, even if they have no natural or good end.

[27] One might object: pain is an extrinsic evil, not an intrinsic evil. For pain's badness depend upon a relation. This objection is interesting but does not go through. Health is intrinsically good, though it likewise depends upon a relation for it to be good. Health is a good in relation to sentient creatures that can have well-being. Thus, pain is extrinsically bad only if health is extrinsically good. We avoid pain for its own sake inasmuch as we seek health for its own sake.

accounts for the phenomenon that there is still something intrinsic about the positive evils that only obtains in relation to certain objects and stands in opposition to the good.

I cannot develop presently how this view works with other positive evils, such as the classic examples of murder or error in belief (e.g., the belief that 2+2=5). This takes us too far afield, and in any case, is not required for the argument to go through. The above view is compatible with the presently assumed understandings of DCO and BG. It follows that the negation of the privation theory is logically consistent with DCO and BG and is prima facie a *live* option.

One might object that the opposition view of evil, while it maintains a larger explanatory scope than the privation theory of evil, ascribes to God to creation of objects which are intrinsically and necessarily evil. This is a concern of moral justification—why would a morally good God create such positive evils? While interesting, this is the topic of theodicy and not a question the present article must address. Even so, nothing about the opposition view of evil changes what standard answers are available to the question of moral justification.[28]

This suffices to conclude my main argument. There is still a strong counterexample of pain to the privation theory. But the thick version of DCO and BG do not entail the privation theory of evil. So, the denial of the privation theory does not require the denial of DCO and BG. Because I claimed there is a positive evil, I provided an additional way to construe positive evils that subscribes to DCO and BG. On the assumption that the specific construal of these doctrines are theologically important, nothing theologically disastrous follows.[29]

---

[28] E.g., free-will, soul-making, etc. It is a merit of the opposition view that intrinsically/ necessarily evil things exist. On the privation only view, a privation is not able to be the proper object of a cause. Thus, God cannot be said to be the direct cause of evil, though he can directly permit or indirectly orchestrate the evil through causing other events. The opposition view includes this but further holds that God can be the direct cause of an evil (i.e., of painfulness).

[29] It is another issue whether God is indeed the proximate and immediate cause of everything like the thick version of DCO maintains. Indeed, this assumption seems more problematic than assuming that God is the remote cause of being. For example, occasionalism, theological determination, and overdetermination are issues. Hugh McCann is acutely aware of these concerns in (McCann 2012, chaps 2 and 5); Leftow address the issue of overdetermination (Leftow 2012). The assumption is usually maintained for views of sovereignty and providence, but it is not clear to me that maintaining God remotely causes and grounds many (not all) things is any less capable of addressing the concerns of God's sovereignty and providence over creation. While some (e.g., McCann 2012; Lee 2000, sec. 2) argue that God must be the proximate primary cause of everything, I do not see any reason to deny the position that God is the creator of everything fundamental to created reality, and how God is not the cause of non-fundamental things, like Adam (fundamental) and the act of the primal sin (non-fundamental). But this is for another time.

It is another question whether a thinner version of DCO (and BG) would entail the privation theory. But as I intimated above, it is even less clearly the case that, if God's remotely creates something, the remotely created object is good. I focused on and assumed the thick version precisely because this is what prominent adherents maintain, and because these would most likely require the privation theory of evil.

5. Conclusion

The privation theory of evil partially serves to protect venerable theological concerns. This is the case with historic and contemporary thinkers across the three major Christian traditions. I discussed the mechanics of the privation theory and argued that the privation theory still has not met the classic counterexample of pain. Even so, I suggest that this is not entirely problematic, for it is not apparent that the theological theses which motivate the privation thesis entail the thesis. Instead, DCO and BG are logically compatible with the denial of the privation theory. Moreover, I briefly provided a different way to account for the positive evil of pain. In offering this argument, I placed the privation theory in the context of important theological concerns and examined exactly how dependent it is upon them.

Bibliography

Anselm. 2007. *Basic Writings*. Edited and translated by Thomas Williams. Indianapolis: Hackett.

Aquinas, Thomas. 2003. *On Evil*. Translated by Richard Regan. Edited with an Introduction and Notes by Brian Davies. Oxford University Press, 2003.

Augustine. 2019. *Confessions*. Translated, with Introduction and Notes, by Thomas Williams. IN: Hackett.

Boethius. 1973. *Theological Tractates, The Consolations of Philosophy.* Trans. By H. F. Stewart, E. K. Rand, and S. J. Tester, in Loeb Classical Library, Edited by Jeffrey Henderson. Cambridge, MA: Harvard University Press.

Calder, Todd C. 2007. "Is the Privation Theory of Evil Dead?" *American Philosophical Quarterly*, Vol. 44, No. 4: 371–381.

Catechism of the Catholic Church. 1993. The Vatican. Accessed April 4th, 2022. Catechism of the Catholic Church (vatican.va).

Furlong, Peter. 2014. "Is God the Cause of Sin? An Examination of the Unadorned Privation Defense." Faith and Philosophy 31 (4): 422–34. https://doi.org/10.5840/faithphil201412922.

Furlong, Peter. 2019. The Challenges of Divine Determinism: A Philosophical Analysis. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108696845.

Gavrilyuk, Paul L. 2020. "Problem of Evil: Ancient Answers and Modern Discontents." In *Evil and Creation: Historical and Constructive Essays in Christian Dogmatics,* edited by David J. Luy, Matthew Levering, & George Kalantzis. Bellingham, WA: Lexham Press.

Gracia, Jorge J. E. 1990. "Evil and the Transcendentality of Goodness: Suárez's Solution to the Problem of Positive Evils." In *Being and Goodness: The Concept of the Good in Metaphysics and Philosophical Theology*, edited by Scott Macdonald, 151–176. New York: Cornell University Press. https://doi.org/10.7591/9781501728372-009.

Grant, W. Matthews. 2009. "Aquinas on How God Causes the Act of Sin without Causing Sin Itself," *The Thomist: A Speculative Quarterly Review*, Vol 73 no 3: 455–496. https://doi.org/10.1353/tho.2009.0015.

Grant, W. Matthews. 2015. "The Privation Account of Moral Evil: A Defense," *International Philosophical Quarterly* 55, no 3: 271–286. https://doi.org/10.5840/ipq201572139.

Grant, W. Matthews. 2016. "The Privation Solution: A Reply to Furlong," *Faith and Philosophy* Vol. 33, No. 2: 223–34. https://doi.org/10.5840/faithphil201632157.

Hart, David Bentley. 2005. *The Doors of the Sea*. Grand Rapids, MI: Eerdmans.

Lee, Patrick. 2000. "Evil as Such Is a Privation: A Reply to John Crosby." *American Catholic Philosophical Quarterly* Vol. 81, No. 3: 469–88. https://doi.org/10.1353/tho.2000.0014.

Lee, Patrick. 2000. "The Goodness of Creation, Evil, and Christian Teaching." *The Thomist* 64 (2): 239–69. https://doi.org/10.1353/tho.2000.0014.

Leftow, Brian. 2012. *God and Necessity*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199263356.001.0001.

Macdonald, Scott, 1990. *Being and Goodness: The Concept of the Good in Metaphysics and Philosophical Theology*. New York: Cornell University Press. https://doi.org/10.7591/9781501728372.

McCann, Hugh J. 2012. *Creation and The Sovereignty of God*. Bloomington, IN: Indiana University Press.

McFarland, Ian A. 2014. *From Nothing: A Theology of Creation*. Louisville, KY: Westminster John Knox Press.

Oderberg, David S. 2019. *Metaphysics of Good and Evil*. New York: Routledge. https://doi.org/10.4324/9780367809577.

Pruss, Alex. 2022. "An Alternative to the Privation Theory of Evil." Unpublished manuscript received April 22.

Samet, Irit. 2012. "On Pain and the Privation Theory of Evil," *European Journal for Philosophy of Religion*, Vol. 4, No. 1: 19–34. https://doi.org/10.24204/ejpr.v4i1.306.

Stump, Eleonore. 2009. "The Problem of Evil." In *The Cambridge History of Medieval Philosophy*, Vol 2, edited by Robert Pasnau, 773-784. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521762182.019.

Suárez, Francisco. 1989. *The Metaphysics of Good and Evil According to Suárez.* Edited and translated by Jorge J. E. Gracia and Douglas Davis. Hamden: Philosophia Verlag.

Swenson, Adam. 2009. "Privation Theories of Pain." *International Journal for the Philosophy of Religion,* Vol. 66, No. 3: 139–154. https://doi.org/10.1007/s11153-009-9202-4.

Williams, Thomas. 2005. "The Doctrine of Univocity is True and Salutary," *Modern Theology* Vol. 21, No. 4: 575–85. https://doi.org/10.1111/j.1468-0025.2005.00298.x.

Published Online First: February 28, 2023

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.65763

Schleiermacher and the Transmission of Sin: A Biocultural Evolutionary Model

HELEN DE CRUZ
Saint Louis University
helen.decruz@slu.edu

JOHAN DE SMEDT
Saint Louis University
johanvdesmedt@gmail.com

Abstract: Understanding the pervasiveness of sin is central to Christian theology. The question of why humans are so sinful given an omniscient, omnipotent, and omnibenevolent God presents a challenge and a puzzle. One element of this puzzle is how sinful tendencies transmit in human communities. Here, we investigate Friedrich Schleiermacher's account of sin which we characterize as a biocultural evolutionary approach. That is, we propose that Schleiermacher conceives of sin as both biologically rooted and as culturally transmitted. We look at empirical evidence to support his account and use the cultural Price equation to provide a naturalistic model of the transmission of sin. This model can help us understand how sin can be ubiquitous and unavoidable, even though it is not biologically transmitted, and even if there is no historical Fall that precipitated the tendency to sin.

Keywords: Friedrich Schleiermacher, Biocultural evolution, Original sin, the Fall, Hamartiology.

1. Introduction: The Ubiquity of Sin

In *Orthodoxy* the lay theologian G.K. Chesterton (1909) attempted to defend orthodox Christianity against the Modernists, British theologians at the turn of the previous century who wanted to update Christianity in the light of science. To this purpose, they proposed to discard many orthodox theological concepts, including original sin. As Chesterton (1909, 24) remarked: "Certain new theologians dispute original sin, which is the only part of Christian theology which can really be

proved." He saw sin as an obvious, empirically indisputable feature of human nature, "a fact as practical as potatoes."

We agree with Chesterton that original sin can be empirically grounded. However, doing so requires a re-examination of the underlying ideas and assumptions about what sin is, how it originates, and how it is transmitted. There is no consensus among Christian theologians on these points. As Oliver Crisp (2015) observes, Christian churches (except for the Oriental Orthodox churches)[1] universally accept the Chalcedonian definition of Christ's two natures as human and divine, but there is no such general agreement on sin. Still, sin plays a key role in Christian thinking and practice, particularly because it prompts the need for divine grace and salvation. The fact that humans invariably fall into sin, in spite of an all-powerful, omnibenevolent, and omniscient God (i.e., an omni-God) presents an enduring puzzle to theologians and philosophers of religion. Why do we sin if an omni-God could have easily created human beings who are sinless? Theodicies that appeal to human free will do not, as a primary aim, account for the pervasiveness of sin. They might explain why we sin, but not why it seems impossible to refrain from doing so.

In the absence of a clear Christian consensus on sin, we look at Augustine's account, which is highly influential in western Christianity, as a useful point of reference. According to Augustine and views influenced by him, our tendency to sin (original sin) is due to the Fall. We biologically inherit original sin from our ancestors and pass it on to our descendants. However, the Fall poses theological as well as empirical problems, leading theologians across the centuries to question Augustine's model and to propose models of original sin that do not depend on a historical Fall, and that do not rely on a biological model of the transmission of original sin. In this paper, we examine Friedrich Schleiermacher's account of the transmission of sin as an attractive alternative to the influential Augustinian account. Section 2 gives a brief overview of the doctrine of original sin with a focus on Augustine, and we note why, in spite of recent and older criticisms, it remains influential. Section 3 sketches Schleiermacher's account of sin, which we term "biocultural" due to its reliance on both biological and cultural factors. In section 4, we use the method of science-engaged theology (see e.g., Perry & Leidenhag 2021) to provide an empirical basis of a Schleiermacherian model, and in section 5, we present an evolutionary account for this, grounded in mathematical modelling.

[1] The Oriental Orthodox churches adhere to Miaphysite Christology, which says that Christ has one united nature. This is a form of Trinitarianism that rejects Christ's two natures. Denominations include the Coptic Orthodox Church of Alexandria, the Syriac Orthodox Church of Antioch, and the Armenian Apostolic Church.

2. The Doctrine of Original Sin

"Sin" (and its equivalents, e.g., "hamartia" in New Testament Greek) is a Christian theological concept.[2] A non-religious person might say murder or theft are morally wrong, but not that they are also *sinful*. Theologians argue that such wrongdoings are sinful because they constitute disobeying or not acknowledging God's commandments.[3] The scriptural basis for sin is the Genesis 3 narrative: the first man and woman disobey God by eating from the forbidden fruit from the Tree of Knowledge of Good and Evil. Their transgression not only causes sin to come into the world, but also death, patriarchy, pain in childbirth, agriculture, and an uneasy relationship with snakes. Other scriptural sources, such as letters by Paul and James, provide further clues that sin has a negative impact on human flourishing, freedom, and relationships. For example, we can be enslaved by sin (Romans 6: 16–17), which tarnishes our ability to live in orderly communities (James 3: 16).

Taking scripture as a common starting point, different Christian traditions have outlined diverging conceptions of sin. Augustine (354–430) formulated an account of original sin that is so influential that it is often termed "the doctrine of original sin."[4] The doctrine of original sin is in fact composed of several related doctrines[5] (see Couenhoven 2005 for review). Central is the historicity of the Fall: Augustine proposed that a historical Fall, precipitated by the "primal sin" (or first sin, *in casu*, eating from the Tree of Knowledge of Good and Evil), caused the subsequent sinfulness of all of humanity. One of the many results of the Fall is "original sin" (*peccatum originatum*), the human propensity to inevitably sin. So, the primal sin, together with some mechanism of inheritance, explains why humans are in the state of original sin: because our ancestors committed the first sin, we're all in a state of sinfulness.

According to the Augustinian account, humans were in an original state of righteousness prior to the Fall. This means, among others, that they were *able* to refrain from sinning. By contrast, humans ever since are *unable* to refrain from

[2] We also see terms relating to "sin" in non-Christian traditions, but their meaning is different in each of these cases, so here we focus on the Christian concept.

[3] We will here not discuss the issue of whether divine commandments should be seen in the light of an autonomous or heteronomous morality (for an overview and historical contextualization of this debate, see e.g., Bertini 2017).

[4] The Augustinian account is the focus of some recent edited volumes and special issues (see e.g., Madueme & Reeves 2014, Cavanaugh & Smith 2017, De Cruz & De Smedt 2021, for collections on this topic).

[5] We will here for brevity's sake still refer to this bundle of doctrines as "the doctrine of original sin."

sinning. This is because of far-reaching metaphysical and moral consequences of the Fall, not only for humanity but for the universe as a whole. As a result of the primal sin, original sin is present in all humans, including newborns. It not only instills in them the inevitable propensity to sin (termed "original corruption"), but also the original guilt associated with the first humans.

Augustine developed this account of original sin in part as a response to the British theologian Pelagius (354–418) and his followers, who believed that humans are inclined to sin because they are influenced by a sinful environment, but that, in principle, they could refrain from doing so. Augustine, by contrast, thought sin was unavoidable, and that only God's grace could save humanity. His account of sin is thus closely tied to the necessity of grace, which is why some contemporary authors (e.g., Smith 2017, Green & Morris 2020, Madueme 2021) prefer the Augustinian account, and have defended it over modern alternatives in spite of its lack of fit with modern science. For example, paleoanthropology does not provide any evidence for a historical Fall, original righteousness, or a single ancestral pair for all of humanity (De Cruz & De Smedt 2013, De Smedt & De Cruz 2020).

In the Christian theological literature, there are three proposed mechanisms for the transmission of original sin: biological, federalist, and social. We can see all three in Augustine, but he is mostly associated with the biological transmission model. According to this model, sin is transmitted through sexual reproduction. At the basis of this lie seed principles, an ancient understanding of how reproduction (and other growth processes) work: male seeds are implanted in female wombs, and these male seeds carry the seeds of the next generations, in a Matryoshka doll-like open-ended series. Augustine thought he observed sinful tendencies even in infants, citing this as evidence that we have the tendency to sin inherited from birth: "I have observed and experienced a little one expressing jealousy. Though he was not yet capable of speech, he glared, pale with envy, at his sibling at the breast . . . Surely one cannot call it 'innocence' when a baby prevents his sibling—who is completely dependent for care, and stays alive only because of that one source of sustenance—from having a share in the plentiful, abundant flow of milk" (Augustine, 4th century CE [1961], book I, chapter 11, 21). To Augustine, the phrase from Romans 5:12, that we have all sinned in Adam, should be taken literally: humans all originate from Adam's body and were already physically present there as seeds (Lamoureux 2015).

While the biological model is associated with the doctrine of original sin, we will here briefly note two alternative minority traditions. One is the federalist position, which was developed in Calvinist theology in the 19th century. According to this view, there is no real transferal of properties in the transmission of original

sin from the first pair to their progeny. Rather, God arranges things in such a way that their progeny is treated as if they had sinned and as if they had the first pair's guilt. The federalist position treats Adam as the federal representative of all of humanity, who by sinning has implicated the rest of humanity. This model has to grapple with two monumental problems: the problem of injustice (how is it just that I am guilty of Adam's sin?) and plausibility (why would God arrange the world this way?) (Crisp 2006). Another alternative is that sin is not biologically, but socially transmitted: sin is transmitted through social learning in human communities. Examples of theologians who have developed social accounts of the transmission of sin include Walter Rauschenbusch (1917), Stephen J. Duffy (1988), and more recently, Matthew Croasmun (2017). We will explore Schleiermacher's social model (or more accurately, his biocultural model that contains elements of both the biological and social model) in the next section.

The doctrine of original sin (with the positing of a historical Fall, single couple, original guilt and corruption, and biological transmission of sin) has been influential in western theological traditions, including the major Protestant confessions and much of Protestant theology, as well as in Roman Catholic thought. Its popularity derives from its perceived ability to solve the following dilemma: how can bad things happen if everything is created and willed by an omni-God? Either God created evil, which is problematic, or God didn't, but then we have to postulate an independent evil force (e.g., the devil), which is equally problematic. In Augustine's view, evil in itself is not a causal independent force, like, say, gravity is. Rather, for Augustine, evil is a *deficiency,* the privation of good. This privation expresses itself in humans desiring to do something bad. Since evil is not something that independently exists, humans must be conflicted or mistaken when they desire to do bad things. Thus, Augustine sidestepped the dilemma by claiming that God did not cause sinful inclinations in us, and that there is no independent evil force tempting humanity. His account requires that we posit a change in human nature (pre-Fall to post-Fall): sin changed humanity and the rest of the cosmos so profoundly that there's a discontinuity between our pre- and postlapsarian condition (Pedersen 2020). In this view, because we inherit not only sin but also guilt through biological means, God is not responsible for our tendency to sin, and did not cause it to happen.

The emphasis on a historical Fall by an original human pair has long been a sticking point among theologians. Particularly in recent decades, some theologians and authors in the field of science and religion (e.g., Venema & Knight 2017, Schneider 2012) have rejected the doctrine of original sin because it is not compatible with science. Among many other problems, the current genomic

evidence does not support the existence of a single ancestral human pair, but rather points toward small ancestral breeding populations ranging from a few hundred to a few thousand individuals. There are also enduring theological worries about the transmission not only of sin, but also of guilt. Why would an omni-God create a world in such a way that the sin of the first couple became the sin of all their descendants? This puzzle remains, regardless of whether one accepts the traditional biological transmission of original sin by Augustine, or the Reformed federalist interpretation of original sin. As the Reformed theologian Benno van den Toren (2016, 13) wonders, "How can a just God attribute the sin of the first couple to all their offspring?" For this reason, some theologians have formulated revised versions of the doctrine of original sin, such as Crisp's (2015) moderate Reformed doctrine of original sin, which denies the transmission of guilt along with sin. Thus, on scientific and theological grounds we have reasons to re-evaluate the doctrine of original sin.

3. Schleiermacher's Biocultural Account of the Transmission of Sin

Social accounts of the transmission of sin present an attractive alternative to the doctrine of original sin. They do not require a historical Fall which may make them more easily compatible with scientific accounts of human origins. They have early roots in Paul's Letter to the Romans, which describes sin as an emergent property that originates from individual sinful acts. Paul posits a choice: we can either be part of a *body of sin* (the community of sinners) or of the *body of Christ*, the church (e.g., 1 Cor 12, Rom 6). Paul's concept of "body" as a metaphor for a community of people is inspired by Stoic philosophy. Stoicism was a philosophical tradition that permeated the culture at the time; recall that Paul and Seneca were contemporaries. For Stoics, the only things capable of acting are bodies (*soma*), so if we are affected by something, be it our emotions or social forces, they must constitute a body too. Sin, in order to affect us, must therefore be a body. In Paul's view, we are part of a social body of sin, which means that we participate in the communal practices that jointly constitute sin (Croasmun 2017, 112–113). In Paul's writings, the body of sin gets an almost agential character. Sin is a slave master, which controls humans (Rom 6:6), rules our bodies by directing our passions (Rom 6:12), and uses our body parts as instruments (Rom 6:13). Paul's account of sin seeks to emphasize that everyone needs salvation: the pull of sin is so powerful that we cannot free ourselves from it (Green 2017).

However, Paul does not outline a mechanism of how sin would work as a social agential force. One such detailed proposal was drawn up by the Prussian

theologian Friedrich Schleiermacher (1768–1834), who developed an intriguing alternative to the doctrine of original sin. As we'll see below, Schleiermacher pays attention to both biological and cultural elements of original sin, which makes his model an excellent candidate for an empirically-informed account of sin. We focus on the presentation of his ideas in his monumental dogmatic theology, *Christian Faith* (*Der Christliche Glaube,* CG, 1830 [2016]), which aimed to put dogmatic theology on firm grounds, fitting it within the larger Enlightenment project of critically analyzing traditional theological concepts in the light of the latest philosophy and science.

The most eye-catching feature of Schleiermacher's account is that does not invoke a historical Fall. He rejected the Fall as well as the fundamental "alteration of human nature that has arisen by means of a first sin committed by the first human beings" (CG §72, 442).[6] While some contemporary theologians formulate non-lapsarian views because the Fall is not easily reconciled with evolutionary theory, paleoanthropology, and other scientific disciplines, Schleiermacher's primary concern was theological. He believed that the Augustinian doctrine of original sin, with its emphasis on a historical Fall, makes no theological sense. Original sin explains why we currently tend to sin, but it doesn't explain why the first humans did. Augustine did not appeal to innocence or gullibility (as some other early Church Fathers did, such as Irenaeus), so the first pair sinned knowingly: given their original righteousness, their capacities for reason were not clouded by original sin, as Schleiermacher emphasized. They went in clear-eyed, making their mistake all the more puzzling. It is more parsimonious to say that the tendency to sin is part of human nature all along. Schleiermacher explicitly appropriated original sin as a part of our psychological makeup, "a susceptibility imparted to every individual" (CG §70, 419; Wyman 1994, 233–234). Thus, we no longer need to posit a large cognitive difference between humanity pre- and post-Fall. Rather, in Schleiermacher's view, each person falls individually.

Without a historical Fall, Schleiermacher still has to explain why we all seem to be in this condition of fallenness (CG, §71). If not through a historical event, why are we this way? Not positing a Fall has the risk of assuming that humans would, in principle, be able to escape the condition of fallenness. That would make grace redundant, a Pelagian position that is theologically unorthodox. Schleiermacher wanted to keep the necessity of grace. He explained how we can have solidarity in sin without a Fall by conceptualizing sin as part of human nature. He aimed to

[6] We refer to the paragraph in *Christian Faith*; the page number refers to the 2016 translation by Tice et al.

solve a paradox: the sins we commit do not originate in us, yet they are something we are fully responsible for. His explanation contains three components: God-consciousness and our falling from it, psychological tendencies rooted in our biology that cause us to sin (the seed of sin), and the cultural transmission of sin in social contexts. Because Schleiermacher invokes both biological features and cultural transmission, his account can be characterized as "biocultural." We use this term rather than more common alternatives in the recent literature such as "gene-cultural coevolution" or "dual inheritance model" because it is more general, and avoids the anachronism of an early nineteenth-century publication that talks about genes. Moreover, "biocultural" points to two kinds of factors that Schleiermacher would have been familiar with: biology and culture that interact with each other.

God-consciousness constitutes the biological component of Schleiermacher's account of sin. It is central to his dogmatic theology and ethics. It is a thick theological concept that includes our reflective awareness of the self ("self-consciousness"), awareness of how one is socially situated ("species-consciousness"), and a feeling of absolute dependence on God. Elaborating on earlier work where he argued that religion originates in feeling (e.g., Schleiermacher 1799 [2006]), *Christian faith* examines what this feeling consists of, and how it originates. God-consciousness arises as a spontaneous product of our creaturely, biological nature. It starts out as a feeling of being dependent on other creatures, which arises from the push and pull of us acting on the world and the world acting on us—which Schleiermacher terms "relative dependence." We are always dependent on our environment to sustain ourselves. We are enmeshed in a web of interdependence: the air we breathe, the soil we stand on, the creatures we eat, the humans we collaborate with.

Once we become aware that there is a Being that underlies this whole, we become aware that God "is designated as the one grounding this interconnected being in all its diverse parts" (CG, § 30.1, 183). This awareness is God-consciousness. It is a kind of self-transcendent sense where we realize our absolute dependence on God. Importantly, when we reach this stage (evolutionary and developmentally), earlier forms of consciousness do not fall away. Nor are they necessarily bad: they are necessary to secure our basic means of existence (we need to eat, associate with other people, etc.) and God-consciousness builds on our creaturely awareness of our surroundings. Yet, it is in the mismatch of God-consciousness and the lower forms of consciousness that Schleiermacher locates sin. Sin is an inability to integrate our religious self-consciousness with our social and bodily self-consciousness (Nelson 2009). Because we are so caught up with our

lives—we're happy to eat and to be with friends, we're sad when a loved one gets ill or dies—we fail to see that God made a good world. Our sensuous nature makes us lose sight of the fact that everything God created is for the good. This tends to obscure the fact that we are absolutely dependent on God.

God-consciousness is crucial for Schleiermacher's account of sin, because without being aware of the good, we would not sin, "sin exists only insofar as a consciousness of sin also exists" (CG §68, 410). God-consciousness only appears in humans, and is not present in other animals. Thus, only when humans become God-conscious are they able to sin. To this end, Schleiermacher sketches a proto-evolutionary account. Although *Christian Faith* was published some three decades before Darwin's *On the origin of species* (1859), evolutionary ideas (at the time called *transmutationism*) circulated widely, and Schleiermacher had access to these. Most pertinent to his cultural sphere was the work by Gottfried Reinhold Treviranus, notably his *Biologie, oder Philosophie der lebenden Natur für Naturforscher und Aerzte* (*Biology, or philosophy of living nature for naturalists and physicians*), which was published in Göttingen in 1802. This work went through six editions from 1802 to 1822; it was highly influential, and made the term "biology" the standard term to refer to the study of living things (previously, it was "natural history"). Treviranus presented a transmutationist theory, based on fossil evidence. Moreover, he portrayed nature as a web of interdependent beings, an idea that recurs in Schleiermacher's work (De Cruz 2022).

Daniel Pedersen (2017: 35–39) argues that Schleiermacher knew about, and accepted transmutationism. A key passage in *Christian Faith* is the following: "we pretty much know, regarding our world that species have existed that are no longer present and that present species have not always existed" (CG §46 (postscript), 254). The only biological theory at the time that incorporated species going in and out of existence was transmutationism. Moreover, §5.1 presents an explicitly transmutationist idea about human origins, where Schleiermacher connects human cognition to that of extant non-human animals:

> Suppose that we go back to the initial, more obscure period of the life of human beings. Everywhere therein we would then find the animalistic life to be almost alone predominant, but the spiritual life would be still entirely suppressed. As a result, moreover, we would have to imagine the state of a human being's consciousness in that obscure period to be very much akin to that of the lower animals. (CG, §5.1, 28)

Without God-consciousness, what we term "sin" would only be a "self-focused activity of flesh" (CG §67, 405), such as gluttony and lust, which originate prior to the emergence of God-consciousness:

> If God-consciousness has not yet developed, there is also not yet any resistance to it ... in the future this self-focused activity of flesh will indeed become resistance to spirit [i.e., sin], but beforehand it cannot actually be observed as sin, but only as a seed of sin at best. (CG §67, 405)

Here, Schleiermacher hints at the tendencies in our biological makeup that in other animals would not be called sin but "seed of sin," such as, for example, lust for power or violence against conspecifics. We inherited this seed of sin from our hominin ancestors. Only when humans became God-conscious, these inclinations could become actual sin. As Derek Nelson (2009, 136) puts it, our biologically inherited sensible self-consciousness is sometimes out of sync with God-consciousness. As a side-effect of how creation works, "God-consciousness outpaces, at times, the gait of the sensible self-conscious will." God-consciousness arises both in individuals and in social structures (communities), which will lead us to the social model below. The puzzle then arises why humans individually ignore their God-consciousness, and sin as a result of this. Since Schleiermacher explicitly denies a historical Fall, and therefore denies that human nature was fundamentally altered, he cannot invoke a biological tendency to sin that resulted from whatever might have happened during or right after the Fall.

To solve this puzzle of why humans sin, Schleiermacher appeals both to our biological makeup and to cultural transmission. Our biological makeup gives us God-consciousness so we have an awareness of the good. When our animal self-consciousness and our God-consciousness don't align, we sin. This happens because we are born in cultures that have plenty of sin. We can learn to express our morality fully in our cultural communities, but this is also where sinful tendencies are transmitted. To make an imperfect modern analogy: we may have genes that code for violent behaviour, which we share with our closest living relatives, the chimpanzees. Some human communities are more or less pacifist and do not allow for public expressions of violent behavior (e.g., the Amazonian Pirahã, the Semai of Malaysia). In these communities, the particular seed of sin that one might call ancestral violence is not transmitted. In other communities (e.g., the historical states of Prussia and Sparta), violence is publicly condoned, encouraged, and thus culturally transmitted.

This raises the question of why similar sinful tendencies (e.g., stealing, violence, xenophobia) arise in human cultures. The reason for this is our biological tendencies, namely our (biologically situated) God-consciousness, and the "seeds of sin" which are also biological and which make us more susceptible to some cultural influences than others, as we will show below. Humans are born into faulty communities with sinful ways of engagement, "the sin of each individual has its source in an earlier existence above and beyond one's own existence" (CG §69, 414). We inevitably pick up some of these ways of engaging within our communities through growing up in a culture where we learn the norms as children, and thus we sin: "any given mode of education is grounded in leanings and experiences that have preceded the existence of the one who is to be educated" (CG §69, 414).

Schleiermacher's social transmission account aims to resolve the tension of sin as something that is unavoidable—part of our evolved human nature and part of human cultures—and also as something we are personally responsible for. We exert some degree of control over the cultural ideas we inherit. For example, across cultures we can find various implicit biases, such as negative attitudes toward people of different ethnic groups, genders, or social classes. Even if one is aware of such biases and tries to withstand them, one will still often fall foul of them (see Vicens 2018 for a conceptualization of implicit bias as sin). Because cultural transmission is such a powerful force in shaping human behavior, we will inevitably end up adopting some socially transmitted sinful dispositions, such as the endemic racism in the United States.

How does this social transmission of sin relate to original sin? Note that while Schleiermacher locates sin in social processes, he does not think they ultimately originate there. Rather, sin originates in our innate sense of the good (tied to God-consciousness), that we deliberately, and each individually, fall from. Because we are members of cultural communities that are imbued with sinful inclinations and behaviors, we will tend to fall and sin (CG §72). Thus, we each fall *individually* by denying our God-consciousness.

The social transmission of sin can be seen in a larger cultural evolutionary framework, where humans are born in sinful communities. We acquire these sinful behaviors through social learning, and thereby set up the conditions for our descendants to inherit them: "What appears from birth as the susceptibility to sin of a generation is conditioned by the susceptibility to sin of earlier generations and itself conditions the susceptibility to sin of generations yet to come" (CG §71, 429). We are all implicated in sin, not only because of the actual sins we commit, but also because we transmit sinful beliefs and behaviors, and expose others to them, not

only young children but also fellow adult citizens. In this way, Schleiermacher repudiates the Augustinian notion of original guilt: we are not guilty of a sin that our ancestors committed (CG §72, 447). At the same time, we are responsible for the sins we commit and help perpetuate, even though they may not have originated in us.

Since human nature was not fundamentally altered through a historical Fall (as Augustine claimed), Schleiermacher's biocultural account of sin locates the origin of evil with God. After all, sin is a result of our biological makeup and social environment. Individual transmission of sin is a result of how God created us. As Pedersen (2020, 142) points out, this account "makes God the author of sin and the cause of evil. Indeed, Schleiermacher does not merely imply this, but explicitly claims it as a consequence of his account."

Like Augustine, Schleiermacher sees sin as a privation of good, not as an independent quality. However, he does see God-consciousness as something that is fully part of human nature. We don't just transmit sin, but also virtues socially. We use reason to discover ethical principles, and we participate in social life where we overcome our individual limitations through sharing, collaboration, and forming communities: we can flourish and morally improve by participating in human pursuits, including science, religion, politics, and the arts. Notably, Schleiermacher thought that we could fully realize ourselves in a plurality of communities and institutions (Schleiermacher 1812–1813 [2002], §61, §97). In our human-made institutions we can bring our lives closer to our moral ideals. Living in these interdependent communities, and using our reason and our innate sense of the good we can devise moral norms, which we can then choose to follow, or not to follow. This discovery of moral norms through reason is the flip side of the social transmission of sin. Sin is the social transmission of the denial of God-consciousness, as we saw above. Virtue is an acknowledgment of God-consciousness, and is equally socially transmitted.

Schleiermacher's account is still relevant for discussion on the naturalistic origins of sin because it acknowledges both biological tendencies (God-consciousness and the seed of sin) and cultural factors in the transmission of sin. Given recent work on the importance of both biology and culture in the transmission of moral norms and behavior (as we will review below), we can reassess his theory in the light of contemporary empirical evidence.

4. Evaluating Schleiermacher's Biocultural Account of the Transmission of Sin

If we assume, as Chesterton (1909, 24) did, that empirical evidence is relevant for accounts of the transmission of sin, what kind of evidence would be germane? John Perry and Johanna Leidenhag (2021) use the term "science-engaged theology" to denote a method which aims to solve specific puzzles that arise on the intersection of science and religion. Rather than asking if Christianity (or religion) as a whole can be reconciled with modern science, it is more productive to ask if a specific theological question can be answered using the tools of a specific scientific (sub)discipline. In our case, when we consider the transmission of sin, we should ask which scientific disciplines might be relevant.

At the turn of the previous century, when modernist theologians such as Frederick Tennant (1866–1957) considered the question of sin, e.g., in *The origin and propagation of sin* (1902) and *The sources of the doctrines of the Fall and original sin* (1903), the focus was on biology, specifically on evolutionary theory. At the time, some forms of evolutionary theory were generally accepted by the scientific community, and the animal ancestry of humanity was beyond reasonable scientific dispute. The many successes of evolutionary theory, finds of early human fossils (Darwin 1871), striking similarities between human and primate anatomy (e.g., Huxley 1863), and between human and primate facial expressions and emotions (Darwin 1872) led modernist theologians and church leaders to re-evaluate Christian theological concepts such as original sin. An example of the spirit of the times were the sermons on evolution (dubbed "Gorilla sermons" by the British press) by the Anglican bishop Ernest Barnes in the 1920s and 1930s. In these, he denied the Fall and original sin, as concepts that were both outdated and not in line with the sciences (Bowler 2007). By contrast, Tennant (1902) sought to salvage some elements of the doctrine of original sin. He recognized humans have an (apparently) inescapable tendency to sin, and attributed this to our animal ancestry. In his view, we inherit our tendencies for self-preservation from our primate ancestors; these inclinations only become sinful once we become morally aware. Rather than falling down, we fall up. But as we become morally aware, we fail to live up to the potential that we have as moral agents. The contemporary theologian Patricia Williams (2001) defends a similar biological model.

While biological evolution is certainly relevant for the theological puzzle-solving of original sin, we will cast a wider net. There is an increasing recognition that cultural practices have a large influence on human behavior. Biology and culture are intertwined throughout prehistory and history—a phenomenon variously referred to as gene-culture co-evolution or dual inheritance. Cumulative

culture explains the demographic success of humans, able to colonize every landmass (Dean et al. 2014). We cannot explain how humans behave by genes (and genetic evolution) alone. We need models of cultural transmission to understand differences between human communities and to explain recurrent patterns of human behavior (Richerson and Boyd 2005; Henrich 2018). In the light of this, we will now evaluate two aspects of Schleiermacher's account: the emergence of God-consciousness and the social transmission of sin.

4.1. Empirical Evidence for the Emergence of Sin Through God-Consciousness

As we have seen, Schleiermacher proposes that sin can only occur when we have some form of moral awareness, which he situates in his thick theological concept of God-consciousness. What would constitute empirical evidence for God-consciousness? After all, it is a theological, not a scientific concept. However, for science-engaged theology it suffices to explore scientific concepts and empirical findings that shed light on, and are compatible with the theological idea of God-consciousness. It does not require full scientific proofs of theological concepts, because these are impossible. Recall that Schleiermacher saw God-consciousness as part of human nature, and as something that arose in the course of human evolution. We can ask when beliefs in gods arose, as this sets up the necessary conditions for sin to arise (no sin without God-consciousness).

Schleiermacher links religion and morality in his concept of God-consciousness. Across cultures, people perceive a relationship between religious beliefs and morality. In the cognitive science of religion, there are several hypotheses that aim to explain this relationship. The *supernatural punishment hypothesis* (SPH) states that the threat of punishment by supernatural agents inhibits self-interested behavior and promotes cooperation. Supernatural agents, such as ancestors, place spirits, gods, and ghosts tend to have privileged knowledge of human affairs, often coupled with extraordinary capacities such as controlling the weather (White et al. 2022). Because of their perceived special properties, humans fear supernatural punishment, for example, through their control of the weather supernatural agents can initiate a drought to punish misbehaving people. This leads them to behave better even if no-one is watching, even when there are no secular punishment systems in place.

One specific version of the SPH is the Big Gods hypothesis, developed by Ara Norenzayan (2013). Big Gods are a special category of supernatural beings: they are omniscient (or at the very least, very knowledgeable), powerful (very often they have created the universe), next to moralistic and punitive. A clear example is

the Abrahamic god, but Big Gods can be found outside of the Abrahamic traditions, for example, in many strands of Hinduism. Norenzayan (2013), noticing a strong correlation between belief in Big Gods and large-scale societies, proposes that it is thanks to belief in such gods that humans were able to associate in larger groups, giving rise to large-scale societies. In small-scale societies humans exert social control through reciprocity, shunning, and other mechanisms. In large-scale societies this is no longer possible, because we no longer know all the members of our group personally. According to Norenzayan, people fearing that god (or the gods) are watching them and would punish them, behave more cooperatively. He identifies Göbekli Tepe as an early testimony to Big God beliefs. This site with a probably religious function was built by hunter-gatherers in present-day Turkey dating to about 11,500 Before Present. It consists of massive stone pillars arranged in circles and carved with animal imagery.

Norenzayan holds that Big God beliefs precede the emergence of large-scale societies. However, the question of what came first is an enduring topic of lively debate among cognitive scientists of religion. Harvey Whitehouse and colleagues (2021) argue that belief in Big Gods only arose after large-scale societies became established. However, their paper (originally published in *Nature*) had to be retracted due to errors in their analysis.[7] In any case, the Big Gods account predicts a late origin of God-consciousness.

An alternative version of the SPH is *broad supernatural punishment* (Watts et al. 2015): across cultures, there is a broad range of supernatural beings, including localized spirits, but also non-agentive forces such as karma. These supernatural entities can also inflict punishment for moral reasons, thereby facilitating cooperation and reducing cheating in human societies. A third alternative (Purzycki et al. 2022) is the *moralization bias.* Humans conceptualize other humans as interested in moral behavior. Since we use the same cognitive apparatus when we represent the minds of gods and those of humans, humans represent the minds of supernatural agents similar to human minds. Because of this, supernatural agents are perceived as also being interested in moral behavior (moralization bias). Purzycki et al. (2022) recruited participants from 15 different cultures. They

[7] The main problems with Whitehouse et al.'s paper concern the quality of the dataset (Seshat: Global History Databank) on which the analysis is based. Due to serious issues with the coding of the dataset (sloppiness, mistakes, and omissions), their paper had to be withdrawn. However, the authors subsequently cleaned up the dataset and rewrote the paper. They insist that their conclusions still hold. A new version of the paper (not published yet at the time of writing) can be found here: https://osf.io/mbnvg/

concluded that the moralization bias is widespread across cultures. It did not matter whether their participants believed in Big Gods (e.g., Shiva, the Christian God) or less powerful supernatural agents such as forest spirits. The participants in this study also associated supernatural beings who were morally concerned with members of their community cooperating more and cheating less.

The latter two hypotheses allow us to potentially push back in time the origins of God-consciousness, though it is difficult to pinpoint archaeologically when it would have emerged. Strong archeological clues for religious beliefs emerge during the Late Pleistocene, particularly in the form of symbolic artifacts, usually consisting of a mix of human and animal body parts. The oldest depictions of such therianthropes date from 43,900 Before Present. They are part of an elaborate rock art panel that depicts therianthropes hunting wild pigs and dwarf bovids, from Sulawesi, Indonesia (Aubert et al. 2019). The oldest figurine depicting a therianthrope is the so-called lion-man, a mammoth ivory figurine with a human body and a cave lion's head, dated to about 39–41,000 Before Present. It was found in Hohlenstein Stadel cave (Southwestern Germany), and stands at about 31 cm (Kind et al. 2014). Some authors (e.g., Lewis-Williams 2002) have argued that patterns painted on cave walls and on some mobiliary art, such as hand stencils and collections of dots, are evidence of altered states of consciousness and shamanic practices. The oldest hand stencil is 39,900 years old and was discovered in Sulawesi, Indonesia (Aubert et al. 2014). This archaeological evidence may point toward belief in supernatural beings. This allows us to put the emergence of God-consciousness at about 44,000 Before Present.

4.2 Empirical Evidence for the Social Transmission of Sin

We will now examine Schleiermacher's social transmission of sin by looking at evidence for social learning in humans. Growing empirical evidence suggests that children acquire cultural (including moral) norms through social learning. Toddlers show "promiscuous normativity," i.e., they can infer from a single instant of how something is done a norm of how it *ought* to be done. This is not just about moral norms, but about any cultural norms, for example, dress code or how to play a game (see e.g., Schmidt et al. 2012). They also enforce social norms on other group members.

Children aged four to six react with disapproval toward people who do not conform to the norms of their community, even if these are merely invented communities for the purposes of a psychological experiment. For example, Roberts et al. (2017) show that American young children disapprove of members of

communities that have been made up ad hoc in a lab that do not adhere to arbitrary norms, such as eating one particular kind of berry rather than another. This tendency declines in older children and disappears in adults. A replication of this experiment in China (Roberts et al. 2018) shows that the tendency to disapprove of norm violations is present in both Chinese children and adults.

Young children are also directly influenced by the behaviors of their peers or parents, for example, both US and Indian children share less after witnessing their parents being stingy (Blake et al. 2016). This research indicates that young children are to an important extent guided by the norms of their community and the behavior of their elders when deciding how to behave. Throughout human evolution, we see the supreme importance of cultural transmission in various hominin species. Stone tool technology, even Oldowan, the simplest one, requires a level of social learning not seen in extant nonhuman primates, presumably involving processes such as explicit teaching and imitation (Morgan et al. 2015). Unfortunately, moral norms are not preserved in the archaeological record, but it is plausible that the transmission of such norms was also cultural in the distant past. Over time, this cultural transmission of norms, which may have involved diverse domains (e.g., prohibitions on incest, obligations to care for group members) could have effects on human genes as well, as they favor the selection of altruistic behavior. Groups with altruistic behavior would do better than groups without it, leading to a cultural group selection of norm-governed moral behavior (see e.g., Tomasello 2016).

If sin is transmitted socially, as Schleiermacher and other adherents of the social model argue, then behaviors we morally disapprove of can be transmitted socially and this is what the empirical record shows. A growing body of empirical evidence suggests that people are very sensitive to their social environment and to the perceived social approval or disapproval of their behaviors. Actions such as bullying or harassment are not socially learnt, rather, people who bully and harass look for social cues to gauge to what extent their behavior is socially sanctioned. In a series of studies on bullying in secondary schools in the US, Betsy Paluck and co-authors (2016) enrolled students who were well-liked by their peers to take part in anti-conflict interventions. These popular students were asked to become the public face of opposition to bullying, and to spread anti-bullying messages among their peers. As a result of their intervention, reports of bullying incidents declined by 30% at these schools. Students were apparently less willing to bully when they perceived that popular peers found it unacceptable. Munger (2016) used a Twitter study with bots that automatically responded to white users who employed racial slurs with the following message “Hey man, just remember that there are real

people who are hurt when you harass them with that kind of language." Targeted users reduced racial slurs for up to a month after this intervention, but only if the bot was presented as a white man (i.e., a perceived ingroup member) and especially if the bot had a high number of followers (i.e., was perceived as socially influential). Blanchard et al. (1994) conducted a similar study in a live setting on campus where students who heard someone condemning racism expressed stronger anti-racist sentiments, whereas those who heard racism condoned expressed weaker anti-racist views.

Moreover, several studies have probed the uptick of hate crimes following two election results in the US and the UK in 2016: the US presidential election and the Brexit referendum in the UK. Hate crimes included spikes in intimidation, harassment, property damage, and hate speech against foreigners, Muslims, disabled and LGBTQ+ people (see e.g., Devine 2021). Both political campaigns significantly used negative rhetoric against aforementioned minority groups. Paluck and Chwe (2017, 990) contend that this increase in hate crimes is not because perpetrators "learn" xenophobia, Islamophobia, etc. from the media or the political candidates, but rather that "potential perpetrators are encouraged to act by the fact that Trump garnered votes and now holds the highest office. They infer from this that they have a better chance of escaping social and legal sanction than before his election." Similarly, in the UK the perpetrators of hate crimes felt supported by the rhetoric from the government and popular press and media that stigmatized entire communities, such as vans hired by the Home Office exhorting illegal immigrants to go home or face deportation (Burnett 2017).

Taken together, the developmental psychological and sociological evidence points to the social transmission of sinful tendencies and behaviors within communities. As Schleiermacher already argued, as a child you are born in a community with sinful tendencies and behaviors that you absorb and that you will inevitably perpetuate and transmit to younger community members. The psychological and sociological evidence is compatible with Schleiermacher's model of the transmission of sin, and can illuminate it.

5. A Cultural Evolutionary Model for the Social Transmission of Sin

To get a sense of the evolution of cultural or moral norms within communities, we can use mathematical modelling. Though presently not often considered in empirically-informed theology, this can be a fruitful source for science-engaged theology. We can model the social transmission of sin using the cultural Price equation. The Price equation is devoid of any specific content, so it can be flexibly

used to model changes over time in traits that are inherited through cultural or genetic evolution, or a combination of both. In the cultural Price equation (El Mouden et al. 2014), cultural change is modeled as a result of cultural fitness. Certain cultural traits are adopted or not adopted by group members, and as a result their distribution in the population changes over time. To take an example, roller skating was at some point popular in the US and western Europe, and had then a high cultural fitness, but it subsequently got out of fashion, and now has a lower cultural fitness.

The cultural Price equation looks at the change of a cultural trait measured at a population level, denoted by $\Delta\bar{z}$, as shown in equation (1).

$$(1) \qquad \Delta\bar{z} = Cov(c,z) + E_c(\Delta z)$$

$\Delta\bar{z}$ is the change of a given cultural trait from one generation to the next, for example, roller skating, the wearing of hats, or religious attendance—the cultural trait is denoted with z. For our present purposes, $\Delta\bar{z}$ denotes changes in the degree to which a given hominin population conforms to specific moral norms. We assume here that hominins who are members of this community are aware of their moral norms, and attempt, in various degrees, to live up to them. People who succeed better at adhering to these norms have higher z-values, whereas those who do less well have lower z-values. The differential success between these individuals will influence the average value of z over time.

To calculate $\Delta\bar{z}$, the cultural Price equation adds up two terms, the first, $Cov(c,z)$ is the component that is concerned with selection, the second $E_c(\Delta z)$ is the component that looks at systematic biases that might influence or distort the selection process (Okasha 2006, 26–28). $Cov(c,z)$ indicates the co-variance of individual z-values and their relative cultural fitness c, which denotes the number of cultural descendants of particular individuals in the next generation, divided by the population mean number of cultural descendants. For example, suppose that Lucy who roller skates can get three other people to do it, and the mean number of cultural descendants for this trait is 2, then Lucy's relative fitness is 3/2, which is a higher than average cultural fitness. $E_c(\Delta z)$ is the expected cultural fitness-weighted change in prevalence of z over time, averaged in the population. For another example, take Ben who can convince his neighbor to also volunteer in the local soup kitchen. If the mean number of cultural descendants for this helping behavior is 2, then Ben's relative fitness is 1/2, which is a lower than average cultural fitness.

In genetic evolution this is assumed to be zero, because random mutations do not lead genetically-coded traits to systematically differ from parents to offspring (in other words, there will be changes but these will not be directional). But in cultural evolution this term is not negligible, as there are systematic distorting influences on how individuals absorb culturally transmitted information. For example, the ideal that all people are equal is often distorted by racist, classist, and gender biases, with as a result that some people are more equal than others. Okasha (2006, 28) reformulates the Price equation to get a better grasp on what $E_c(\Delta z)$ is as follows:

$$(2) \qquad E_c(\Delta z) = E(\Delta z) + Cov\ (c, \Delta z)$$

The first term on the right side of equation (2) represents individual distorting biases, also called cognitive attractors. These are tendencies of human minds to systematically distort culturally transmitted information, to selectively remember or misremember. For example, a cultural trait that is difficult to keep in memory or that is very hard to do without extensive practice will have more limited cultural success, simply by virtue of human cognitive limitations: fewer people are able to play *La cathédrale engloutie* by Claude Debussy (1910), than they are capable of playing *Chopsticks* (A. de Lulli, 1877), a simple waltz enthusiastically hammered out by piano novices across the world. The human mind not only distorts due to limitations in memory or ability, but also in the kinds of things that our minds respond well to. For example, some culturally transmitted concepts such as zombies or vampires (dead people who are somehow still alive) are counterintuitive, and therefore surprising and arresting, making them good candidates for cultural transmission (Boyer 2001). In the moral domain, moral ideas that accord well with our intuitive ideas about justice, fairness, and not doing harm are also more likely to spread culturally, hence the cross-cultural independent invention of the golden rule, do not do onto others what you don't want done to you (see Flanagan 2017 for an exploration of moral foundations and their role in the cultural evolution of moral norms). On the other hand, we also have evolved biases that dispose us to sinful cultural traits, which Schleiermacher referred to as "seed of sin." Social factors will either facilitate or impede these biases. For example, recently far-right ideologies thrive due to social contagion (Youngblood 2020). They are facilitated by evolved cognitive biases such as wariness of outgroup members, tendencies of men to dominate women, and identifying violence with courage.

The second term in equation (2) is of particular relevance to our present discussion, as it can potentially capture the cultural transmission of sin. Note that $Cov(c,\Delta z)$ is subtly different from cultural selection (recall, that is $Cov(c,z)$). Here, what is modeled is the extent to which people will adopt or abandon cultural traits based on the influence of group members who display such traits, i.e., the perceived cultural prevalence of the traits. The most common ways in which this happens is conformist bias and prestige bias. In conformist bias, people will uncritically accept certain cultural practices if they are perceived as ubiquitous. Prestige bias is adopting certain traits because influential or successful people practice them. In the previous section we have seen empirical evidence for conformist and prestige bias in the adoption of moral norms (e.g., do no bully) or certain problematic behaviors (e.g., hate crimes). Both of these biases are at work in the transmission of sin, or in resisting it. As we saw, high-status individuals (in high school, or on Twitter) have a large impact on reducing problematic behavior such as bullying and using racist slurs.

The cultural Price equation (1) predicts that the proportion of z will continue to change over time as long as there is covariance between the cultural fitness of individuals and the extent to which they possess z (their individual z-values). To go back to roller skating, as long as enthusiastic, skilled skaters are more likely to get others to take up the sport, the proportion of roller skaters in the population will change over time. Likewise, people who are moral exemplars will have more cultural descendants. For example, Confucius was and still is perceived as a moral exemplar by many East Asians, leading his specific moral ideas (Confucianism, also called Ruism) to be widespread across East Asia, and recently also in the west (Olberding 2011). To take another example, in contemporary hunter-gatherer cultures, generous hunters who share their spoils are more socially influential, and more likely to enjoy the altruism of others, compared to less generous hunters (Bird & Power 2015). As long as some people have more moral influence than others, or there is some form of covariance between a person's moral standing and her ability to influence cultural descendants, the moral views within a population (whether for good or ill) will be subject to change.

But the expected fitness change in a population based on factors other than cultural selection, $E_c(\Delta z)$, also plays a crucial role in equation (1). It includes both people's following of lower moral standards due to the perceived influence of prestigious individuals (e.g., indulging in drugs because famous musicians do it), but also due to a perceived prevalence (e.g., thinking it is OK to dodge one's taxes or be racist because "everyone does it"). However, this term also includes individual innovation, which changes the original information, but can lead to an

improvement just as much as a deterioration of the original culturally transmitted information. As an example for the good, take the arguments against slavery and against the trade in fellow humans by abolitionists such as John Wesley (1703–1791) and William Wilberforce (1759–1833). At the time, this was a cultural innovation as many people thought slavery (or at the very least the trade in slaves) was economically and morally acceptable. This accords with Schleiermacher's ideas that humans are able to resist at least some sinful tendencies (in this case, greed, racism, and dehumanization) and that they are able to leave a better legacy for future generations.

The cultural Price equation captures mathematically how the transmission of moral norms is the result of individual and collective biases and of individual innovation. As modeled here, cultural fitness is independent of genetic fitness. It is possible for someone to have great cultural fitness, i.e., many cultural descendants, but only a modest genetic inclusive influence, e.g., no genetic children. For example, Mother Teresa has enduring cultural influence and offspring in the Missionaries of Charity, a religious congregation she founded in 1950, while having no genetic descendants. However, gene-culture co-evolutionary theories predict that changes over time in z can lead to different selective pressures at the genetic level. To get a full picture of how this gene-culture co-evolution would work goes beyond the scope of this paper, and beyond the simple mathematical model outlined above. However, we have shown with this model how contemporary mathematical approaches to cultural evolution can illuminate theological ideas, viz. the transmission of sin.

6. Conclusion

In this paper, we have reviewed Friedrich Schleiermacher's biocultural model of the transmission of sin. We have shown through a science-engaged theological approach how diverse bodies of empirical evidence from a range of disciplines, notably social psychology, cognitive science of religion, and mathematical modelling, can be used to evaluate Schleiermacher's approach. From this we show that his attention to both the biological propensities of humans and their attunement to cultural norms provides a compelling explanatory framework for the transmission of sin. With this paper, we have shown that an alternative Schleiermacherian model of the transmission of sin can be empirically grounded, using a science-engaged theological approach.

Acknowledgments

The authors thank audiences in Lyon (France) and Pasadena, CA for their suggestions and feedback to presentations of this paper, as well as two anonymous reviewers for their insightful comments.

Bibliography

Aubert, M., Lebe, R., Oktaviana, A. A., Tang, M., Burhan, B., et al. 2019. "Earliest hunting scene in prehistoric art." *Nature*, 576(7787), 442-445. https://doi.org/10.1038/s41586-019-1806-y.

Aubert, M., Brumm, A., Ramli, M., Sutikna, T., Saptomo, E. W., et al. 2014. "Pleistocene cave art from Sulawesi, Indonesia." *Nature*, 514(7521), 223-227. https://doi.org/10.1038/nature13422.

Augustine. 1961. *Confessions* (R. S. Pine-Coffin, Trans.). London: Penguin.

Bertini, D. 2017. "Moral heteronomy. History, proposal, reasons, arguments." *Dialegesthai. Rivista telematica di Filosofia*, 17. https://mondodomani.org/dialegesthai/

Bird, R. B., & Power, E. A. 2015. "Prosocial signaling and cooperation among Martu hunters." *Evolution and Human Behavior*, 36, 389–397. https://doi.org/10.1016/j.evolhumbehav.2015.02.003.

Blake, P. R., Corbit, J., Callaghan, T. C., & Warneken, F. 2016. "Give as I give: Adult influence on children's giving in two cultures." *Journal of Experimental Child Psychology*, 152, 149–160. https://doi.org/10.1016/j.jecp.2016.07.010.

Blanchard, F. A., Crandall, C. S., Brigham, J. C., & Vaughn, L. A. 1994. "Condemning and condoning racism: A social context approach to interracial settings." *Journal of Applied Psychology*, 79(6), 993–997. https://doi.org/10.1037/0021-9010.79.6.993.

Boyer, P. 2001. *Religion explained. The evolutionary origins of religious thought*. London: Vintage.

Bowler, P. J. 2007. *Monkey trials and gorilla sermons. Evolution and Christianity from Darwin to Intelligent Design*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/9780674028609.

Burnett, J. 2017. "Racial violence and the Brexit state." *Race & Class*, 58(4), 85–97. https://doi.org/10.1177/0306396816686283.

Cavanaugh, W. T. & Smith, J. K. A. (Eds.) 2017. *Evolution and the Fall.* Grand Rapids, MI: Eerdmans.

Couenhoven, J. 2005. "St. Augustine's doctrine of original sin." *Augustinian Studies*, 36(2), 359–396. https://doi.org/10.5840/augstudies200536221.

Chesterton, G. K. 1909. *Orthodoxy*. London: John Lane, The Bodley Head.

Crisp, O. D. 2006. "Federalism vs realism: Charles Hodge, Augustus Strong and William Shedd on the imputation of sin." *International Journal of Systematic Theology*, 8(1), 55–71. https://doi.org/10.1111/j.1468-2400.2006.00183.x.

Crisp, O. D. 2015. "On original sin." *International Journal of Systematic Theology,* 17, 252–266. https://doi.org/10.1111/ijst.12107.

Croasmun, M. 2017. *The emergence of sin: The cosmic tyrant in Romans*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190277987.001.0001.

Darwin, C. 1859. *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. London: John Murray. https://doi.org/10.5962/bhl.title.82303.

Darwin, C. 1871. *The descent of man, and selection in relation to sex*. London: John Murray. https://doi.org/10.5962/bhl.title.24784.

Darwin, C. 1872. *The expression of the emotions in man and animals*. London: John Murray. https://doi.org/10.1037/10001-000.

Dean, L. G., Vale, G. L., Lalan, K. N., Flynn, E., & Kendal, R. L. 2014. "Human cumulative culture: A comparative perspective." *Biological Reviews,* 89(2), 284–301. https://doi.org/10.1111/brv.12053.

De Cruz, H. 2022. "A taste for the infinite: What philosophy of biology can tell us about religious belief." *Zygon: Journal of Religion and Science*, 57(1), 161–180. https://doi.org/10.1111/zygo.12765.

De Cruz, H., & De Smedt, J. 2013. "Reformed and evolutionary epistemology and the noetic effects of sin." *International Journal for Philosophy of Religion*, 74(1), 49–66. https://doi.org/10.1007/s11153-012-9368-z.

De Cruz, H., & De Smedt, J. 2021. "Introduction to the symposium on evolution, original sin, and the Fall." *Zygon: Journal of Religion and Science*, 56(2), 447–453. https://doi.org/10.1111/zygo.12697.

De Smedt, J. & De Cruz, H. 2020. *The challenge of evolution to religion.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108685436.

Devine, D. 2021. "Discrete events and hate crimes: The causal role of the Brexit referendum." *Social Science Quarterly*, 102(1), 374–386. https://doi.org/10.1111/ssqu.12896.

Duffy, S. 1988. "Our hearts of darkness. Original sin revisited." *Theological Studies*, 88, 597–622. https://doi.org/10.1177/004056398804900401.

El Mouden, C., André, J. B., Morin, O., & Nettle, D. 2014. "Cultural transmission and the evolution of human behaviour: A general approach based on the Price equation." *Journal of Evolutionary Biology*, 27, 231–241. https://doi.org/10.1111/jeb.12296.

Flanagan, O. 2017. *The geography of morals. Varieties of moral possibility.* New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190212155.001.0001.

Green, A., & Morris, J. 2020. "Living within our limits: A defense of the Fall." *Journal of Analytic Theology*, 8, 371–389. https://doi.org/10.12978/jat.2020-8.061713121418.

Green, J. B. (2017). ""Adam, what have you done?" New Testament voices on the origins of sin." In *Evolution and the Fall*, edited by W. T. Cavanaugh and J. K. A. Smith, 98–116. Grand Rapids, MI: William B. Eerdmans.

Henrich, J. 2018. *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton University Press.

Kind, C.-J., Ebinger-Rist, N., Wolf, S., Beutelspacher, T., & Wehrberger, K. 2014. "The smile of the lion man. Recent Excavations in Stadel Cave (Baden-Württemberg, south-western Germany) and the restoration of the famous Upper Palaeolithic figurine." *Quartär*, 61, 129–145.

Lamoureux, D. O. 2015. "Beyond original sin: Is a theological paradigm shift inevitable?" *Perspectives on Science and Christian Faith*, 67(1), 35–49.

Lewis-Williams, D. 2002. *The mind in the cave: Consciousness and the origins of art.* London: Thames & Hudson.

Madueme, H. 2021. "The theological problem with evolution." *Zygon: Journal of Religion and Science*, 56, 2, 481–499. https://doi.org/10.1111/zygo.12690.

Madueme, H. and Reeves, M. 2014. *Adam, the Fall, and original sin. Theological, biblical, and scientific perspectives*. Grand Rapids, MI: Baker.

Morgan, T. J., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., et al. 2015. "Experimental evidence for the co-evolution of hominin tool-making teaching and language." *Nature Communications*, 6(1), 1–8. https://doi.org/10.1038/ncomms7029.

Munger, K. 2017. "Tweetment effects on the tweeted: Experimentally reducing racist harassment." *Political Behavior*, 39(3), 629–649. https://doi.org/10.1007/s11109-016-9373-5.

Nelson, D. R. 2009. "Schleiermacher and Ritschl on individual and social sin." *Journal for the History of Modern Theology/Zeitschrift für Neuere Theologiegeschichte*, 16(2), 131–154. https://doi.org/10.1515/ZNTH.2009.007.

Norenzayan, A. 2013. *Big gods: How religion transformed cooperation and conflict.* Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400848324.

Okasha, S. 2006. *Evolution and the levels of selection*. Oxford: Clarendon Press. https://doi.org/10.1093/acprof:oso/9780199267972.001.0001.

Olberding, A. 2011. *Moral exemplars in the Analects. The good person is that*. New York: Routledge.

Paluck, E. L., & Chwe, M. S. Y. 2017. "Confronting hate collectively." *PS: Political Science & Politics*, 50(4), 990–992. https://doi.org/10.1017/S1049096517001123.

Paluck, E. L., Shepherd, H., & Aronow, P. M. 2016. "Changing climates of conflict: A social network experiment in 56 schools." *Proceedings of the National Academy of Sciences USA*, 113, 566–571. https://doi.org/10.1073/pnas.1514483113.

Pedersen, D. 2020. *Schleiermacher's theology of sin and nature.* London: Routledge. https://doi.org/10.4324/9780429199059.

Pedersen, D. 2017. *The eternal covenant: Schleiermacher on God and natural science*. New York and Berlin: De Gruyter. https://doi.org/10.1515/9783110542301.

Perry, J., & Leidenhag, J. 2021. "What is science-engaged theology?" *Modern Theology*, 37(2), 245–253. https://doi.org/10.1111/moth.12681.

Purzycki, B. G., Willard, A. K., Klocová, E. K., Apicella, C., Atkinson, Q., et al. 2022. "The moralization bias of gods' minds: A cross-cultural test." *Religion, Brain & Behavior*, 12(1–2), 38–60. https://doi.org/10.1080/2153599X.2021.2006291.

Rauschenbusch, W. 1917. *A theology for the social gospel.* New York: MacMillan.

Richerson, P. J. & Boyd, R. 2005. *Not by genes alone, how culture transformed human evolution*. Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226712130.001.0001.

Roberts, S. O., Gelman, S. A., & Ho, A. K. 2017. "So it is, so it shall be: Group regularities license children's prescriptive judgments." *Cognitive Science*, *41*, 576–600. https://doi.org/10.1111/cogs.12443.

Roberts, S. O., Guo, C., Ho, A. K., & Gelman, S. A. 2018. "Children's descriptive-to-prescriptive tendency replicates (and varies) cross-culturally: Evidence from China." *Journal of Experimental Child Psychology*, 165, 148–160. https://doi.org/10.1016/j.jecp.2017.03.018.

Schleiermacher, F. 2006 [1799]. *On religion. Speeches to its cultured despisers* (R. Crouter, Trans.). Cambridge: Cambridge University Press.

Schleiermacher, F. 2002 [1812–1813]. *Lectures on philosophical ethics*. (R. Louden, Ed., L. A. Huish, Trans.). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511814952.

Schleiermacher, F. 2016 [1830]. *Christian faith* (N. Tice, C. L. Kelsey., E. Lawler, Trans). Louisville: Westminster John Knox Press.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. 2012. "Young children enforce social norms selectively depending on the violator's group affiliation." *Cognition*, 124(3), 325–333. https://doi.org/10.1016/j.cognition.2012.06.004.

Schneider, J. 2012. "The fall of "Augustinian Adam": Original fragility and supralapsarian purpose." *Zygon: Journal of Religion and Science,* 47(4), 949–969. https://doi.org/10.1111/j.1467-9744.2012.01307.x.

Smith, J. K. A. 2017. "What stands on the Fall? A philosophical exploration." In *Evolution and the Fall*, edited by W.T. Cavanaugh and J. K. A Smith, 48–64. Grand Rapids, MI: William B. Eerdmans.

Tennant, F. R. 1902. *The origin and propagation of sin.* Cambridge: Cambridge University.

Tennant, F. R. 1903. *The sources of the doctrines of the Fall and original sin.* Cambridge: Cambridge University.

Tomasello, M. 2016. *A natural history of human morality*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/9780674915855.

Treviranus, G. R. 1802. *Biologie, oder Philosophie der lebenden Natur für Naturforscher und Aerzte*. Göttingen: Johann Friedrich Röwer. https://doi.org/10.5962/bhl.title.10546.

Van den Toren, B. 2016. "Human evolution and a cultural understanding of original sin." *Perspectives on Science and Christian Faith*, *68*(1), 12–22.

Venema, D., & McKnight, S. 2017. *Adam and the genome. Reading scripture after genetic science.* Grand Rapids, MI: Brazos Press.

Vicens, L. C. 2018. "Sin and implicit bias." *Journal of Analytic Theology*, *6*, 100–111. https://doi.org/10.12978/jat.2018-6.110407210818a.

Watts, J., Greenhill, S. J., Atkinson, Q. D., Currie, T. E., Bulbulia, J., & Gray, R. D. 2015. "Broad supernatural punishment but not moralizing high gods precede the evolution of political complexity in Austronesia." *Proceedings of the Royal Society B: Biological Sciences*, 282(1804), 20142556. https://doi.org/10.1098/rspb.2014.2556.

White, C., Marin, M., & Fessler, D. M. T. 2022. "The dead may kill you. Do ancestor spirit beliefs promote cooperation in traditional small-scale societies?" *Journal of Cognition and Culture,* 22(3–4), 294–323. https://doi.org/10.1163/15685373-12340135.

Whitehouse, H., François, P., Savage, P. E., Hoyer, D., Feeney, K. C., et al. 2021. "Big Gods did not drive the rise of big societies throughout world history." Preprint. https://doi.org/10.31219/osf.io/mbnvg

Williams, P. 2001. *Doing without Adam and Eve. Sociobiology and original sin*. Minneapolis: Fortress Press.

Wyman, W. E. 1994. "Rethinking the Christian Doctrine of Sin: Ernst Troeltsch and the German Protestant liberal tradition." *Journal for the History of Modern Theology/Zeitschrift für Neuere Theologiegeschichte*, 1, 226–250. https://doi.org/10.1515/znth.1994.1.2.226.

Youngblood, M. 2020. "Extremist ideology as a complex contagion: the spread of far-right radicalization in the United States between 2005 and 2017." *Humanities and Social Sciences Communications*, *7*(1), 1–10. https://doi.org/10.1057/s41599-020-00546-3.

Published Online First: September 19, 2022

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.67873

Vices, Virtues, and Dispositions

LORENZO AZZANO
University of Santiago de Compostela
lorenzo.az@hotmail.it

ANDREA RAIMONDI
Thapar Institute of Engineering and Technology
& andrea.raimondi@thapar.edu

Abstract: In this paper, we embark on the complicated discussion about the nature of vice in Virtue Ethics through a twofold approach: first, by taking seriously the claim that virtues (and certain flavours of vices) are genuinely dispositional features possessed by agents, and secondly, by employing a pluralistic attitude borrowed from Battaly's pluralism (2008). Through these lenses, we identify three varieties of viciousness: incontinence, indifference, and malevolence. The upshot is that the notion of vice is not as categorically homogeneous as that of virtue: some states of viciousness consist in interference of present virtuous dispositions, or mimicking of absent vicious ones, whereas others can be considered genuine dispositions themselves. Furthermore, this set-up can provide an interesting, albeit highly idealized story as to how, through the interference in one's environment, one gets acquainted with vice in various degrees. Finally, this approach can be illuminating vis-à-vis Virtue Ethics in general; e.g. we can employ it to discuss more productively Johnston's (2003) objection to Hursthouse's (1999) account of moral conduct.

Keywords: Virtue ethics, Vices, Dispositions, Moral progress, Vice pluralism

Introduction

There is much discussion in Virtue Ethics about the nature of virtues; but what are we to think of vices? A quick exploration reveals an even more gerrymandered landscape than in the case of virtues: are vices merely absences of virtues, or do they have a reality in their own right? We embark on this discussion through a twofold approach: first, by taking seriously the claim that virtues (and certain flavours of vices) are genuinely dispositional features possessed by agents, and secondly, by employing a pluralistic attitude (borrowed

by Battaly 2008) in which three varieties of viciousness are individuated: incontinence, indifference, and malevolence. Our approach can illuminate the nature of vices in a number of ways, and perhaps even illuminate some issues on Virtue Ethics in general, more specifically Johnston's (2003) objection to Hursthouse's (1999) account of moral conduct.

Finally, the exploration of this tripartite distinction through dispositional lenses (especially by focusing on the so-called interference scenarios) can provide an interesting, albeit highly idealized story as to how, through the interference of one's environment, one gets acquainted with vice in various degrees. Examples are provided with the relevant literature.

1. Virtue Ethics and dispositions

1.1. Introduction to Virtue Ethics

According to Virtue Ethics moral issues are primarily about someone's character, rather than about someone's actions; about what kind of person to be, rather than what to do (see Anscombe 1958, *inter alia*). A good moral character consists of qualities the exercise of which permits one to pursue and achieve certain ideals deemed good, from which the person and those around would benefit. We call these ideals the Good; a person who pursues them, we call a virtuous person; and the means through which they pursue them, the virtues. Naturally which quality is deemed a virtue vastly depends on what we believe the Good to be. If, for example, we are persuaded of the importance of fairness and equitableness we would claim that justice is Good and, consequently, that it is good, or *virtuous*, to be just. Thus, the quality of being just corresponds to justice as a virtue. Which are the virtues? Besides justice, Plato, in the *Phaedro*, famously mentions temperance, fortitude, and discernment as those virtuous character traits that, thereafter, would have come to be known as the cardinal virtues. But the list is far from exhaustive: we now recognize a broad variety of virtues, such as courage, compassion, perseverance, integrity, generosity, prudence, and many others.

Our choice in the selection of virtues vastly depends on our preferences in Virtue Ethics, and most specifically, the relation between the Good and the virtues. For our purposes, concerning the discussions of both virtues and vices, it may be preferable to adopt a somewhat pluralistic attitude; as noticed by Battaly, different views capture different aspects of what constitutes a virtue, and it would be unproductive to try to find a sole winner, and hence a unique concept of virtue or a unique list of virtues (Battaly 2008).

For instance, drawing from Aristotle's original account, Eudaimonists such as Foot (1978; 2001), Annas (1993), and Kraut (1989), argue that what confers a

quality its status as a virtue is its relation to the Good understood as *eudaimonia*, "flourishing", or "happiness", thus, because each virtue is entirely defined in terms of its contribution to eudaimonia, the only qualities counting as virtues are the ones that latch onto it. On this view, the normativity of virtues depends, perhaps ontologically, on the normativity of the Good. Things are different in Agent-Based Virtue Ethics, where the source of normativity lies in the degree to which the virtue is entrenched in the person, and their motivations for acting virtuously (see Hursthouse 1999; Zagzebski 2004; Adams 2006). On Zagzebski's account "[w]e do not have criteria for goodness in advance of identifying the exemplars of goodness" (2004, 41), hence which of the qualities count as virtuous depends on a sort of primitive agential response, refined over time, to wanting to be like some and not wanting to be like others. The source of normativity, then, is our propensity to take a liking, which is in turn based on exemplary emotions and motives, and not human flourishing. As a consequence, Agent-Based virtue ethicists can welcome more qualities in the domain of the virtues. In a similar fashion, Swanton (2003) believes that we can be liberal about what counts as a virtuous character trait, whose moral value is already pre-theoretically grasped independently of any prior (explicit) account of the Good. Finally, Chappell (2014) and Murdoch (1971) develop a Platonic Virtue Ethics according to which virtues are not defined in terms of the Good: rather, a virtue is whatever quality the exercise of which "pierce[s] the veil of selfish consciousness and join[s] the world as it really is" (Murdoch 1971, 91). Virtues, then, have a negative role in driving our attention away from desires and passions, thus enabling us to see reality for what it is, or as Plato would have said, the Forms around us.

The Eudaimonist, Agent-Based, and Platonist options are of course not the only ones available on the market, but we can already see the different extent of the domain of virtues: the Platonic account is far more liberal than the Eudaimonistic, but, perhaps, more restrictive than the Agent-Based account, in that Platonic virtues are, to a large extent, intellectual in nature.

1.2. Virtues as dispositions

In a space where many flowers have bloomed, there seems to be one common element: virtues are dispositions (see Annas 2005, *inter alia*). As a matter of fact, the identity between virtues and dispositions is a central tenet of virtue ethics; yet this identification is so foundational that its rationale is occasionally poorly explored: why are character traits (thus, both virtues and vices) best characterized as dispositions? Clarifying this issue will be crucial for our purposes in the rest of the paper.

There are several *prima facie* reasons for the identification of virtues/vices with dispositions. In our mind, they do carry some weight, although they are not *per*

se conclusive. Firstly, the idea is that both virtues and dispositions are associated with both triggering and manifestation conditions, specifying what kind of behaviour they are meant to elicit in a given circumstance. These are meant to generate patterns in the behaviour of their bearer, which, albeit not exceptionless, are most often than not detectable. Individuals with such-and-such dispositions tend to behave in a certain way, and the same is true of individuals with such-and-such character traits. That said, the disposition can still exist, and be instantiated, in the absence of such behaviour, viz. when the triggering conditions are not met. This association is meant to showcase an important feature of Virtue Ethics: one-time or occasional behaviour is neither necessary nor sufficient to warrant the ascription of a virtue (or, as we will argue, for the ascription of a vice).

Here is an example which may showcase some of these features: we stumble upon our colleague Luigi speaking kindly of his colleagues. We might thus want to claim that Luigi is a generous person, but perhaps that wouldn't be wise; for generosity is best identified with a disposition that may, or may not, underpin that behaviour. Perhaps Luigi has only so spoken for the first and only time (in that case, we would say that he is in a momentary generous state), or merely to promote his self-interest (in that, we wouldn't describe him as generous at all); hence, albeit character traits are connected with the kind of behaviour they produce, they are not identical to it. In order for us to attribute generosity to Luigi, his generosity needs to be displayed with a certain regularity; but not necessarily: Luigi can be described as truly generous even in circumstances when his generosity does not give rise to the relevant generous actions or thoughts. He might, for instance, be alone in his room reading and thus in a situation where the conditions that are relevant for his generosity are not met. Just like dispositions, character traits need to be triggered to produce the associated action, thought, or emotion –which may or may not happen.

There are other similarities between dispositions and character traits, that might be worthwhile to consider, e.g. the fact that they both come in degrees. Famously things can be more or less disposed to a certain outcome (e.g. Bird 2007, Vetter 2015), which is reflected in everyday ascriptions of dispositions (e.g. when we say that "gold is a highly malleable metal"). Virtue Ethicists' talk of excellence may reveal that something similar occurs for virtues as well: Whereas Virtue Ethicists of the Aristotelian kind often use excellences and virtues (*arête*) interchangeably, many others use excellency to qualify the virtue instead. For example, arguing for a version of Agent-Based Virtue Ethics, Swanton claims that a virtue "is a disposition to respond to, or acknowledge, items within its field or fields in an excellent or good enough way" (2003, 19; see also Crisp 2010). Hence, virtues, like dispositions, can be possessed to a higher or a lesser degree, with excellency being the highest; and each virtue can be exercised more or less. A

highly generous person, for example, might help others more often than not, whether because they do so more frequently under the same triggering circumstances, or because they do so in more triggering circumstances. What is more, talk of gradable dispositions provides some ground for a theory of moral praiseworthiness: an agent comes to have the relevant virtue to a higher degree by confronting morally challenging situations more often than not. To be courageous, for example, calls the agent to manifest its braveness. But, to be sure, it is not easy to be brave when the situation calls. And, as Aristotle argues in Book II of the *Nicomachean Ethics*, this is why we attribute valour to those that, when the time comes, manifest their courage. According to Aristotle, virtues are not natural qualities because, unlike natural qualities, those who bear them are responsible for their virtue, and because of this we deem them praiseworthy. Agent-Based virtue ethicists might also be content with the fact that our primitive, theoretical grasp of the nature of virtues stands on a solid foundation, without requiring any relation to a pre-established Good. In fact, if virtues are nothing but dispositions to think, feel, and act in a certain way, the nature of a virtue is exhausted by the dispositions themselves.

Before we move on to the discussion of vices, we would like to stress out the metaphysical relevance of the identity of dispositions and character traits such as virtues and vices. After all, we did preface this section with the claim that these are *prima facie* reasons for this identification. The reason is very simple: that dispositions have stimuli and manifestations, that they come in degrees, that they are (somehow) associated with possibility, are *per se* quite uncontroversial statements, mostly orthogonal to the issues discussed in the metaphysics of dispositions proper. The same claims could be endorsed, properly reconstructed, in a metaphysical set-up where no genuine disposition is actually present, perhaps as claims about conditionals, or laws of nature, or something of that sort. Dispositional talk might be, so to speak, just a *façon de parler* with no real metaphysical import; thus our claim to have a "dispositional account" of vices would be but a spurious application of the metaphysics of dispositions, one of the many Bird (2016) tried to warn us about in describing the pernicious tendency of the metaphysics of dispositions to overextend itself. In a way, everyone agrees that there are dispositions (and perhaps everyone also agrees that virtues are dispositions, as above). But without a proper metaphysics, claims like this mean very little.

This is of course not the place to develop an extensive background on the metaphysics of dispositions, nor to qualify it as an account of the realist variety about dispositions (for some recent efforts, besides Bird 2016, see Azzano 2019, Vetter 2020, and Tugby 2021). Yet we would like to put forward some crucial assumptions about virtues and vices as dispositions that will make our claims not as toothless as the above criticism suggests: crucially, we think that (at least

some) dispositions are intrinsic features of their bearers that may be possessed independently from the environment they find themselves in, and may continue to be instantiated in massively interfered scenarios, viz. in scenarios in which their manifestation is constantly prevented from occurring. We are skeptical of the claim that all dispositions are intrinsic (see McKitrick 2003), but it is a crucial feature of interference scenarios that the relevant dispositions persist in being instantiated independently from the external circumstances (on the correlation between interferences and intrinsicness, see Choi 2003).

Interference scenarios are commonplace in the literature on dispositions (starting from Johnson 1992, Martin 1994, Bird 1998), and constitute everyday occurrences: fragile objects are often protected by packing material to prevent them from breaking; radioactive material can be exposed to moderating boron rods to prevent it from melting, and so forth (see Fara 2005; Choi 2011; *inter alia*). Some of these interferences are not random occurrences, but rather systematic, in the sense that their occurrence co-varies with a disposition's triggering circumstances with some modal strength. Thus, the claim that there are, or there can be, systematically interfered (intrinsic) dispositions amounts to the claim that these dispositions do not manifest as they should, even in the most paradigmatic circumstances. The way we see it, a glass wrapped in bubble wrap, that wouldn't crack even after the most violent strike, still possesses the potential for breaking that we would normally call fragility. More generally, dispositions are not superficial features to be equated with conditional facts about their bearers, but are genuine features possessed by objects independently from their environment. Sometimes, this environment involves interferences: as such, dispositional ascriptions may be true whereas the correspondent counterfactual conditionals are false, or *vice versa*; we will see many of such cases below. Claims like this will come as no news for the reader well-versed in the relevant literature[1] but no less important: after all, Martin's (1994) interference counterexample to the dispositional analysis of dispositions, based on similar cases, is the *locus classicus* in the literature on the metaphysics of dispositions, and as later recognized in Choi (2003), the claim that dispositions are intrinsic is an important element contributing to the impossibility of factoring dispositional talk into conditional talk (also see Eagle 2009). Details, of course, differ from account to account: but with some approximation when we claim that virtues and vices are dispositions, we mean something theoretically primitive *vis-à-vis* counterfactual talk, or other conceptual resources within the nomic family; in other words, to pursue our project we will need to take talk of dispositions seriously.[2]

[1] The asymmetry between dispositional and conditional talk were perhaps first explicitly discussed in Goodman (1954).

[2] This paper operates at a certain degree of abstraction concerning the nature of virtues *qua* dispositions. There's space for manoeuvre, of course. An interesting suggestion is the following:

2. Vices as dispositions

We now have at least a tentative grasp on what a virtue is, and in what sense it is a disposition. What about *vices*? This is the main topic of the rest of the paper, to be discussed upon the background provided so far.

2.1. Vices in Virtue Ethics

Vices have never taken the centre stage of Virtue Ethics. This is primarily because, as *per* any moral theory, the aim of Virtue Ethics is to illustrate how to be moral. Nevertheless, inevitably, discussions of virtues unavoidably suggest considerations about their mirroring images: vices. The list of vices is perhaps even longer than that of virtues. The most famous are perhaps the deadly sins: sloth, pride, envy, covetousness, lust, gluttony, and wrath (see Bloomfield 1952; Taylor 2006). But someone might be vicious by being lazy, inconsiderate, intolerant, selfish, arrogant, hypocritical, ungrateful, and so on.

Curiously, and most interestingly for our current purposes, there is perhaps even more disagreement about the nature of vices than there is about the nature of virtues. Traditionally, e.g. according to the Aristotle of the Book II of the *Nicomachean Ethics*, vices occur when the capacity to come to correct judgments about practical matters is compromised, and virtues fail to hit the mean. Similarly, Plato in Book IV of the *Republic* claims that a vice consists in having too much or too little of something which, in his harmonic form, would be consistent with virtue. As Hurka stresses (2001, 104), for Aristotle and Plato there are no *pure* vices; vices are rather excesses and defects associated with virtues; they are, more precisely, contraries of virtues. What this picture suggests, is that while virtues may very well be dispositions, vices need not be. Similarly, patristic scholars, in according to Christian metaphysics, denied the reality of evil; e.g., Lombard, following Augustine, claims that because evil does not exist in its own right, vices are absences of virtues, rather than realities in themselves (II.35.4.1). Thus, to explain why someone has acted viciously by reacting with too much fear we only need to appeal to his practical reason. It would be improper to drag the virtuous disposition in the explanation: it would be unlovely to claim that vicious behaviour is the manifestation of virtuous dispositions, for virtuous dispositions

---

some friends of dispositions believe that directionality, viz. pointing to a manifestation, is a crucial feature of genuinely dispositional properties (e.g. Armstrong 1997, Molnar 2003, *inter alia*). Despite disagreements about the nature of this directionality, it is worth pointing out that Aristotelian virtues enjoy a similar teleological flavour: for Aristotle, eudaimonia is the highest good for humans, and the virtues are "the state that makes a human being good and makes him perform his function well" (1998: 42-1106a); virtues can achieve this result because they are directed toward eudaemonia; they are virtues *for* the Good. We leave this topic for another time.

have allegedly only virtuous behaviour as their proper manifestation. If this were not the case, they would not be reliable means to achieve the good, or would not constitute the Good. Rather, what we may say is that the virtuous dispositions "misfired": the disposition was present, it was exercised, but the relative manifestation did not occur. In Book VII of the *Nicomachean Ethics* Aristotle informs us that vices so understood profile the morally weak person, a person that has knowledge of what they ought to do but, because they are overpowered in their practical deliberation, fail to attain the good (1146b32). This vicious person, often referred to by Aristotle as the *incontinent* (*akratic*) has, we might say, their virtuous disposition interfered by other character traits or features of the environment.

Let us call a person whose virtuous dispositions are systematically interfered, the Incontinent: the Incontinent can act viciously, on occasion, but not necessarily so: they act viciously when their virtuous disposition is interfered. As already hinted before, interfering scenarios are ubiquitous in the literature of dispositions. In the case of the Incontinent, the idea is that the person possesses other character traits (perhaps other dispositions as well), the presence of which prevents the virtuous dispositions from manifesting correctly. According to Aquinas too, interference with the right order of reason and passion is the mark of vices (Bejczy 2020, 2008–2009). In a sense, dispositions are still part of the explanation of vicious actions, but such actions do not, *per se*, stem from a vicious disposition; nor it is the case that the interfered dispositions must be vicious dispositions. The Incontinent merely pursues passions and desires, which are not *per se* vicious. For example, they might fail to drink in moderation (despite knowing how to), prompted by their desire to try several new beers from a new brewery they are just visiting: this desire is in no way vicious, and nonetheless when they do so their actions are blameable (1111a25-b4). Ultimately, the Incontinent is vicious by falling short of their virtues (if vicious at all).

Friends of the Agent-Based approach might complain about this understanding of vices, which makes them derivative from virtues, a mere by-product. One *prima facie* reason to be suspicious is that, if virtues are traits, so must be vices; and if the former are traits that attain the good, the second must presumably be traits that attain the bad. An Agent-Based theorist might accuse the Aristotelian of putting forward a disunified metaphysics. But there is more to the accusation. As Zagzebski (2004) and Adams (2006) claim, a vice understood as acting against the requirement of virtues must be understood as an act grounded in the appropriate vicious motivations. The idea is the following: without looking at the motivation behind the malevolent manifestation we cannot discriminate against someone that merely went wrong in their rational assessment from someone who wanted, and had reasons to, act viciously. This is an old distinction, of course. Aristotle himself recognises this problem: although

the morally weak person may perform a vicious action, it is problematic to deem them vicious: after all they are, in their choice, in conflict with themselves, and may very well show regret for their action; yet there is another sense according to which the vicious person is one which lacks any regret and is at peace with their decision (1151a10; 1150b29-30; see also Battaly 2014, 66). Likewise, mediaeval theologians, most prominently Aquinas, claim that vices are not absences of good qualities, but themselves bad qualities (Bejczy 2020, 2008–2009). Thus enter the Malevolent, different from the Incontinent in the sense that they possess such bad qualities and thus engage in vicious actions, feelings, and thoughts, not by mistake, but through deliberation. They act on choice rather than appetite (*epithumia*): whereas the Incontinent ultimately cares about being good, the Malevolent is bad through and through. To use Hurka's terminology (2001, 93), the Malevolent is a mirror image of the virtuous, a *purely* vicious person, in the sense that their vices (*kakiai*) stem from genuine dispositions.

We can now see how discussion about vices is more complicated than the discussion about virtues: whereas the existence of virtues qua dispositions appear to be somewhat uncontroversial, it is far more complicated to argue for the correspondent claim about vices, with different philosophical sensibilities pulling in different directions. Take for example, gluttony: friends of vice-as-incontinence would notice, and rightly so, that in order to avoid gluttony one must moderate their food consumption, avoiding excesses; moderation in food consumption does not require assuming that there are vice-conducive and non-vice-conducive dishes, so that moderation consists of assuming the latter but not the former. A moderate person is someone who has the proper attitude toward consuming any food: so if a virtuous agent is one who eats conscientiously, the gluttonous is merely one who fails to do so. On the other hand, friends of genuine, positive, vices might point out that there is a matter of rational deliberation that does not merely amount to a lack of moderation. For example, a gluttonous person is not merely one who non-conscientiously eats great quantities of salad, but if they went on their way to seek for great quantities of enticing dishes: in gluttony, quality seems to matter as much as quantity. In a sense, even the gluttonous person has to act conscientiously, in the sense that they have to deliberate over what is most vicious to eat: they have, then, a genuine disposition toward gluttony such that they will opt for more vice-conducive meals.

This is but one example of a vast taxonomy. What is more, the Incontinent and the Malevolent, so understood, hardly exhaust the taxonomy of the vicious. Following a popular distinction, we will also introduce a third variety of vice, which we will call the Indifferent. This heterogeneity may strike the reader as problematic; failures to find an all-encompassing account of vices, running parallel to accounts of virtue introduced before, is something that Annas (1977)

and Hampton (1990) already emphasised, in relation to Aristotle and Virtue Ethics more generally. The way we see it, again to echo Battaly (2008; 2015), is that a pluralistic account of vices is preferable, and that there is no unique and univocal notion of vice that triumphs over the others. There's no point harping on whether the Incontinent or the Malevolent is really vicious, a question which risks turning into a semantic debate very quickly: for even the vices of the Malevolent are just *primus inter pares*. On the contrary, we will show that the tripartite distinction between the Incontinent, the Indifferent, and the Malevolent, yields an interesting account of how we acquire vices *qua* dispositions, in which the competing contenders for the role of "real" vice are just stages of the process of becoming vicious. In other words, these varieties of viciousness may turn out to be a feature, as opposed to a bug, of our account.

2.2.Varieties of viciousness

We know by now that the Incontinent can be deemed vicious but, in a sense, is not completely so: they can be gluttonous or cowardly, but can be so on occasion, while being moderate or brave on others. The Incontinent tends toward the good, and their failure to achieve it might be categorised as a mistake. Vices characterising the Incontinent are of the most familiar kind: they involve a disproportion in thinking, feeling, and acting: lust, anger, drunkenness, pride, and avarice, but also workaholism, chauvinism, and idolatry (see Adams 2006, 38). Such a person is surely vicious, but their motives should be taken into account. Thus, Aristotle, when comparing them with the Malevolent, claims that the Incontinent is like the epileptic, while the properly vicious is like someone with a chronic disease (1151a33). While the first can be cured, the second cannot; they are continuously bad (1150b34). The Malevolent acts viciously with motives, thoughts, and desires that are vicious. Not only that: the Malevolent is meant as a mirror image of the virtuous: as the virtuous desires to act under the guidance of virtue because they are the source of good, and consequently desires the Good not for themselves only but for others too, so does the Malevolent love their own evil and take pleasure in other people's pain. Just like the virtuous, and unlike the Incontinent, the Malevolent is characterised by a lack of internal conflict.[3] Satan is a good example of this pure evil, since he is not only evil, but he is on a quest to propagate evil, from which he takes great pleasure (see Adams 2006, 38–41).

[3] Although this point too is somewhat controversial. In Book IX Aristotle seems to suggest that the Incontinent and the Malevolent are similar to the extent that the latter, just like the former, is hardly at peace with their vices, and rather suffers as a result of their misdeeds. Aristotle states that "the vicious are full of regret" (1166b25) because they cannot escape the memories of their past actions, and that "like the morally weak, are in conflict with themselves" (1166b 7-8).

The ontological difference between the Incontinent and the Malevolent is quite simple: while the Incontinent may very well possess virtuous dispositions, the Malevolent does not, and possess genuinely vicious dispositions. There's a third option, of course: one can lack dispositions either way. Thus enter our third variety of viciousness, what we call the Indifferent: the basic idea is that while, for the Incontinent, there still is a drive for the Good, in the Indifferent such a disposition to do good is removed. Battaly (2014, 73) captures the character of the Indifferent as someone lacking moral motivations; as *per* Williams (1993, 5–13), the Indifferent is someone in the vicinity of the amoralist. There's a difference, however: unlike the amoralist, the Indifferent has knowledge of the Good: if indifference is a vice, it is because the agent under consideration, despite knowing what is good, decides not to act accordingly. Unlike the Incontinent, the Indifferent is not making any rational mistake, nor will they regret their inaction. Finally, unlike the Malevolent, they have no desire to bring about evil.

This concludes our preliminary exploration of these varieties of vice: we will now illustrate how a dispositional understanding can illuminate them.

2.3. A dispositional spectrum of viciousness

Let us begin with the Incontinent. We do not characterise incontinence cases as one-shot cases of deliberating and acting inappropriately: to some extent, the Incontinent is disposed to act against the command of virtues as a result of their passion. Again, contrary to the Malevolent, they do not act on a decision but are frequently a victim of the circumstances.

These circumstances involve certain other character traits of the agent which function as interferences: these character traits, themselves dispositions, can standardly be understood as passions. With this, we do not simply mean to acknowledge the obvious fact that the situation a person finds themselves in is a significant determinant of how they will behave; rather, what we claim is that the presence of these triggering circumstances can be understood as a specific class of virtue-interferences. More specifically, *mimicking* (Johnston 1992; Bird 1998). Mimicking circumstances are such that there exists an interfering factor in virtue of which x exhibits the manifestation of a disposition without possessing it. As Bird nicely puts it:

> For example, consider a robust iron cooking pot that is definitely not fragile. However it is attached to a powerful bomb with a very sensitive detonator. Should the pot be struck or dropped the detonator will cause the bomb to go off and the pot to shatter as a result. So the counterfactual, 'if the pot were struck, it would break' is true but it is not a fragile pot. (2007, 29)

The idea is that we can explain the behavior of the Incontinent without primarily appealing to any of its virtuous dispositions (just like, say, the sturdiness of the bomb-strapped iron pot is not primarily responsible for its destruction). Sins of incontinence are ones in which virtuous dispositions "misfire" only in the roundabout way that vicious dispositions are mimicked, viz. their behaviour is replicated in their absence. We can offer at least an example of this phenomenon, viz. what is sometimes dubbed the "bystander effect" (BE). In its original framework (see Latané & Darley 1968: 1969: 1970), BE refers to the phenomenon in which the presence of bystanders influences an individual's likelihood of helping a person in an emergency situation. There are a number of cases, reported in the news (see Rosenthal 1964; *inter alia*), where violence occurs but none of the bystanders does anything to prevent the incident. Before the introduction of BE the common explanation of these occurrences referred to the indifference or apathy of the agents involved. The introduction of BE disproved this hypothesis (Latané & Darley 1969) emphasising, instead, two facts: firstly, each individual bystander is disposed to help someone in danger, *if taken in isolation*; secondly, (with a certain degree of idealization) the larger the group of bystanders, the less likely one is to properly manifest this disposition. It's easy to see how we can understand BE as a case of mimicking: a bystander is disposed toward helping those in need involved in a dangerous situation; in other terms, they instantiate the virtuous disposition to manifest bravery, yet external circumstances, such as the presence of bystanders, cause to behave as if they possess a contrary vicious disposition, although they don't (the crowd in BE functions as Bird's bomb-plus-detonator).

It is important to notice how our account can only function previa a serious ontology of dispositions, in which dispositional ascriptions cannot be taken as shortcuts for counterfactual conditional or other similar items. Virtuous and vicious dispositions are genuine and intrinsic features that are possessed independently from a context that might turn the correspondent conditional false. This is the case of the Incontinent: although the virtuous disposition to, say, help people in need were to be instantiated, the correspondent conditional "if someone were in danger, I would help them" is false; contrariwise, although no vicious disposition is present, the conditional "if someone were in danger, I would ignore them" is true. Similarly, the presence of the virtuous disposition in a massively interfered scenario is what allows us to find truthmakers for sentences like "the agent could have resisted the temptation to act in such-and-such a manner", the truthmaker being the virtuous disposition itself (see Jacobs 2010 for dispositions as truthmakers for counterfactual conditionals). Things are different for the Malevolent, for which no such truthmaker can be found, thus turning this truth into a falsehood: the Malevolent could not have resisted the temptation and acted virtuously, for they lack the relevant virtues.

Three things are interesting to notice about the Incontinent, so understood. Firstly, as the number of people who are present in an emergency situation increases, the less likely it is that any individual will help someone in need: that means that the mimicking can generate cases where not only does it seem as if the individual has the vicious disposition, but also that this alleged disposition is possessed to a higher or a lesser degree. Secondly, although BE specifically concerns emergency situations, we can see how this case generalises to others where there are contextual interfering factors. We do believe, in fact, that many vices of the Incontinent can be so explained: we can interpret lust, gluttony, drunkenness, smokiness, or workaholism as cases where someone cares disproportionally about sex, drinks, food, tobacco, and productivity either because of peer pressure or because the circumstances call for excessive reaction, like in the example mentioned before where the Incontinent, despite knowing how to drink in moderation, might fail to do so prompted by the abundance of new beers in the new brewery. There are also reasons to think that the present account can be extended to intellectual vices, as per Cassam (2016; 2019). Cassam puts forward a theory of vice *à la* Driver (2001), which he dubs Obstructivism, according to which *x* counts as a vice if *x* is an "intellectual limit", that is, a trait, a thinking style, or an attitude that obstructs the acquisition, retention, and transmission of knowledge. And this seems to be exactly what is going on in the case we just discuss, for the Incontinent fails to retain their virtuous behaviour in the face of events.

Unlike the Incontinent, the Malevolent lacks a virtuous disposition and rather possesses a vicious one. Thus, because the vices of the Malevolent stem from genuine vices *qua* dispositions, the Malevolent is vicious more often than not and, certainly, more often than the Incontinent. It stands to reason that sometimes their behaviour will be different. But some other times, it will align, and on some very special occasion, it will align systematically, due to systematic interferences (in our account mimicking circumstances) that give the Incontinent their name. Yet the underlying ontology of dispositions allows us to keep the two separate. After all, the Malevolent will lack any virtuous disposition to help a stranger, and will rather possess a vicious disposition to, say, passively observe the violence for their pleasure; by way of BE, the Incontinent and the Malevolent will in fact behave in very similar ways. Yet their dispositions, and thus their motives, for doing so are very different. To be sure, one has to appreciate how systematic these mimicking circumstances must be for the similarities to arise between the Incontinent and the Malevolent; consider this other (less dramatic) example: suppose that Luigi has the disposition to drink in moderation; nevertheless, confronted by peer pressure, he behaves as if he would have acted out of vices, in virtue of the mimicking mechanism described above. Now suppose that Luigi *exclusively* goes out drinking when accompanied by his department colleagues;

it now seems that in every instance in which Luigi goes out drinking, he fails to do so in moderation. If such cases are possible, and we think they are, a Malevolent Luigi and an Incontinent Luigi would be virtually indistinguishable. There is, again, a difference in their motives. But we are in no way arguing that the two are identical, just that there are similarities that justify talking of our Incontinent as a stage toward higher vicious stages (more on that later).

The case of the Indifferent is more elusive than the previous two. On some occasions, one can conceive the Indifferent as acting just like the Incontinent and the Malevolent, but for different reasons; e.g. in the bystander example discussed above concerning BE, one can imagine the Indifferent failing to take action not because of BE (like the Incontinent), nor because of malice (like the Malevolent), but out of sheer indifference. As Adams points out, vices of indifference are "one of the deepest mysteries about moral excellence and badness" (2006, 45). This is because indifference captures cases where "a vice, incompatible with good moral character, is constituted by an absence of concern for an important good, rather than by the presence of an opposition or hostility" (2006, 45). We think we can make progress in understanding indifference by looking at it through dispositional lenses. Recall, again, the mimicking situation of the bystander: the Incontinent bystander fails to manifest bravery, yet they are full of regrets about their (in)action. The Indifferent, however, is someone who, facing someone else's suffering, remains unmoved: it is not the case that they do not recognise something as evil, or that they fail to recognize the right thing to do; they are just indifferent to the other's pain, or other items such as false beliefs held by others that might cause them harm. In this guise, Hurka (2001, 94–95) considers callousness an exemplar vice of indifference, together with apathy and shamelessness; Adams (2006, 44–47) adds ruthlessness to the picture. Ontologically, all these cases have a unique make-up: they don't involve the instantiation of a vicious disposition, as per the Malevolent, but neither do they involve the possession of a virtuous disposition: simply put, the Indifferent lacks motivation either way.[4] It is crucial to notice that our notion of the Indifferent is one of a highly idealized agent. A state of indifference is one which otherwise virtuous (or vicious) agents might be lead temporarily or progressively; we can say a little bit more about this by considering again the phenomenon of interferences to dispositions, and more specifically by considering cases of *finking* (Martin 1994, Lewis 1997). To quote Bird (2007, 25):

[4] Our characterization of the Indifferent is a byproduct of our previous characterization of the Malevolent, viz. as one who knows which is the Good, and yet deliberates differently. In a similar fashion, the Indifferent is not ignorant of the Good, but lacks motivation either way.

> [A]n electro-fink is a device that can make an electric wire live or dead. It also detects whether the wire is being touched by a conductor. Let us take 'live' to mean 'disposed to conduct a current when touched by a conductor'. Let the wire be live; it is properly connected to an electric generator. Let the electro-fink operate by making the wire dead (cutting the connection to the generator) whenever it is touched by a conductor. Thus the wire is live. But were the wire to touch a conductor, the electro-fink would cause it to become dead and it would not conduct a current. So something can be live (disposed to conduct a current when touched by a conductor) yet it is false that if it were touched by a conductor it would conduct a current.

Put in another way, the circumstances that act as a stimulus for the disposition are also the condition for its loss. As our guiding example, we take the case of one virtue, charity, and highlight a triggering situation, an encounter with a beggar, that acts as a fink, causing the loss of the disposition.[5]

The phenomenon of beggars is widely researched and discussed in many areas, from sociology (McIntosh & Erskine 2000), to anthropology (Butovskaya, Diakonov, Salter 2002), to, of course, philosophy (Moen 2014; Allais 2015). Although the phenomenon is wide-spreading, some occurrences are poorly understood. For example, although the disposition of high-status individuals to act with charity is almost a cross-cultural universal, people more often than not fail to give to beggars. They remain unmoved by their sorrow, that is, indifferent. By investigating the people's experiences and attitudes in these circumstances, McIntosh and Erskine report that, although people are disposed to donate, consistent views about the begging encounters are in fact rare (2000, 3). Motives behind this attitude are illustrative of the charity disposition being finked, and,

[5] There's of course another reading of this scenario (we would like to thank an anonymous reviewer for stressing this point): the beggar scenario might be framed not as a finking scenario, but one in which the disposition does not fire at all.

We think this is not a very appealing proposition. Suppose that I fail to give money to a beggar because of, say, race- or class-based prejudice. One might decide to incorporate the interfering factor in the stimulus, so that this might not count as an interference scenario. Thus generosity might be characterised (at least partially, since we are probably dealing with a multi-track disposition here), as a *disposition to give money to beggars, unless victims of our race- or class-based prejudice*. But is this what generosity is?

One of the lessons that we can learn from the literature on interference counterexamples to the conditional analysis of dispositions, is that although such scenarios might be dealt with by properly qualifying stimulus and/or manifestation specifications, this runs the risk of compromising the nature of the disposition itself, up to and including complete trivializations (e.g. *ceteris paribus* clauses). This seems to be the fate of generosity, once understood as a massively qualified disposition to give money to beggars. As a matter of fact, this kind of procedure is particularly malignant in the case of virtues *qua* dispositions: if every interference to generosity can be incorporated *qua* stimulus, as an excuse to not exercise it, can one ever fail to be generous?

almost across the board, reflect the troubling nature of the encounter. People more often than not see the triggering circumstances as not genuine (2000, 4). This is echoed by Moen (2014) who argues against giving to beggars from the premise that, compared to the poor, beggars have access to the "high-society" and so a pound in their money is a pound wasted. Their needs are not "genuine". Another consideration focuses on the moral status of the triggering circumstance. For example, some other people see the beggars as "moral strangers" (McIntosh & Erskine 2000, 5), someone whose presence challenges and undermines their set of values and the understanding of their own lives (O'Neill 1999). Something in the vicinity of this is envisaged by Kant who, arguing against giving to beggars, claims that:

> A poor man who begs is constantly depreciating his personhood and abasing himself; he makes his existence dependent on other people, and accustoms others, by the sight of him, to the means whereby we neglect our own worth. (2001, 605)

How is the Indifferent at fault, here? Like the Incontinent, the Indifferent fails to be virtuous, but the reason for the failure is here different from the one of the former. There is no rational failure involved, and the subjects recognise the circumstance where they would donate if they were exposed to them. Nevertheless, these circumstances, which should allegedly act as triggers for the virtuous disposition, are now perceived as vicious in nature, or at least estranged from the good. The same circumstance that triggers a disposition for the good is such that, if the disposition is exercised, it would lead to something considered bad. Because there is no good in possessing a disposition to donate whose manifestation results in a morally bad occurrence, like an electro-fink, the triggering event acts by removing the person's disposition to donate. We are, of course, not claiming that we develop the same indifference in other cases in which the disposition is triggered; only that some of these cases are akin to finks, causing the loss of the disposition so that, if the only circumstances encountered are the beggars one, the person can be deemed to have lost their virtue of charity. According to Adams (2006, 44–45), for instance, the Indifferent is not opposed to the good but, more modestly, ends up disregarding it, in itself or in others. What is captured by this Agent-Based approach is the motivational element behind the occurrence of this vice, the same element that is reported by McIntosh and Erskine (2000).

Taking stock: by being Incontinent, someone is being vicious by finding themselves in a situation where they act *as if* they were vicious. On the other hand, by being Indifferent someone is being vicious by having their virtuous dispositions removed.

With the *incontinent-indifferent-malevolent* triad in place, our attempt to bring ontological unity to the vice house, under the flag of dispositions, can be considered complete. A final but crucial conclusion that we can draw from this observation is that, from our pluralistic standpoint, vice is not a categorically univocal phenomenon. A vice can be the presence of something, or the lack of something else; what unifies our treatment of vice is rather the dispositional treatment of both vices and virtues *qua* dispositions.

This treatment is only possible, incidentally, once a certain degree of realism concerning dispositions is assumed. The Incontinent and the Malevolent may manifest very similar (vicious) behaviors, and may satisfy the same counterfactual conditionals; yet their dispositional ascriptions are underpinned by very different underlying properties: this background metaphysical difference is exactly what allows for a variety of vicious agents to be distinguished.

2.4. Turning to vice

So far, we have studied our three varieties of viciousness in isolation; but what about their relation? There's an argument to be made for the claim that we should not conceive the Incontinent, the Indifferent, and the Malevolent as self-contained unrelated flavours of vice, but rather as progressive stages in one's descent to vice. After all, just like there is a way to become virtuous, there is also a way to become vicious: as such, the aforementioned varieties of viciousness might be understood as progressive stages of vice. There is, of course, in any hypothetical transition from a stage of incontinence to a stage of malevolence (passing, perhaps, through a state of indifference), a psychological aspect to be studied concerning how the relevant dispositions, and the correspondent motivations, are acquired and lost by the agent; the way we see it, the study of this aspect primarily belongs to moral psychology, and, as such, cannot be investigated in this paper. We can, however, gesture towards an account which focuses on the ontological underpinning of this transition.

A crucial thing to keep in mind is that virtues and vices, when both understood genuinely as dispositions, can be acquired and lost: we do have at least an inkling of a theory as to how that happens; according to the Aristotelian tradition, such character traits, just like abilities, can be progressively acquired through exercise, and lost without it. Let us begin with the Incontinent; such an agent may very well be similar to a virtuous person, character-wise: yet environmental pressure might lead them to what we may call, a lack of exercise concerning virtuous dispositions (think about the moderation of our colleague Luigi, who is faced with peer-pressure whenever he goes out for drinks with his other colleagues). Mimicking circumstances such as the ones above constitute an excellent example

of such environmental pressure, but others may exist. The important part is that the Incontinent is therefore faced with a veritable training in vice, and finds themselves more and more acquainted with behaviour that we might call vicious. Eventually, the lack of exercise vis-à-vis their virtuous disposition might lead to losing such character traits: in an highly idealised scenario, this is what marks the transition from a stage of incontinence to one of indifference. At this point, the contrary happens concerning vicious dispositions: constantly interfered behaviour counts as exercise towards their acquisition, which eventually leads the agent to the path of becoming the Malevolent.

Of course this is a highly idealised scenario, not only because it ignores the mechanisms of moral psychology underpinning these shifts in dispositions and motivations, but because, being both virtues and vices dispositions, and thus gradable (as above), such shifts take time, display nuances, and do not need to be fully completed: incontinence, indifference, and malevolence, are not absolute moral states, but limiting cases across a spectrum on which an agent may position itself.

A final but crucial conclusion that we can draw from this observation is that, from our pluralistic standpoint, vice is not a categorically univocal phenomenon. A vice can be the presence of something, or the lack of something else; what unifies our treatment of vice is rather the dispositional treatment of both vices and virtues qua dispositions, and the positioning of the varieties of vice on a progressive spectrum.

3. Vices, virtues, and dispositions

As a final remark, we want to briefly sketch a way we can put these results to work by offering a different angle on Johnson's objection (2003) to Hursthouse's virtue-based theory of moral conduct (1999). Like any other normative theory in ethics, Virtue Ethics needs to offer a guiding principle for good actions. According to Hursthouse, Virtue Ethicists must endorse the following:

> (V): An action A is right for S in circumstances C if and only if a completely virtuous agent would characteristically A in C. (1999, 28)

The idea is that, provided that moral progress is possible, becoming virtuous is a matter of doing what a "completely virtuous agent" would do. This agent is, admittedly, an idealised character that possesses *all* the virtues *completely* (Slote 1995; 2001; *inter alia*), meaning that "is aware of all of the morally relevant features of the circumstances pertaining to the act under evaluation and … (in the circumstances) acts (or refrains from acting) in ways that express the relevant set of virtues." (Timmons 2001, 280). Here's the problem in Johnson's own words:

> Consider this person: he is mendacious, lying even about unimportant things such as the films he has seen or books he has read. [. . .] Suppose now that a friend calls him on the carpet for lying, and as a result he decides that he must change.[. . .] His task will not be simple. Lying has become habitual and has permeated his attitudes so deeply that no "decision to do better" can by itself change him. What sorts of steps might he take? [. . .] there are many things he might do [. . .] he decides to begin writing down lies that he tells, no matter how insignificant, to become more aware of his habits and to keep track of improvements.[. . .] Common sense would regard these kinds of things as what he morally ought to do in circumstances such as these, at least insofar as they will improve his character. Yet all are utterly uncharacteristic of completely virtuous agents. (2003, 817–818)

This is worrisome. If moral progress is possible, then a moral "novice" must be able to perform some actions toward moral progress. The actions that a novice must perform are only those prescribed by the appeal to the virtuous agent. Nevertheless, the virtuous agent won't do any of the actions that the novice would need to do to progress. Hence, moral progress is not possible. We think, however, that our account shows that it is not correct to claim that the virtuous agent won't do the things that the novice would.

What sets the two apart is that the former, but not the latter, is *completely* virtuous, and is aware of all the morally relevant features of the circumstances pertaining to the "right" actions. There are no circumstances where the virtuous agent should "keep track of his improvement" and, hence, no actions that the novice can take as an example to progress morally. So, both have a virtuous disposition, just possessed to a different degree, with different degrees of *morally* relevant circumstances. But, as we have seen in our examples of the Incontinent we can possess a virtuous disposition and still end up acting viciously. So, the possession of the virtuous disposition is not sufficient to exclude, from the range of actions performed by the ideal virtuous agent, those aimed at avoiding vicious-triggering circumstances. Moreover, notice that those are circumstances of which the ideal virtuous agent might not be intrinsically aware of. They are only aware of the *morally* relevant features. Nevertheless, many of the relevant features are clearly non-moral: the number of people in the group, a crucial difference-maker for incontinence, is a non-moral feature; likewise, the features that might switch the moral evaluation in the context of giving to beggars, like the age or the gender of the beggar (Butoskaya, Diakanov, Salter 2002), seem non-moral too. Thus, despite being *completely* virtuous, the ideal agent must perform (at least some) actions that the novice too must perform. The difference might lay in the fact that while, for the latter, these actions result, or tend to result, in moral progress, for the former they constitute a sort of "moral maintenance". Now

Johnson might revise, in the light of our account, the definition of *completely virtuous* as to include the fact that the ideal agent cannot have their dispositions interfered the way we allowed. However, we find this solution rather wanting. Following this revision, the ideal agent is such that it is impossible for them to do other than what their dispositions tend them to do, for they are intrinsically unfinkable and unmaskable. But a virtuous agent is such properly because, while she could do the bad, he chooses to do the good. And this is why she is an example for moral novices. As Johnson himself admits "Certainly any conception of a completely virtuous agent should be a human, rather than an impossibly god-like, ideal" (2003, 812). It seems that, after all, knowing the limits of morals and all the ways to be vicious, can contribute to a better understanding of the virtues, and the demand they put on us.

Conclusions

Our purpose in this paper was to bring some degree of ontological clarity to the divided house of theories about vice. We pursued this clarification in the spirit of Battaly's Pluralism, that is, by maintaining that the notion of vice is not categorically homogeneous: some states of viciousness consist in interferences, and mimicking of actually absent vicious features, whereas others, like virtues, are more ontologically substantial.

More to the point, we argued that some vicious phenomena can be explained via dispositional occurrences, such as dispositional finks and mimics, so that we are not necessarily forced to understand them only under the lens of genuine dispositions' instantiation. Because Virtue Ethics ought to discuss vices as well, we hope to have done a service to Virtue Ethics more in general: in fact, we believe that the two projects are strongly connected, and that a philosophical analysis of vices can contribute to a better understanding of virtues, as it aims to recognise and articulate the limits of our moral conduct.

Acknowledgements

We would like to thank Michele Paolini Paoletti and Maria Silvia Vaccarezza for their work on the *TheoLogica* special issue on Sin and Vice; furthermore we would like to thank two anonymous reviewers for useful comments on various drafts of this paper.

Bibliography

Aquinas T. 1888–1906. *Summa Theologiae, Opera Omnia Iussu Leonis* XIII ed, vols 4–12. Commissio leonina, Rome.

Allais, L. 2015. "What Properly Belongs to Me: Kant on Giving to Beggars." *Journal of Moral Philosophy*, 12(6), 754–771. https://doi.org/10.1163/17455243-4681042.

Annas, J. 1977. "Plato and Aristotle on Friendship and Altruism." *Mind* 86 (344): 532–554. https://doi.org/10.1093/mind/LXXXVI.344.532.

Annas, J. 1993. *The Morality of Happiness*. New York: Oxford University Press.

Annas, J. 2005. "Virtue Ethics." In *Oxford Handbook of Ethical Theory*, edited by D. Copp, 515–36. Oxford: Clarendon Press. https://doi.org/10.1093/0195147790.003.0019.

Anscombe, G. E. M. 1958. « Modern Moral Philosophy." *Philosophy* 33, 124:1–19.

Aristotle. 1998. *The Nicomachean Ethics*. Translated by David Ross. Oxford: Oxford University Press. https://doi.org/10.1017/S0031819100037943.

Azzano, L. 2019. "The Question of Realism for Powers." *Synthese* 196 (1):329–354. https://doi.org/10.1007/s11229-017-1478-9.

Battaly, H. 2008. "Metaethics Meets Virtue Epistemology: Salvaging Disagreement About the Epistemically Thick." *Philosophical Papers* 37, no. 3:435–54. https://doi.org/10.1080/05568640809485230.

Battaly, H. 2014. "Varieties of Epistemic Vice." In *The Ethics of Belief*, edited by In J. Matheson & R. Vitz, 51–76. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199686520.003.0004.

Battaly, H. 2015. "Epistemic Virtue and Vice. Reliabilism, Responsibilism, and Personalism." *Moral and Intellectual Virtues in Western and Chinese Philosophy*, edited by Mi, C., Slote, M. & Sosa, E., 109–130. London: Routledge. https://doi.org/10.4324/9781315683935-15.

Bejczy, I. 2020. "Virtue and Vices." In *Encyclopedia of Medieval Philosophy*. 2nd ed, edited by Lagerlund, H. Springer. https://doi.org/10.1007/978-94-024-1665-7_515.

Bird, A. 1998. "Dispositions and Antidotes." *The Philosophical Quarterly*, 48: 227–234. https://doi.org/10.1111/1467-9213.00098.

Bird, A. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford University Press.

Bird, A. 2016. "Overpowering: How the Powers Ontology Has Overreached Itself." *Mind* 125 (498): 341–383. https://doi.org/10.1093/mind/fzv207.

Bloomfield, M. W. 1952. *The Seven Deadly Sins: An Introduction to the History of a Religious Concept*. East Lansing: State College Press.

Butoskaya, M., Diakanov, I., & Salter, F. 2002. "Effects of Sex, Status and Ethnic Similarity on Willingness to Donate to Beggars: a questionnaire study." *Anthropologie* (1962-), 40(2), 119–128.

Cassam, Q. 2016. "Vice Epistemology." *The Monist*, 99, 159–180. https://doi.org/10.1093/mind/fzv207.

Cassam, Q. 2019. *Vices of The Mind. From the Intellectual to the Political*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198826903.001.0001.

Chappell, T. 2014. *Knowing What to Do*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199684854.001.0001.

Choi, S. 2003. "Improving Bird's antidotes." *Australasian Journal of Philosophy* 81 (4): 573 – 580. https://doi.org/10.1080/713659764.

Choi, S. 2011. "What Is a Dispositional Masker?" *Mind*, 120: 1159–1171. https://doi.org/10.1093/mind/fzr083.

Crisp, R. 2010. "Virtue Ethics and Virtue Epistemology." *Metaphilosophy* 41 (1–2): 22–40. https://doi.org/10.1111/j.1467-9973.2009.01621.x.

Driver, J. 2001. *Uneasy Virtue*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511498770.

Eagle, A. 2009. "Causal Structuralism, Dispositional Actualism, and Counterfactual Conditionals." In *Dispositions and Causes*, edited by In Handfield, T., 65–99. Oxford: Oxford University Press.

Fara, M. 2005. "Dispositions and Habituals." *Noûs*, 39: 43–82. https://doi.org/10.1111/j.0029-4624.2005.00493.x.

Foot, P. 1978. *Virtues and Vices*. Oxford: Blackwell.

Foot, P. 2001. *Natural Goodness*. Oxford, Clarendon Press. https://doi.org/10.1093/0198235089.001.0001.

Goodman, N. 1954. *Fact, Fiction & Forecast*. Harvard University Press.

Greco, J. 2010. "A Virtue Epistemology." *International Philosophical Quarterly* 50 (3):399–401. https://doi.org/10.5840/ipq201050358.

Hampton, J. 1989. "The Nature of Immorality." *Social Philosophy and Policy* 7 (1): 22. https://doi.org/10.1017/S026505250000100X.

Hawthorne, J. & Cortens, A. 1995. "Towards Ontological Nihilism." *Philosophical Studies* 79(2): 143–165. https://doi.org/10.1007/BF00989707.

Holroyd, J. & Kelly, D. 2016. "Implicit Bias, Character, and Control." In *From Personality to Virtue*, edited by A. Masala & J. Webber, 106–133. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198746812.003.0006.

Hurka, T. 2001. *Virtue, Vice, and Value*. Oxford: Oxford University Press. https://doi.org/10.1093/0195137167.001.0001.

Hursthouse, R. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.

Jacobs, J. D. 2010. "A Powers Theory of Modality: or, How I Learned to Stop Worrying and Reject Possible Worlds." *Philosophical Studies* 151 (2): 227–248. https://doi.org/10.1007/s11098-009-9427-1.

Johnson, R. N. 2003. "Virtue and Right." *Ethics* 113: 810–34. https://doi.org/10.1086/373952.

Johnston, M. 1992. "How to Speak of the Colors." *Philosophical Studies*, 68: 221–263. https://doi.org/10.1007/BF00694847.

Kant, I. 2001. *Lectures on Ethics*, in Schneewind, J. B., Heath, P. (eds). *The Cambridge Edition of the Works of Immanuel Kant*. CUP.

Kraut, R. 1989. *Aristotle on the Human Good*. Princeton: Princeton University Press. https://doi.org/10.1515/9780691225128.

Latané, B., & Darley, J. M. 1968. Group Inhibition of Bystander Intervention in Emergencies. *Journal of Personality and Social Psychology*, 10(3), 215–221. https://doi.org/10.1037/h0026570.

Latané, B., & Darley, J. M. 1969. "Bystander "Apathy."" *American Scientist*, 57, 244–268.

Latané, B., & Darley, J. M. 1970. *The Unresponsive Bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

Lewis, D. K. 1997. "Finkish Dispositions." *Philosophical Quarterly*, 47, 143–58. https://doi.org/10.1111/1467-9213.00052.

Lombard P. 1971–1981. *Sententiae in IV Libris Distinctae*, 2 vols, 3rd edn, ed. Collegium S. Bonaventurae. Collegium S. Bonaventurae, Grottaferrata.

Martin, C. B. 1994. "Dispositions and Conditionals." *Philosophical Quarterly*, 44, 1–8. https://doi.org/10.2307/2220143.

McIntosh I, Erskine A. 2000. "Money for Nothing?: Understanding Giving to Beggars." *Sociological Research Online* ; 5(1):107–115. https://doi.org/10.5153/sro.449.

McKitrick, J. 2003. "A case for extrinsic dispositions." *Australasian Journal of Philosophy* 81 (2):155–174. https://doi.org/10.1080/713659629.

Moen, O. M. 2014. "Should We Give Money to Beggars?" *Think*, 13(37), 73–76. https://doi.org/10.1017/S1477175613000481.

Murdoch, I. 1971. *The Sovereignty of Good*, London: Routledge.

O'Neill, J. 1999. "What Gives (with Derrida)?" *European Journal of Social Theory* 2. (2): 131–145. https://doi.org/10.1177/13684319922224374.

Rosenthal, A. M. 1964. *Thirty-Eight Witnesses: The Kitty Genovese case*. Berkeley, CA: University of California Press.

Slote, M. 1995. "Agent-Based Virtue Ethics." *Midwest Studies in Philosophy* 20 (1): 83–101. https://doi.org/10.1111/j.1475-4975.1995.tb00306.x.

Slote, M. 2001. *Morals From Motives*. Oxford University Press. https://doi.org/10.1093/0195138376.001.0001.

Sosa, E. 1991. *Knowledge in Perspective*. New York: Cambridge University Press.

Strawson, P. F. 1954. "Particular and General." *Proceedings of the Aristotelian Society* 54: 233–260. https://doi.org/10.1093/aristotelian/54.1.233.

Swanton, C. 2003. *Virtue Ethics: A Pluralistic View*. Oxford: Oxford University Press. https://doi.org/10.1093/0199253889.001.0001.

Taylor, G. 2006. *Deadly Vices*. Oxford University Press. https://doi.org/10.1093/0198235801.001.0001.

Tugby, M. 2021. "Grounding Theories of Powers." *Synthese* 198 (12): 11187–11216. https://doi.org/10.1007/s11229-020-02781-2.

Timmons, M. 2001. *Moral Theory: An Introduction*. Rowman & Littlefield Publishers.

Vetter, B. 2015. *Potentiality: From Dispositions to Modality*. Oxford, England and New York, NY, USA: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198714316.001.0001.

Vetter, B. 2020. "Explanatory Dispositionalism: What Anti-Humeans Should Say." *Synthese* 199 (1–2): 2051–2075. https://doi.org/10.1007/s11229-020-02872-0.

Williams, B. 1993. *Morality: An Introduction to Ethics*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781107325869.

Zagzebski, L. 2004. *Divine Motivation Theory*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511606823.

Published Online First: April 03, 2023

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.67763

Presentism, Timelessness, and Evil[1]

BEN PAGE
Pembroke College, University of Oxford & Eton College
benpage@me.com

Abstract: There is an objection to divine timelessness which claims that timelessness shouldn't be adopted since on this view evil is never "destroyed," "vanquished," "eradicated" or defeated. By contrast, some divine temporalists think that presentism is the key that allows evil to be destroyed/vanquished/eradicated/defeated. However, since presentism is often considered to be inconsistent with timelessness, it is thought that the presentist solution is not available for defenders of timelessness. In this paper I first show how divine timelessness is consistent with a presentist view of time and then how defenders of Presentist-Timelessness can adopt the presentist solution to the removal of evil. After this, I conclude the paper by showing that it's far from clear that the presentist solution is successful and that unless one weakens what is meant by the destruction/vanquishing/eradication/defeat of evil, one can only make the presentist solution work by adopting a number of additional assumptions that many will find unattractive.

Keywords: God and Time, Evil, Timelessness, Divine temporality, Presentism

Some object to divine timelessness by claiming that this view implies that evil is never "destroyed" (Ward 2001, 162), "vanquished" (Craig 2001a, 66; 2001b, 214; 2008, 609–610; Copan & Craig 2004, 162, n. 29), "eradicated" (Peckham 2021, 108) or defeated (Mullins 2014, 127–132; 2021a, 107) and as such divine timelessness should be rejected.[2] Exactly what "destroying," "vanquishing," "eradicating," and defeating amounts to is sadly never much discussed by advocates of the argument, but at the very least it seems they all involve the requirement that evil no longer exists.[3] Since the concept of "defeat" is familiar within the literature on

[1] I wish to thank Brian Leftow, Sam Lebens, Tim Pawl, Alex Pruss and David Worsley for discussion of this paper, since it has been much improved because of this. The opinions expressed here are those of the author and do not necessarily reflect the views of Eton College.

[2] I take this argument to be against divine timelessness, and so not one that requires other attributes like simplicity.

[3] Mullins comes closest to providing a discussion of his terminology, namely defeat. However, whilst he does cite Adams (Mullins 2014, 127, n.37; 2021a, 100), the notion itself gets relatively

evil (e.g. Adams 1999; 2006; Stump 2010, chap. 13), it is the one I shall employ here, where the definition we shall run with requires that whatever else defeat amounts to, it is necessary that evil must no longer exist.[4] The argument can then be put as follows:

1. If God is timeless then evil is never defeated.
2. Christian teaching has it that evil is defeated.[5]
3. Therefore, God cannot be timeless.

The concern of this paper is to begin to assess this argument's potency against divine timelessness.[6] In order to do this I will first show how defenders of the argument go about supporting (1); namely by arguing that presentism makes it the case that there is no evil in the new creation and that divine timelessness is incompatible with presentism. I will then challenge both of these points. Firstly, by showing that divine timelessness is compatible with presentism, such that defenders of timelessness can adopt the presentist solution. And then by raising some concerns about the presentist solution, suggesting that it may not achieve the result advocates of the argument think it does.

Presentism to the rescue!

Before suggesting how a defender of divine timelessness might respond to the argument, it's worth noting how advocates of the argument think they can affirm the defeat of evil, such that evil no longer exists. The key is the adoption of a

---

little attention, with the reader being told that "God must make it the case that their suffering comes to an end . . . God must take them out of the environment where they are vulnerable to horrors . . . God must heal them." (2021a, 107)

[4] It's not clear that Adams, who is the most well-known defender of "defeating" evil, would agree with such a requirement, since her definition doesn't seem to include it: "The evil *e* can be defeated if it can be included in some good-enough whole to which it bears a relation of organic (rather than merely additive) unity; *e* is defeated within the context of the individual's life if the individual's life is a good whole to which *e* bears the relevant organic unity" (1999, 28). Stump, who also focuses on defeat, does not include it either: "To say that her suffering was defeated for her is to say that there was a benefit from her suffering, that that benefit came primarily to her, that it would not have come without her suffering, and that it significantly outweighed her suffering" (2022, 5). Nevertheless, since the above argument requires the non-existence of evil I shall assume it here, but it should be clear given what has been said that many may find fault with the argument from the get-go and therefore wish to use another notion other than 'defeat' to run it, such as Hollingsworth's 'Elimination of Evil (EOE)" (forthcoming, 1).

[5] I focus on Christian teaching here, since this is what the authors of the articles I'm responding to focus upon. Nonetheless, a claim like this is not one only Christians may accept, for instance see Lebens and Goldschmidt (2017), who argue for something similar from a Jewish perspective.

[6] I say begin, since there is more to be said regarding eternalist views of timelessness, as well as the justification for (2). These are topics I aim to address elsewhere (Page, *manuscript a*).

particular metaphysical view concerning the nature of time, namely presentism. For although presentism has proved difficult to define,[7] it is often taken to imply that only present entities exist.[8] Given this, only the present temporal instant exists, and therefore no past or future temporal instants exist.[9] As such, in the present instant the evil acts that I performed in the past do not exist, and neither do any future acts of evil that I will perform. Rather only those acts of evil that are being done in the present instant exist. This is taken to help explain the defeat of evil, since according to the Christian tradition believers in the new creation will be perfected such that they will no longer do evil, and if all that exists in the new creation is the present instant of perfected non-evil doing believers, then evil will not exist either.[10] The result of this, so say these theorists, is that the necessary condition for defeat can be met, evil exists no more, and therefore those who want to follow Christian teaching should be presentists.

Timelessness, what's the problem?

From this alone it may not be clear what the problem for defenders of divine timelessness is, since so far I've had no need to mention God's relationship to time. However, the advocates of the argument claim that timelessness is incompatible with presentism, and therefore the solution offered above is unavailable for defenders of timelessness. Instead, it is supposed that the defender of divine timelessness must adopt eternalism, which then allows advocates of the argument to claim that eternalism is incompatible with their notion of defeat, since on eternalism all instants of time exist, and so even in the new creation the temporal instants containing my earlier acts of evil exist. Given these two problems, it is suggested that we should be divine temporalists, since it is claimed that only divine temporality is compatible with presentism, and only presentism allows for the defeat of evil.[11]

---

[7] For some, such as Tallant & Ingram (2021), the situation is much worse than this as they argue that there is no distinctive core amongst views which are called presentist.

[8] This has lead some to define presentism to be a thesis about *what exists*, with Crisp defining presentism as, "for every *x*, if *x* existed, exists, or will exist, then *x* is a present thing" (2004a, 18; Sider 2001, 11–17; Rea 2003, 246; Bourne 2006, 13; Markosian 2004, 47). However other presentists think that presentism should be defined as a thesis concerning *what it is to exist* (e.g. Merricks 2007, 124–135; Tallant 2014).

[9] I speak of instants as a *façon de parler*, and so one should take it to be neutral concerning whether time is discrete or continuous, substantivalist or relationalist.

[10] One can substitute "new creation" for "heaven," if they wish, but I will speak of new creation throughout.

[11] This whole concern is evident in the writings of Mullins (2014, 123ff), Peckham (2021, 99), and Craig (2008, 609–610, 600). Whilst in other places Craig and Mullins don't make the incompatibility claim regarding presentism and timelessness explicit, they think the argument we are addressing is problematic given eternalism (Craig 2001a, 66; 2001b, 214; Copan & Craig

There is much to argue with here, such as whether an eternalist really cannot make sense of the defeat of evil, but challenging this will have to wait for another time. Rather, what I will argue in the next few sections is that timelessness is compatible with presentism, and therefore advocates of timelessness can adopt the presentist solution.

Presentism and timelessness

Mullins writes, in "current debates over God's relationship to time, it is widely agreed that God cannot be timeless if presentism is true, but that God can be timeless if four-dimensional eternalism is true" (2014, 123).[12] Mullins might be right about the wide agreement, for it is at least clear that many who affirm divine timelessness are eternalists (Rogers 2000, chap. 5; Helm 2010; Mawson 2008), and it is certainly easier to see how timelessness and eternalism fit together. Nevertheless, I think it's far from clear that divine timelessness is incompatible with presentism. Perhaps that puts me in the present minority, but I am by no means alone in thinking this (e.g. Leftow 1991; 2018; McCann 2012, chap. 3).[13] Here, I can only begin to explain why I take the two to be compatible, for it is beyond the scope of this paper to show how a timeless God can cause things in a temporal world, how concerns about omniscience can be overcome, etc., and I do so by providing what I take to be a coherent model for understanding how a timeless God can exist alongside a presentist world.[14]

2004, 162, n.29; Mullins 2021a), and contend elsewhere that presentism is incompatible with timelessness (Craig, 2001c, 139, 282; Mullins 2016). Ward (2001, 162) may present this concern as well, but Ward's words leave it very unclear as to whether he thinks divine timelessness requires an eternalist view of time. Hollingsworth (2023) also argues that presentism is required for the elimination of evil (2023, 2), or perhaps merely that presentism is the best out of some potential options (2023, 13, 16), but he doesn't explicitly say that this requires divine temporality.

[12] Divine temporalists, such as Mullins (2016, 30, chaps. 4–5) Craig (2001c, 139, 232), Hasker (2011, 14–15), Peckham (2021, 99–110), have provided arguments for this claim. I cannot address these here, so all I can say is that I'm not persuaded by them and have responded to one of them elsewhere (2023). Additionally, even some eternalists who take God to be timeless, like Rogers, also think timelessness and presentism are incompatible (2000, 59). Rogers's main reason for this seems to be that she thinks presentism will just imply that God is timeless, as will become evident in the following paragraph, and her other arguments for thinking the two incompatible have been responded to by Leftow (2009).

[13] In so far as other defenders of divine timelessness want to allow for tensed facts (Stump and Kretzmann 1981, 455–458; Wierenga 1989, 175–190), and an absolute temporal now (Stump 2018, 119), this may give us reason to think that they too think presentism is compatible with timelessness, although it would be too quick to state this decisively, as they may only mean to support something like a growing block or moving-spotlight theory instead. It's also clear that many historical classical theists thought that presentism and divine timelessness were compatible (Mullins 2016, 74–76).

[14] For more on providing models in philosophy of religion see my (*forthcoming*).

However, let me first address an initial concern one might have, namely that this prospect is doomed from the outset, since it might be thought that presentism implies that *everything* that exists must exist in the present, and since the present moment is temporal, then God too must be temporal. Rogers makes such a claim when she writes that on presentism "all that exists is the present moment. That means that God exists only in the present moment, since that is all there is" (2011, 11). And Rhoda says the same when he writes, "if presentism is true and God exists, then like everything else God exists now, in the present. The theistic presentist is thus committed to a temporal concept of God" (2009, 53). As we shall see momentarily, I think a presentist can think that everything that exists must exist in *a* present without thinking that everything is temporal, but it is also worth pointing out that many presentists have restricted views of presentism ~~so~~ to allow for the existence of non-temporal entities (Crisp 2004b, 46; Craig 2000, 227; Merricks 2007, 120, n. 1; Leftow 2018, 175).[15] For instance, Bourne writes, "presentism is a theory about what actually exists in time; it says nothing about the existence of anything else. Presentism, like any other theory of time, can have more in its ontology than just objects located in the present" (2006, 79–80).[16] Similarly, Lewis, an ardent four-dimensionalist, thought that presentism was a claim that allowed for non-temporal entities (2004, 4), for surely a presentist can be a Platonist.[17] As such, it seems that there is a very respectable view of presentism that allows for the existence of entities which do not exist in the present.[18]

However, I don't think we have to rely upon such views and can instead uphold that everything which exists must exist in a present, and still think a timeless God is possible.[19] To see this, we will model the situation on a possible

[15] Yet not all presentists allow such restrictions, such as Ingram (2019, chap. 1).

[16] Although not a presentist, Cameron makes the same point, writing, "As I see it, it is compatible with presentism that there exist some things that are not present entities, because they simply do not exist in time at all—such as numbers, or (on some views) God." (2015, 7)

[17] One might worry that the definition Crisp gives of presentism in footnote 8 does not allow for such abstracta, but he is willing to adjust the definition to overcome this concern (2004a, 18; 2004b, 43–46).

[18] Leftow suggests this restricted form of presentism was in fact the dominant view until very recently, with universal presentism being largely a late twentieth-century phenomenon. (2018, 175)

[19] Much of this section and the "not too fast' section is indebted to Leftow (2018), which builds upon some elements found throughout his earlier work (1991). Here I take the thrust of the view and elaborate on certain elements, as well as try to make it more perspicuous, since many often claim to fail to understand what Leftow means when it comes to God's relationship with time (e.g. Rogers, 2009, 321; Mullins, 2016, 153, n.93).

world that has two island universes within it.[20] In virtue of these universes being islands, they have two temporally unconnected timelines, since there are no temporal relations between each timeline, even though both timelines are contained within one possible world. Given this, we can say that the islands are extrinsically timeless, since they have no temporal relations to anything external to them. Suppose further that both universes are presentist in structure, such that in each universe all that exists will exist in *its* present. It's important to notice here that in virtue of these being island universes, each island will have a unique present, as if the two islands had presents which were temporally related then these wouldn't be island universes. As such in our first island universe, U1, t_1 is present, whereas in our second island universe, U2, t^*_4 is present. We thus have one possible world with two presents.

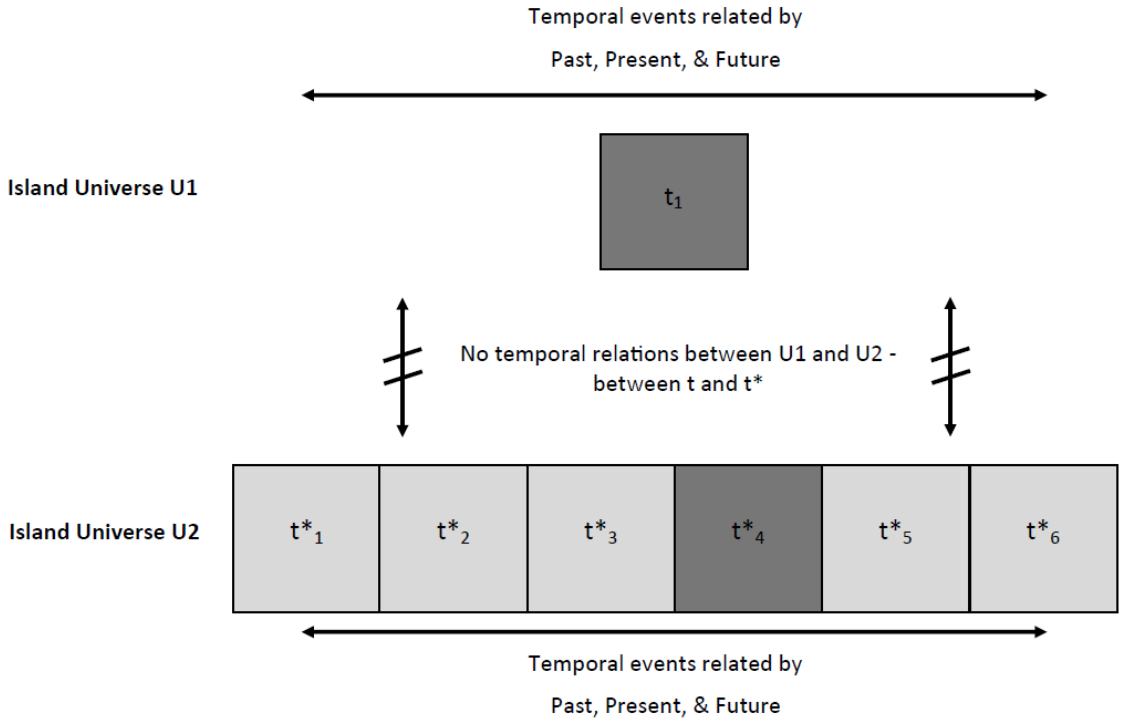
We can also have it that one of these islands is intrinsically timeless. In order to claim this first note that an instant is timeless if it is not possibly succeeded by other instants and/or does not possibly flow, whilst an instant is temporal if it is possibly succeeded by other instants and/or possibly flows.[21] With this background we can claim that Island U1 has the interesting characteristic of being necessarily one instant *long*,[22] and therefore there is necessarily no succession or flow in U1, and the instant t_1 never begins or ceases.[23] We can think of U1 as being what Latham and Miller call a one-instant or stopped presentist world (2020, 145; Tallant 2008, 118), which, as they note, are typically appealed to as timeless worlds. However, U2 is more than one instant long, and has a present which changes, such that there is succession and flow, and therefore we can say that U2 is a temporal universe. All of this we can put in diagrammatic form, such that our possible world with its two island universes looks as follows:

[20] Lewis (1986, 71–72) is famous for thinking this is an impossibility, but see Bricker (2020, chaps. 4 & 6), Baron and Tallant (2016), and Skow (2022, 282) for arguments affirming the possibility of island universes.

[21] This possibility claim allows the first instant of time to be temporal even though it has not been succeeded by other instants and even if the first instant is never succeeded but possibly could have been.

[22] One might worry that talk of an "instant" is misplaced for something timeless. Leftow (2002) provides some responses to these concerns, but if this "instant" talk is off putting, one could opt for Baron and Miller's suggestion and talk of this island universe having no temporal dimension (Baron & Miller 2014, 2872), and thereby remove the need to speak of an instant. However, it's not clear whether one will be able to speak of an eternal *present* in this case, and so also the claim that everything that exists exists in a present.

[23] Leftow (2002, 28) provides a reason why we shouldn't say that t_1 begins or ceases given the temporal structure of U1.



Here the dark highlighted box represents the present of each timeline, and the lightly highlighted boxes in U2 represent that this timeline is more than one instant long, even though only t^*_4 exists in U2, since none of the other instants are present in U2.

Thus, I take it that we have before us an island universe which is both extrinsically and intrinsically timeless, in U1, and another island universe which is extrinsically timeless, in virtue of not being related to U1, but intrinsically temporal, in U2.[24] On this picture all that exists exists in *a* present, and it allows us to say, along with traditional defenders of timelessness, that there is an eternal present.[25] In terms of our model of a timeless God and creation, we can say that t_1 in U1 is God's non-changing eternal present, whilst U2 is God's created presentist world.[26] God undergoes no succession, nor has any temporal relations

[24] U2 *could* be extrinsically related to other things, if other such things existed, but for our purposes we shall take it that only U1 exists outside of U2.

[25] For instance see: Augustine, *Confessions* XI, xiii (16), (1998) 230; Boethius, *Consolation of Philosophy* V.6 (2001), 146; Anselm, *De Concordia* 1.5 (2007), 367; Aquinas, *De Potentia Dei*, q.1, a.5, ad.2. Leftow (2002) also gives a nice discussion as to how God could still be thought of as timeless even though He has a distinct eternal present.

[26] One might worry that island universes don't allow for causal relations and so this isn't a very good model since surely God causes the universe. Yet this isn't something one need accept, and as Bricker (2020, 110) notes, there are weaker notions of island universes that allow causation to be possible between them even though temporal relations are not. Alternatively, one could fall back on the claim that it is a feature of models that they do not capture *everything* about reality, and that this might be the case here.

with creation, and neither begins nor ceases to exist, since as I've already said, t_1 in U1 doesn't begin or cease, and so exists permanently.

All of what has been described so far strikes me as possible, and yet if it is, then I claim that we have a good model for thinking about how a timeless God can exist alongside a presentist world.[27] If, however, one is less convinced, then I am willing to fall back on a form of presentism which allows for the existence of timeless entities, and then claim that my talk of island universes helps us model how those things which exist outside time relate to a presentist world. From now on, I'll refer to this model as Presentist-Timelessness, and whilst there is more to say about how the model explains a timeless God's relationship to a presentist world, what I've said so far I take to be sufficient by showing how in principle the two are not incompatible.[28]

Presentist-Timelessness and Evil

The result of all this is that advocates of the argument we are assessing are incorrect to think that eternalism is the only game in town for someone who thinks that God is timeless. One can affirm divine timelessness and presentism. The question, therefore, is whether the presentist solution to defeating evil is available to someone who adopts Presentist-Timelessness about the Divine. On the face of it it seems so, for as will be remembered, presentism was thought to be required for the defeat of evil since in the future when the new creation exists no evils of the past will exist, for given presentism all that will exist is the present instant of the perfect new creation. Yet on the Presentist-Timelessness model the same will hold. For consider U2 and the evil performed at t^*_1. This evil no longer exists in U2 since only t^*_4 is present in U2, and as such only t^*_4 exists in U2. Thus at $t^*_{\text{new-creation}}$ none of the prior evil performed in U2 will exist either, for all that will exist in U2 is $t^*_{\text{new-creation}}$, in which there is no evil. Presentist-Timelessness can solve the problem in the same way the divine temporalist can.

Not so Fast

However, things are a little more complicated on Presentist-Divine timelessness than I have thus far alluded to, and so more needs to be said. To do this, I'll focus on the model of the two presentist island universes, although something similar

[27] Elsewhere I argue that there are multiple different ways of spelling out a type of presentism which can allow for there to be multiple temporally unconnected presentist timelines (Page, manuscript b).

[28] I suspect the most fruitful way to object to this way of setting out the compatibility would be to argue that island universes are impossible, or the more restricted claim that presentist island universes are impossible. I, however, am yet to see a convincing argument for either claim.

will hold for those who postulate only one temporal present and additional timelessly existing entities. On this view we should notice that there is no single present simpliciter, something some have used to characterise presentism (e.g. Zimmerman, 2005, 431; Pooley 2013, 327).[29] For there are two presents, one in U1, and another in U2, where neither of these presents are temporally related to each other. As such there is no tensed simpliciter language which will apply to the whole possible world that we are considering, and as such tenses will be relativised to each island universe.[30] Thus something that is present in U1 will not be present in U2, since U2's present differs from that of U1, and likewise something present in U2 will not be present in U1, since U1's present differs from that of U2. This is just what is implied by island universes with presentist timelines.

However, there is more too. For if we take presentness and existence to be linked, such that "Something is temporal, and existence only plays the role of absolute temporal presentness" (Leftow, 2018, 175),[31] then in addition to relativising the present to island universes, we will have to do the same with existence.[32] Thus, just as when we talk of the present we will have to be more specific and speak of present$_{U1}$ and present$_{U2}$, so we will need to qualify talk of temporal existence, resulting in exists$_{U1}$ and exists$_{U2}$.[33] There is therefore no temporal existence simpliciter. Nonetheless, we can make use of a notion, namely EXISTS, to say something like "EXISTING in a time-series is existing, but not in that series" (Leftow, 2018, 185). This, however, will be a tenseless notion since the tenses of one timeline will not apply to the other.[34]

---

[29] Zimmerman (2005, 431) seems to be worried that A-theories of time, of which presentism is a species, may have trouble distinguishing certain aspects from B-theories if they do not embrace the notion of present simpliciter. I show in (Page, *manuscript b*) that those who adopt multiple temporally unconnected presentist timelines can in still distinguish themselves from B-theories.

[30] This has some similarities to Fine's (2005, 278–280) notion of an external relativist view of tensed facts, but is different in that Fine thinks there is no unique present *within* a timeline, but an advocate of temporally unconnected timelines needs only to hold that there is no single unique present across both timelines, even if there is a unique present within each.

[31] One can restrict this so it only applies to temporal entities and allows timeless entities as well (Leftow 2018, 175).

[32] The route given here follows Leftow, but elsewhere I show that one does not have to go this route (Page, *manuscript b*). Crisp's version of presentism will not require the use of EXISTENCE and yet still gives the same broad picture.

[33] One may worry that this type of account requires that there are different kinds of existence, and thus one had better be an ontological pluralist if they wish to adopt it. This however would be a mistake. On this view, what it is to exist is tied to being present and the only reason we have different subscripted existences is because there are different presents in each timeline. Therefore, since there is no present simpliciter that ranges across all timelines, so too there is no existence simpliciter, even though it is still the case that for something to exist is for it to be present.

[34] For an argument as to why, see Leftow (2018, 188).

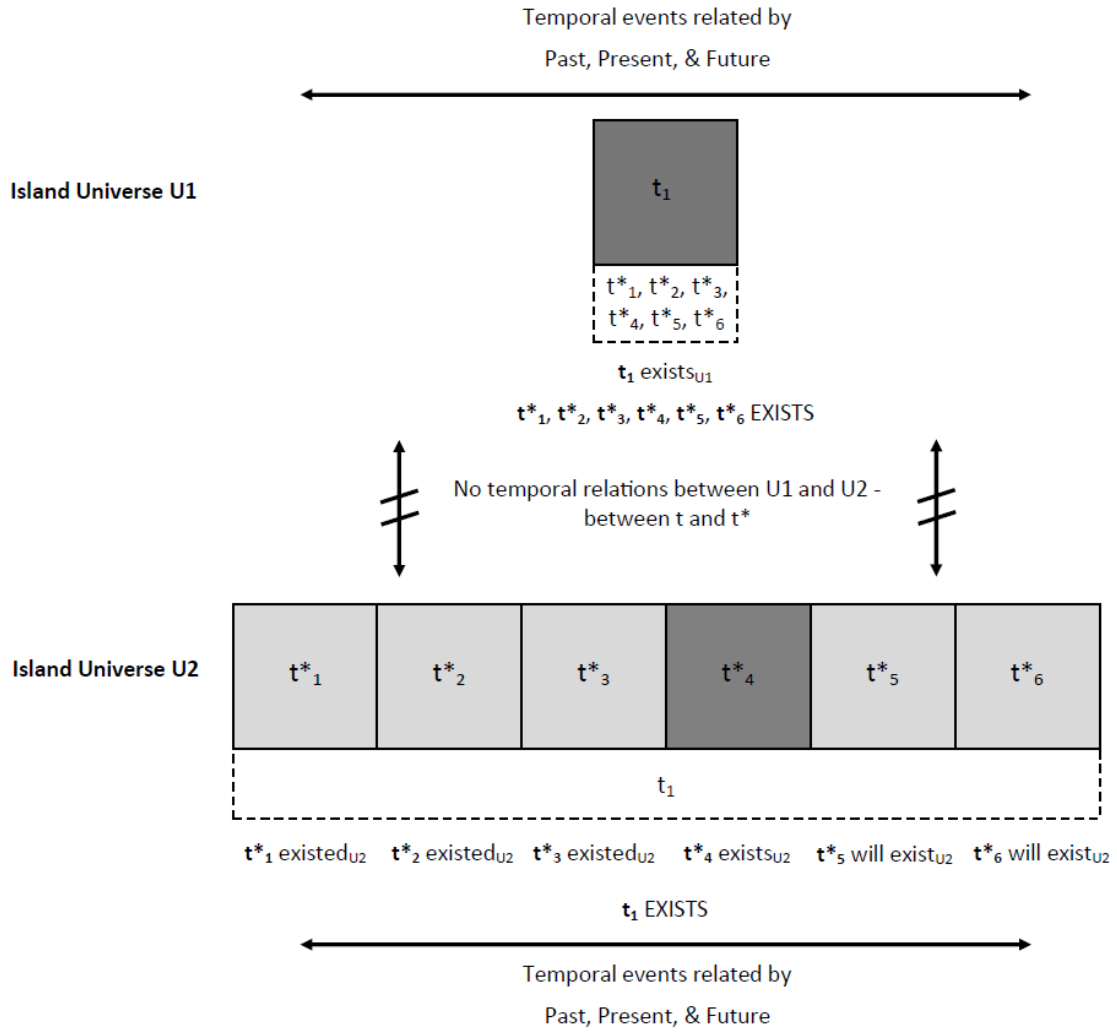
To see this more clearly let us think about what we should say if we found ourselves at time t_1 in island universe U1 and wanted to talk about time t^*_4 of U2. Time t_1 is present$_{U1}$ and therefore exists$_{U1}$, but time t^*_4 isn't present$_{U1}$, but rather present$_{U2}$, and therefore doesn't exist$_{U1}$ but exists$_{U2}$. But nothing in island universe U1 can be present$_{U2}$ nor exist$_{U2}$. Nevertheless, we should be able to say something about the whole possible world, which includes U1 and U2, when in U1. Since the tenses of U1 can't apply to t^*_4 when we are in U1, and the tenses of U2 can't be used in U1 since they don't apply, we must employ a tenseless language and a tenseless notion of EXISTENCE. Hence within U1 we can say that t^*_4 EXISTS, with this signifying that t^*_4 exists but doesn't "exist" *in* the timeline we are currently considering. Our use of EXISTS, therefore, expresses tenseless facts about other timelines, in this case, that t^*_4 EXISTS in U1.[35]

Let us now ask a further question about t^*_4 EXISTING in U1, namely whether it can begin or cease to EXIST in U1? Leftow claims it can't (2018, 187–188), and very briefly, here's why. One reason we might give for saying t^*_4 ceased to EXIST in U1 is because t^*_4 ceased to exist$_{U2}$ in U2. But if t^*_4 ceased to EXIST in U1 when t^*_4 ceased to exist$_{U2}$ in U2, it would look like the timelines weren't temporally unconnected. But they are temporally unconnected, and so t^*_4 cannot cease to EXIST in U1 for this reason. The alternative is to say that t^*_4's ceasing to exist$_{U2}$ in U2 doesn't in any way explain why t^*_4 would no longer EXIST in U1. But then it seems we are left with having to say that there's no reason why t^*_4 would cease to EXIST at one time rather than another in U1. Assuming we don't want insolvable mysteries, what we should say instead is that t^*_4 never begins or ceases to EXIST in U1. As such, t^*_4 always EXISTS in U1, no matter if t^*_4 is past, present or future in U2. Yet t^*_4 is not special in this regard, and much the same will be the case for all the other moments of time in U2, that is if they will be present at some point in W2, then they will never begin or cease to EXIST in U1. They all EXIST in U1 permanently.[36]

[35] Much the same will be the case of the view which restricts presentism to only temporal entities, that is we will have some tenses appropriate of temporal entities but not timeless ones, and therefore tenseless ways of speaking will also be required. Thus, although we use the present tense to read a sentence like "2+2=4," many philosophically minded will say that when being precise, what we are really saying is "2+2=4" tenselessly.

[36] Additionally, note that since U1 is only one instant long, anything that EXISTS in U1 will permanently exist since there are no additional moments in U1 in which something could cease to EXIST.

If this is right, then we can draw a fuller diagram of the situation at hand:



Here we can see that in U1 only t_1 $exists_{U1}$, and that t^*_1, t^*_2, t^*_3, t^*_4, t^*_5, t^*_6 merely EXIST in U1, with this represented by them being placed in the dashed portion of t_1. By contrast, in U2 only t^*_4 $exists_{U2}$, since t^*_1, t^*_2, and t^*_3 all once $existed_{U2}$, but are now in the past and so no longer $exist_{U2}$, and t^*_5 and t^*_6 will come to $exist_{U2}$ at some point in the future, but at present do not $exist_{U2}$. However, t_1 always EXISTS in the U2 timeline, and neither begins nor ceases to EXIST. Let me also make clear, so to avoid confusion, that despite the image making it appear as though, for example, t^*_1 exists twice, once in U1 and once in U2, it doesn't. t^*_1 "exists" only once, in namely U2. To put it another way, suppose a stabbing occurred in t^*_1, the stabbing would only be happening to an individual in U2, they would not also be getting stabbed a second time in U1 as well. Yet this doesn't preclude t^*_1 and the stabbing EXISTING in U1, since this just tells us that t^*_1 EXISTS in a time-series, but not the one we are currently in. Additionally, one should not get the impression that t^*_1 and the stabbing EXIST in some type of abstract way in U1. It

does not, it is just as concrete as it is in U2. Perhaps the easiest way to think about this is to follow Leftow's suggestion that U2 will "*look* eternalist" (2018, 193) to U1, with all the time-slices EXISTING, it's just that U2 is in fact presentist, since in U2 only what is present$_{U2}$ exists$_{U2}$.

Back to Presentist-Timelessness and Evil

The result of this complication is that all the time-slices of U2, what I'll call Creation, EXIST permanently in U1, that is God's eternity.[37] Nevertheless, we should remember the evil acts performed in Creation do not exist twice, but only once, in Creation. For instance, the stabbing exists at t^*_1 in Creation, but it does not exist when $t^*_{new\text{-}creation}$ is present in Creation. As such when the time is $t^*_{new\text{-}creation}$ in U1, we can say that there is no evil in Creation, even though all of the time-slices of Creation EXIST in God's eternity. Thus, evil is defeated if it requires evil to no longer *exist*, but it isn't defeated if it is required that evil no longer *exist and EXIST*. Yet, since the objection given is typically taken to be about the existence of evil in Creation, with Craig for instance writing, "creation is never really purged of evil on this view" (2001a, 66; 2001b, 214; Copan & Craig 2004, 162, n. 29),[38] the advocate of Presentist-Timelessness can say that on their account Creation is purged.[39] Additionally, the Presentist-Timelessness view allows that Creation can be renewed, such that the old creation will *exist* no more when the new creation is present, so answering a worry of Mullins (2014, 131; 2021a, 107). And finally, the view doesn't run counter to Peckham's claim that the Christian scriptures assure us that "the present evil state of affairs will pass away" (2021, 108). For on this view, the present evil in Creation will pass away when the new creation is present. Given this, I think defenders of Presentist-Timelessness can employ the presentist solution to defeating evil, *so long as* presentism itself can do the job.[40]

---

[37] One might worry that creation would therefore become eternal on this view, a concern much like Mullins's (2021b, 92–94) creation objection to timelessness. Elsewhere I show it is a mistake to think this (Page, manuscript c).

[38] This is cited with approval by Mullins (2014, 131; 2021a, 107). Elsewhere Craig writes, "On an A theory of time, once the eschaton arrives, evil, being part of the past, disappears forever from creation, thanks to the objectivity of temporal becoming." (2008, 609)

[39] The advocate of Presentist-Timelessness might say that evil leaves a trace in some sense *beyond* Creation, in that it EXISTS in God's eternity. But it seems evil leaves some trace beyond Creation on most temporalist views too, in the sense that God perfectly recollects all the evil that has occurred in Creation.

[40] Note that Mullins (2014; 2021a) talks a lot about persistence and the need for endurantism to overcome the type of worries this article is addressing. As far as I can see nothing about Presentist-Timelessness rules out endurantism, and so an advocate of this position can adopt endurantism as well.

Does Presentism Really Do the Job?

But does presentism itself, irrespective of whether God is timeless or temporal, really do the job of making it the case that there is no evil in Creation? I'm not convinced that it does. To see this, think about the truthmaker objection to presentism which Armstrong sets out nicely when he writes, "[What] truthmaker can be provided for the truth <Caesar existed>? The obvious truthmaker, at least, is Caesar himself. But to allow Caesar as a truthmaker seems to allow reality to the past, contrary to [presentism]" (2004, 146). Assume now that you are like most presentists and feel the force of this worry, such that you think truthmakers need to be posited to account for these past facts.[41] One suggested truthmaker, say of the past fact of X being murdered by Y in 1900, are Lucretian properties, namely something like the world instantiating the property 'having contained Y's murdering X in 1900'.[42] This property exists and it will exist when the time in Creation is $t^*_{\text{new-creation}}$, and so too will other properties which also serve as the truthmakers for past evils.[43]

Orilia, in a non-theological context, calls this the ugly truthmaker objection to presentism (2016, 233–235; 2018, 154–156), and provides the following response:[44]

> No matter how close the presentists' truthmakers are to the non-presentists' past events, only the latter involve, so to speak, the "real action." And it is only with the real action that there is, in the unfortunate cases, real suffering. (2018, 155; 2016, 234–235)

By "real suffering" it seems Orilia has in mind phenomenological pain/suffering, and so translating his response into the context of this paper, his reply is:

4. If X doesn't bring about some type of phenomenological pain/suffering then X isn't evil.

---

[41] Rhoda (2009, 42) is an example of someone who claims that most presentists *have* and *should* feel the force of this worry and therefore posit truthmakers. Orilia (2018, 154) thinks the same, and thus raises this type of objection to his moral superiority argument in favour of presentism (2016, 230–238; 2018).

[42] For a recent defence see (Tallant & Ingram 2020).

[43] If we are a Platonist, perhaps we can say it has EXISTENCE, but in that case we will be in the same situation as Presentist-Timelessness.

[44] Note, however, that the context in which Orilia asks and answers the objection is different, for Orilia is arguing against this objection in order to defend his view that presentism is morally superior to eternalism (2016, 230–238; 2018). Whereas the context here concerns whether presentism *alone* can make the case that some time-slice contains no evil whatsoever, even if there have been past evils. One can deny this latter claim, whilst also agreeing with Orilia that presentism is morally superior.

5. The truthmakers of past evils don't bring about phenomenological pain/suffering.
6. Therefore, the truthmakers of past evils aren't evil.

If this is right, then it seems presentism might in fact be able to do the job of removing all pain and suffering when the new creation is present despite the existence of these past truthmakers. But I don't think Orilia's reply is one that many theists will agree with. The main reason for this is that I doubt many will endorse (4), for I suggest most will think that there can be evils which do not bring about phenomenological pain/suffering. I certainly am not willing to endorse such a claim, for I think that even whilst under general anaesthetic, when I have no phenomenological awareness, one can still perform evils on me which I may never become aware of.

I suspect presentists should also disagree with this response from Orilia, since it seems that one should at least want to say past evils are evil. But in virtue of what does a presentist say that the brutal killing of some person in the past was evil? As Lebens and Goldschmidt write,

> Presentists have to accept that what makes present evils bad is very different to what makes past evils bad. Present evils are bad because they're happening. Past evils, according to the presentist, are bad because, even though they're not happening, and they don't really exist, it's a bad thing for the present to instantiate certain sorts of backward-looking properties. That's an odd consequence of presentism. (2020, 376; 2017, 10)

If presentists don't say something like this, then they will have to say that *past evils* aren't in fact bad, and it was only in the present moment in which those evils occurred that they were bad. I take it that this will be too big a bullet to bite for most presentists. As such, for presentists the existence of a present truthmaker for this past evil or the present obtaining a fact of this past evil is lamentable and therefore bad or evil.[45] For on the presentist picture all that happens is that what *is* evil has "been relocated,"[46] namely from an event that is presently occurring, to the existence of something such as presently backward-looking truthmakers or facts. Given this, presentism alone will not make it the case that there is no evil

[45] My thanks to Brian Leftow for helping me see, through discussion, exactly what it is that is evil here, although I have since remembered that Lebens and Goldschmidt also made this point too (2017, 10; 2020, 375–376).

[46] I owe this terminology to Brian Leftow in personal correspondence.

in the new creation, as there will be truthmakers for past evils that occurred within creation.[47]

Theists might also have reason to reject premise (5) of Orilla's response. For on a fairly popular view taken by divine temporalists, the divine memory plays the role of the truthmakers for past truths (Rhoda 2009; Zimmerman 2010, 801–806).[48] This alone won't get us to the denial of (5), but if one also rejects divine impassibility, with impassibility being a doctrine that Craig thinks is defended by very few on the contemporary scene (2011), then these existing divine memories may well cause pain and suffering to God just as they likely did when they were occurring.[49] This option is obviously not open for those who think God is timeless, since a timeless God does not have any memories, since nothing is past to Him. Yet defenders of timelessness might still deny (5) since passibility is compatible with timelessness,[50] and if what EXISTS in God's eternity can play the role of truthmaker for past truths within Creation, then perhaps what EXISTS in God's eternity means He eternally feels pain. As such an advocate of Presentist-Timelessness has the ability to deny (5) too.[51]

---

[47] Whilst I've focused here on Lucretian properties, I take it other such truthmakers would result in the same outcome. See Ingram and Tallant (2022) for some examples. One might wonder whether a privation theory would help respond to the concerns raised here. I'm unconvinced, but since I don't think that my interlocutors hold to a privation account I'll ignore that complication here.

[48] Note that those who object to (4) can claim that the evil here is relocated to God's memories, and then they need not go down the route of questions concerning passibility and impassibility.

[49] The ability of God to suffer in some way due to the actions of His creatures is exactly why impassibility is rejected by many. For an influential defence of this see Fiddes (1988), in which he writes, "The sorrow of God because his people reject his loving care leads to a unique kind of pain which is ascribed to God, a state of feeling which is characterized by the prophets as *a blend of love and wrath*. This is presented as a pathos which is God's own pathos." (1988, 20)

[50] To see this note that immutability does not entail impassibility, and so a timeless God who is immutable need not be impassable (Leftow, 2005, 59–66). Timelessness would only rule out impassibility *if* the way in which Creation externally affected God would result in God being temporally related to creation. It's far from clear that this must be the case.

[51] Orilia does discuss another argument against his argument for presentism based on its moral superiority, namely the "ugly history" or "radical" objection, which holds that the moral value of a world does not just depend upon that which exists (2016, 235–238; 2018, 156–159). Whilst this is important given his context, it is not important here since, all this notion of defeat requires is that there is a time-slice in which there exists no evil, rather than us thinking about the moral status of a presentist world and all of its time-slices compared with an eternalist one. For even if the moral value of a presentist world was overall worse than an eternalist one, so long as there was some time-slice in the new creation which had no *existing* evil, evil would still be defeated on presentism and not on eternalism, since despite the eternalist world being more morally valuable overall, evil can still be said to exist. I admit that this does seem odd, and that if instead we define defeat in such a way that it is to do with the moral status of the world and its history then Orilia's response to this objection, and Graziani (2021) rebuttal becomes important. Note too that there is some reason to think that Philosophical Theologians should think that moral

The advocate of Presentist-Timelessness might wish to claim that their view is superior to the above divine temporalist position, since although both will admit that not all evil is completely gone, the evil is relegated to EXISTENCE within God's eternity in Presentist-Timelessness, whilst the divine temporalist, in virtue of thinking God exists within time, might have to say that Creation still contains evil in virtue of God finding Himself within Creation's timeline. Here I suggest the divine temporalist should reply that God's memories shouldn't be thought of as part of creation, and that they exist in metaphysical time and not the physical time of Creation (Padgett, 1992; DeWeese, 2004). If such a move is allowed it seems the divine temporalist will be in a similar situation to the advocate of Presentist-Timelessness in that evil no longer exists in Creation, but nevertheless still leaves a trace for God.

Perhaps it can be claimed that the notion of defeat we are working with should be weakened a little, so to remove even this trace. For it might be suggested the objection we are considering really just concerns evils not existing "as sturdily as ever at its various locations in space-time" (Craig 2001a, 66; 2001b, 214; 2008, 610; Copan & Craig 2004, 162, n. 29), and on this view, if the truthmakers of past evils are Platonic Lucretian properties, divine memories, or merely EXIST, then evil doesn't exist within Creation's space-time and therefore the relevant evils are defeated. Whilst I'm inclined to agree, it does seem some might think this move is a little *ad hoc*, since we have conveniently removed those evils we cannot eliminate from our definition of defeat. Maybe we can find some additional motivation for such a restriction, perhaps through scriptural data, but until we do, this type of move looks a little dubious.

So to avoid this, one might suggest another route for removing this trace of evil, namely by claiming that there are in fact no truthmakers for past-truths, and thus there are no presently existing truthmakers for any past evils.[52] This move won't help the advocate of Presentist-Timelessness, for the evils in Creation still EXIST in eternity, but does it help the divine temporalist so that they can say that on their view in the new creation there is no evil whatsoever? I suspect that on most views it won't.

In order to see this it will first be helpful to note that I am assuming the temporal God we are considering is omniscient, and therefore will have a perfect memory of all past events.[53] If one wishes to hold, like Ward (2001, 107) that an

status of the world and its history is important, for the removal of certain parts of history is what Lebens and Goldschmidt (2017; 2020) attempt to accomplish in some of their work.

[52] Craig (2017, 387–392) and Merricks (2007) are examples of people who hold there are no presently existing truthmakers for past truths.

[53] One might wonder whether human memory in the new creation will cause problems similar to those that I am about to pose, but I think it is much less obvious that humans must actively remember all past events, especially when compared with God.

omniscient God can ignore things or forget, or instead that God is not essentially omniscient given Kenotic theology (Davis 2006; Forrest 2000, 130–132), and so His knowledge does not include all past memories, then the argument I am about to make will not hold.[54] However, I'm not inclined to agree with either of these views, and take it that a perfect cognizer cannot be ignorant of the past and so will ignore this reply here.[55] Therefore, given God's omniscience, His memories of past evils will still exist in the new creation, even though now, since we are rejecting the need for presently existing truthmakers for past truths, these memories will no longer be truthmakers of those truths. I will also assume, along with many of my interlocutors (Mullins 2016, 202; Davis 1983, 4; Craig 2011; Ward 2001, 162–163; Peckham 2021, chap. 2), that God is passible, such that He can be emotionally moved by things, and that typically the types of emotions passibilists want to allow God to feel are negative ones such as pain, grief, etc.

With that as background, suppose that there are evils in Creation which were gratuitous.[56] I take it that when this evil occurred in Creation God knew about it and felt anguish over its occurrence, given omniscience and passibilism.[57] Yet I'm inclined to think that if God felt anguish over this evil when it occurred, then He will also feel some anguish over the memory of it in the new creation, for it is an evil that brought about no greater good. For it would seem odd if God stopped feeling anguish over this evil at some later point in His life, since nothing about the evilness of this evil seems to change, and no greater good comes of it. Yet since I think experiencing anguish can be considered an evil, for surely the world would be better if there was no such anguish, it turns out that there is evil at the present time of the new creation, namely God's present anguish over His memories of past gratuitous evils.[58]

---

[54] One reason Ward (2001, 162) and Davis (2001, 107) might go this route is that they don't think evil should be permanently present to God, and whilst on divine timelessness it always is, on divine temporalism it won't, so long as God is not required to perfectly recollect the past.

[55] Craig, for instance, agrees writing, "it is logically impossible for God, being essentially omniscient, to forget what He experienced as present." (2001c, 39)

[56] Whilst I don't think this view is the majority position, there does seem to be a growing number of theists who think that gratuitous evils are compatible with theism, for instance, see: van Inwagen (2006, chap. 6), Howard-Snyder & Howard-Snyder (1999), Hasker (1992), and Mooney (2019).

[57] If one does not like the term anguish, substitute it for righteous anger (Psalm 7:11; Deuteronomy 9:22), hate (Proverbs 6:16; Psalm 11:5), grief (Genesis 6: 5–6), or sorrow (Mark 14:34).

[58] The fact that God experiences anguish does not mean that His overall emotional state is not one of joy. For I take it that one can have conflicting emotions at the same time, and with different strengths. One might also be able to run the problems I state here, although their force would seem weaker, on the view which takes God to be all happy, since the infinite happiness of God, although infinite is still decreased by anguish, such that God would be happier without it (Stenberg, 2019, 435–349). However, I cannot pursue this further here.

One might reply that there is a difference here between experiencing something as present and experiencing something as past, and that this might imply that God doesn't feel anguish when He experiences something as past, such as when He recollects the past. I find this very difficult to believe, since it seems obvious to me from experience that memories can bring about suffering. Perhaps it will be the case that experiencing something as past will typically bring about less suffering, but I think there will be suffering nonetheless and that the vividity with which one remembers the occurrence will likely having a bearing on how much suffering one feels when recalling it.[59] Yet since God's memory is perfect, He will recollect things as vividly as they possibly can be, and as a result I take it that He will experience close to the same levels of suffering as He experienced when the evil was present.[60]

Another reply might instead claim that although God does not forget, perhaps He never chooses to recollect certain memories and therefore feels no anguish from them.[61] However, I think that God's memories form a subset of God's knowledge and that God's knowledge is more perfect if it is always present to Him rather than requiring recollection. As such, I think God can and does hold all his memories within His present attention, with his infinite intellectual capabilities making this possible. Yet because of this, God will always feel anguish regarding the memory of this past gratuitous evil since this memory is always in His present attention.

A more obvious response is to claim that there are no gratuitous evils, and that therefore every evil in fact brings about some greater good. Thus, in the new creation God doesn't experience any anguish from memories concerning gratuitous evils. Sadly, I'm not convinced this will be sufficient either.

To see this first note that there are two different views concerning what type of greater good is required so to overcome gratuitous evils. On the first view the greater good just needs to benefit reality in some way (Mawson 2011), whilst on the second view it needs to benefit the individual who suffered the evil (Adams 1999, 28; Stump 2022, 5). Start with the former view and think about the following case. My daughter dies in some truly terrible way, and in virtue of this certain governmental policies are put in place which prevents any similar evil, which

[59] I say typically, because perhaps in some cases the memories of past evils might bring about more suffering as they are being remembered than when the evil was present, with this perhaps being the case in some trauma victims.

[60] That God's memory can act in this way is important for the divine temporalist since it is one of the ways to at least *weaken* an argument for timelessness, which is sometimes called the incompleteness of temporal life (Leftow, 1991, 278–279 – whilst Leftow formulates this argument, he doesn't name it this).

[61] Perhaps this is why human memory is not an issue in the new creation, since they choose not to recollect, or maybe their finite intellectual capacities are so overcome by the joy of presently being with God that they don't have the intellectual capability to recall the past as well.

would have occurred, from occurring again. This is a great good, and it comes about due to my daughter's death, but it seems to me wrong to think that in virtue of this it is no longer proper for me to feel anguish over her death. This great evil happened to her, and she, after all, has not benefited from it. Anguish is what I will feel, and I think I should feel, even if the anguish is lessened by the knowledge that some great good has come about for the world in virtue of her death. I take it that the same will be true of God, namely that He will feel some anguish for His creatures that have suffered a great evil and who did not themselves receive any good from it and His memories of this occurrence too. This, I take to be consistent with God being justified in allowing the evil to take place so to bring about this great good, but it's just that God will suffer in some way in virtue of His memories of it.

What then about cases where the individual is the one who receives the great good? Think now about the following case. A woman is raped, a terrible evil, but in virtue of this becomes a mother to a child that she loves and devotes her life to, something she comes to think of as a great good.[62] It still seems to me that even in this case, where the woman herself is the recipient of the great good, that she can feel anguish when she recalls the act of rape by which she fell pregnant and resulted in what she considers to be this great good. So too in God's case, He can feel anguish over His recollection of past evils, even though the person who experienced the evil has received a great good.[63] As before, I suspect the anguish will be of a greater severity than when the event is actually occurring, but it will be anguish nonetheless. So once again we will still have evil in the new creation in virtue of God's memories and thus the trace of evil is still not removed.

Perhaps, a defender of presentist-timelessness who also takes God to be passible can claim that they are in a better situation than the divine temporalist who thinks God is passible. To see this first note that we can think about our thinking, and think about thinking about our thinking, with this process being iterated. As such, God in the new creation can be thinking about all the evil that occurred in Creation, and then think about thinking about this evil and so on. I've suggested it's likely that God can feel anguish over all the evil that occurred in Creation, and I think God can also feel anguish when thinking about thinking about these evils, and so on. Perhaps at each iteration the level of anguish God feels will be less, but the level may never reach zero, and therefore at each

[62] For an interesting case study which has *some* similarities to the case given here, see Kantengwa (2014).

[63] Note that I don't think saying that it was necessary that an evil occurred in order to bring about a great good will remove one's anguish here. For one can still feel sorrow or anguish that this is the case, and/or that someone still had to go through that experience.

iteration God feels some anguish.[64] Now I take it that a passible temporal God and a passible timeless God will both receive anguish from this in the same way. However, I think there is an additional way in which a temporal passible God can generate anguish, namely in virtue of His memories. Suppose then that it is now t'_2 in the new creation, God will have memories of thinking about the evils that occurred in Creation at t'_1 in the new creation, and memories of thinking about thinking about evil at t'_1, etc. At t'_3 in the new creation God also gets a further iteration to do with memory, in that he now has memories of His thinking about the evils that occurred in Creation at the time t'_1 in the new creation, but also remembers remembering thinking about the evils that occurred in the new creation at t'_1 in the new creation at the time t'_2 in the new creation, and so on. Given this, as each new moment passes in the new creation God gets more and more memories, and more remembering's of remembering. Once more, perhaps we can say that each iteration of memory generates less anguish, but that it will nevertheless be the case that the anguish never converges on zero. As such, a temporal passible God has another place in which God accrues anguish compared to a timeless passible God, since a timeless passible God cannot accrue anguish in this way since He has no memory, and thus we might think a passible presentist-timeless God has less anguish overall and so should be preferred.

I suspect more details are required here but let me note one reply that a temporalist passibilist might make. They might say that although a timeless God doesn't get any anguish from memories, the anguish a temporal passible God feels about Creation when God exists at a time in the new creation will be less that a timeless God feels about Creation. The reason for this, so they say, is that for the temporal God the evils of Creation have passed, whereas for the timeless God they are permanently present, and that evils which are remembered cause less anguish than those evils which are in one's present.[65] The defender of divine timelessness, may reply along the following lines, claiming that we cannot speak of these evils as truly "present" to God, since the present of Creation and the present of eternity differ. As such the evils of Creation just occur to God, with this use of occurrence being tenseless. The question will then be whether the anguish one experiences from something "occurring" or being "present" is the same or different. If it is the same then the question will be whether the anguish a timeless passible God feels is the same in amount and/or intensity as that of a temporal passible God and His multitude of memories, or if it differs then we will want to know in what way it differs and why there is this difference. At

[64] If it did reach zero at some point we'd have to answer the question as to where God stops feeling anguish and the worry about arbitrary stopping points.

[65] This thought denies what was said in footnote 59, and it does seem inappropriate to compare God's emotional state to that of someone experiencing trauma.

present it doesn't seem obvious to me who will win out, or whether we'll be left with a tie.

In any case, it is clear from what I've said that I think there will be evil in the new creation if one takes God to be temporal and passible. The options at this stage, with us still assuming that there are no truthmakers for past truths, seem to be the following. First, admit the stain of evil never leaves the new creation *completely*, but that Creation itself is still purged, thus leaving one in a similar situation to presentist-timelessness or views where truthmakers for past truths are Platonic abstracta. Second, argue that the suffering God experiences in virtue of His memories isn't in fact evil and so although there is suffering there is no stain of evil in the new creation. If one can pull this off, then this would mean that divine temporalism is better off than a presentist-timeless God since the temporalist does not have to say that evil EXISTS. Or finally, the temporalist could say that God is impassible and so isn't moved by Creation, and as a result neither the evils that occurred in Creation nor the memories of them will cause God to suffer.[66] This too will result in divine temporalism being better off than presentist-timelessness since they do not need to affirm that evils EXIST, but I don't know of any divine temporalist who holds to impassibility.

The result of all this is that a presentist view of time, along with divine temporalism and a number of controversial views will provide the result that there is no evil in the new creation. However, the question will be whether this benefit outweighs the costs that such a view requires.

Summing Up

Although advocates of the argument we are addressing think the defeat of evil is accomplished by adopting presentism, I've argued here that this only seems to be the case if one adopts some additional and controversial assumptions.[67] Since many of these are far from popular, I suspect many will think that we should instead change our understanding of defeat, such that it does not require that

[66] Note that an advocate of Presentist-Timelessness can also reap some of the rewards of the impassibility move, since this will shield God from experiencing evil in virtue of what occurs in Creation and thinking about this occurrence. Nevertheless, evil will still EXIST in eternity.

[67] Let me also note a more general concern, namely that thinking an existing presentist time-slice in the new creation is just the wrong way of thinking about how evil is defeated. For it seems, at least given what the Bible says about God and the defeat of evil, that more needs to be done on God's part than merely create a new presentist time-slice. (For some further thoughts along these lines see: Lebens, forthcoming; Hollingsworth, forthcoming, 13–16). I suspect the correct response here is just to claim that presentism should merely be taken as a necessary condition of defeat, and that it is probably not the most important condition either.

there is *no* evil existing.[68] However, it might be the case that this redefinition means that Presentist-Timelessness does just as well as divine temporalist accounts, and/or may even leave the door open for eternalists to claim that evil can be defeated on their view of time.[69] If this is the case, then the argument will lose all its sting, but investigating whether it is will have to wait for another time. For now, we can conclude that the argument, as typically presented, is likely much less forceful against those who hold to presentist-timelessness, and that given the way defeat has been defined, most divine temporalists who adopt presentism, in virtue of their other commitments, cannot fully defeat evil either!

Bibliography

Adams, M. M. 1999. *Horrendous Evils and the Goodness of God*. Ithaca: Cornell University Press. https://doi.org/10.7591/9781501735929.

Adams, M. M. 2006. *Christ and the Horrors*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511607585.

Anslem. 2007. *Anslem: Basic Writings*. Translated by T. Williams. Indianapolis: Hackett Publishing.

Aquinas. *De Potentia Dei*. https://aquinas.cc/la/en/~QDePot

Armstrong, D. M. 2004. *Truth and Truthmakers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511487552.

Augustine. 1998. *Confessions*. Translated by H. Chadwick. New York: Oxford University Press.

Baron, S. and K. Miller. 2014. "Causation in a timeless world." *Synthese* 191: 2867–2886. https://doi.org/10.1007/s11229-014-0427-0.

Baron, S. and J. Tallant. 2016. "Monism: The Islands of Plurality." *Philosophy and Phenomenological Research* 93: 583–606. https://doi.org/10.1111/phpr.12270.

Boethius. 2001. *Consolation of Philosophy*. Translated by J. C. Relihan. Indianapolis: Hackett Publishing.

[68] Another option may be to instead become "hyper-presentists," adopting Lebens and Goldschmidt's proposal (2017; 2020). However, I suspect defenders of this argument against timelessness won't like this either since "hyper-presentism" is significantly different from presentism, with some advocates of the argument having explicitly voiced their dislike of hypertimes (Mullins, manuscript; Craig, 2008, 599). Note also, that I think that even if an advocate of Presentist-Timelessness adopted "hyper-presentism," they would still have to think that evil EXISTS in God's eternity.

[69] Even if it does, a presentist could still run an argument like Orilia's (2016, 251–252; 2018, 160) which claims that God would have a good reason to prefer creating a presentist world because it is morally superior, and therefore if we think God has created we have additional reasons for thinking the metaphysics of time is presentist. The question will be whether this argument is successful, for some discussion see Graziani (2021), and even if it is, it will not rule out Presentist-Timelessness unless one can show that such a view is incoherent.

Bourne, C. 2006. *A Future for Presentism*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199212804.001.0001.

Bricker, P. 2020. *Modal Matters*. New York: Oxford University Press.

Cameron, R. 2015. *The Moving Spotlight*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198713296.001.0001.

Copan, P. and W. L. Craig. 2004. *Creation out of Nothing*. Grand Rapids, MI: Baker Academic.

Craig, W. L. 2000. *The Tensed Theory of Time*. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-9345-8.

Craig, W. L. 2001a. 'Response to Paul Helm: William Lane Craig', in G. E. Ganssle, ed., *God and Time*. IL: InterVarsity Press.

Craig, W. L. 2001b. *Time and Eternity*. Wheaton, IL: Crossway. https://doi.org/10.1007/978-94-017-1715-1.

Craig, W. L. 2001c. *God, Time, and Eternity*. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-1715-1.

Craig, W. L. 2008 "Time, Eternity, and Eschatology," in *The Oxford Handbook of Eschatology*, edited by J. L. Walls. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195170498.003.0036.

Craig, W. L. 2011. "#213 Divine Impassibility and the Crucifixion.' *Reasonable Faith Q & A*, URL: https://www.reasonablefaith.org/writings/question-answer/divine-impassibility-and-the-crucifixion.

Craig, W. L. 2017. *God and Abstract Objects*. Dordrecht: Springer. https://doi.org/10.1007/978-3-319-55384-9.

Crisp, T. M. 2004a. "On Presentism and Triviality," *Oxford Studies in Metaphysics*, edited by Dean W. Zimmerman, 15–20. Oxford: Oxford University Press.

Crisp, T. M. 2004b. "Reply to Ludlow," *Oxford Studies in Metaphysics*, edited by Dean W. Zimmerman, 37–46. Oxford: Oxford University Press.

Davis, S. T. 1983. *Logic and the Nature of God*. London: Macmillan Press. https://doi.org/10.1007/978-1-349-06352-9.

Davis, S. T. 2001. "Free Will and Evil," in *Encountering Evil*, edited by S. T. Davis. Louisville, KY: Westminster John Knox Press.

Davis, S. T. 2006. *Christian Philosophical Theology*. New York: Oxford University Press. https://doi.org/10.1093/0199284598.001.0001.

DeWeese, G. J. 2004. *God and the Nature of Time*. Hampshire UK: Ashgate.

Fiddes, P. S. 1988. *The Creative Suffering of God*. New York: Oxford University Press.

Fine, K. 2005. *Modality and Tense*. New York: Oxford University Press. https://doi.org/10.1093/0199278709.001.0001.

Forrest, P. 2000. "The Incarnation: a philosophical case for kenosis." *Religious Studies* 36: 127–140. https://doi.org/10.1017/S003441250000514X.

Graziani, E. 2021. "Presentism and the Pain of the Past: A Reply to Orilia." *Philosophical Inquiries* 9: 53–66.

Hasker, W. 1992. "The Necessity of Gratuitous Evil." *Faith and Philosophy* 9 :23–44. https://doi.org/10.5840/faithphil1992911.

Hasker, W. 2011. "A non-classical alternative to Anselm," *Philosophy of Religion: The Key Thinkers*, edited by in J. J. Jordan. London: Continuum.

Helm, P. 2010. *Eternal God* 2nd ed. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199590391.001.0001.

Hollingsworth, A. 2023. "Eschatology, the Elimination of Evil, and the Ontology of Time." *TheoLogica: An International Journal for Philosophy of Theology and Theological Philosophy* 8, 1: 1–19. https://doi.org/10.14428/thl.v8i1.74563.

Howard-Snyder, D. and F. Howard-Snyder. 1999 "Is Theism Compatible with Gratuitous Evil?" *American Philosophical Quarterly* 36: 115–130.

Ingram, D. 2019. *Thisness Presentism*. New York: Routledge. https://doi.org/10.4324/9780429452277.

Ingram & Tallant. 2022. "Presentism," in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta and U. Nodelman. URL: https://plato.stanford.edu/archives/win2022/entries/presentism/.

Kantengwa, O. 2014. "How Motherhood Triumphs Over Trauma Among Mothers With Children From Genocidal Rape in Rwanda." *Journal of Social and Political Psychology* 2: 417–434. https://doi.org/10.5964/jspp.v2i1.334.

Latham, A. J. and K. Miller, K. 2020. "Time in a one-instant world." *Ratio* 33: 145–154. https://doi.org/10.1111/rati.12271.

Lebens, S. and T. Goldschmidt. 2017. "The Promise of a New Past." *Philosophy Compass* 17: 1–25.

Lebens, S. & T. Goldschmidt. 2020. "Hyper-Past Evils: A Reply to Bogdan V. Faul." *Open Theology* 6: 374–377. https://doi.org/10.1515/opth-2020-0119.

Lebens, S. *Forthcoming*. "Time for Some Robust Atonement." *Faith and Philosophy*.

Leftow, B. 1991. *Time and Eternity*. Ithaca: Cornell University Press.

Leftow, B. 2002. "The Eternal Present," in *God and Time*, edited by G. E. Ganssle, & D. M. Woodruff. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195129656.003.0002.

Leftow, B. 2005. "Eternity and Immutability," *The Blackwell Guide to Philosophy of Religion*, edited by W. Mann. New York: Blackwell.

Leftow, B. 2009. "Anselmian Presentism." *Faith and Philosophy* 26: 297–319. https://doi.org/10.5840/faithphil200926316.

Leftow, B. 2018. "Presentism, Atemporality, and Time's Way." *Faith and Philosophy* 35: 173–194. https://doi.org/10.5840/faithphil2018326100.

Lewis, D. 1972. *On the Plurality of Worlds*. Malden, MA: Blackwell.

Lewis, D. 2004. "Tensed Quantifiers," in *Oxford Studies in Metaphysics*, edited by Dean W. Zimmerman, 3–14. Oxford: Oxford University Press.

Markosian, N. 2004. "A defence of presentism," in *Oxford Studies in Metaphysics*, edited by Dean W. Zimmerman, 47–82. Oxford: Oxford University Press.

Mawson, T. J. 2008. "Divine Eternity." *International Journal for Philosophy of Religion* 64: 35–50. https://doi.org/10.1007/s11153-008-9162-0.

Mawson, T. J. 2011. "Theological Individualism." *European Journal for Philosophy of Religion* 3: 139–159. https://doi.org/10.24204/ejpr.v3i1.385.

McCann, H. J. 2012. *Creation and the Sovereignty of God*. Bloomington: Indiana University Press.

Merricks, T. 2007. *Truth and Ontology*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199205233.001.0001.

Mooney, J. 2019. "How to Solve the Problem of Evil: A Deontological Strategy." *Faith and Philosophy* 36: 442–462. https://doi.org/10.5840/faithphil20191121132.

Mullins, R. T. 2014. "Four-Dimensionalism, Evil, and Christian Belief." *Philosophia Christi* 16: 117–137. https://doi.org/10.5840/pc20141617.

Mullins, R. T. 2016. *The End of a Timeless God*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198755180.001.0001.

Mullins, R. T. 2020. "The Problem of Arbitrary Creation for Impassibility.' *Open Theology* 6: 392–406. https://doi.org/10.1515/opth-2020-0110.

Mullins, R. T. 2021a. "Personal identity over time and life after death," *Death, Immortality, and Eternal Life*, edited by R. Byerly. New York: Routledge. https://doi.org/10.4324/9781003058380-7.

Mullins, R. T. 2021b. "Classical Theism," *T&T Clark Handbook of Analytic Theology*, edited by J. M. Arcadi & J. T. Turner Jr. London: T & T Clark. https://doi.org/10.5040/9780567681317.ch-007.

Mullins, R. T. *Manuscript*. *From Divine Timemaker to Divine Watchmaker*.

Orilia, F. 2016. "On the Existential Side of the Presentism-Eternalism Dispute." *Manuscrito* 39: 225–254. https://doi.org/10.1590/0100-6045.2016.v39n4.fo.

Orilia, F. 2018. "The Moral Desirability of Presentism," *God, Time, Infinity*, edited by M. Szatkowski. Berlin: De Gruyter. https://doi.org/10.1515/9783110594164-010.

Padgett, A. G. 1992. *God, Eternity and the Nature of Time*. London: Macmillan. https://doi.org/10.1057/9780230376519.

Page, B. 2023. "The creation objection against timelessness fails," *International Journal for Philosophy of Religion* 93: 169–188. https://doi.org/10.1007/s11153-022-09844-z.

Page, B. *forthcoming*. *Modelling the Divine*. Cambridge University Press.

Page, B. (*manuscript a*). "Evil Eternalism? "

Page, B. (*manuscript b*). "Are Multiple Temporally Unconnected Presentist Timelines Possible?"

Page, B. (*manuscript c*). "O PRECREATION, Where Art Thou?"

Peckham, J. C. 2021. *Divine Attributes*. Grand Rapids, MI: Baker Academic.

Pooley, O. 2013. "Relativity, The Open Future, and the Passage of Time." *The Aristotelian Society* 113: 321–363. https://doi.org/10.1111/j.1467-9264.2013.00357.x.

Rea, M. C. 2003. "Four-Dimensionalism" in *The Oxford Handbook of Metaphysics*, edited by M. J. Loux, & D. W. Zimmerman. New York: Oxford University Press.

Rhoda, A. R. 2009. "Presentism, Truthmakers, and God." *Pacific Philosophical Quarterly* 90: 41–62. https://doi.org/10.1111/j.1468-0114.2009.01328.x.

Rogers, K. A. 2000. *Perfect Being Theology*. Edinburgh: Edinburgh University Press.

Rogers, K. A. 2009. "Back to Eternalism: A Response To Leftow's "Anselmian Presentism"." *Faith and Philosophy* 26: 320–338. https://doi.org/10.5840/faithphil200926317.

Rogers, K. A. 2011. "The classical view of God," in *Philosophy of Religion: The Key Thinkers*, edited by J. J. Jordan. London: Continuum.

Sider, T. 2001. *Four-Dimensionalism*. New York: Oxford University Press. https://doi.org/10.1093/019924443X.001.0001.

Skow, B. 2022. "The Moodless Theory of Modality: An Introduction and Defence." *Australasian Journal of Philosophy* 100: 279–295. https://doi.org/10.1080/00048402.2021.1894585.

Stenberg, J. 2019 "The All-Happy God." *Faith and Philosophy* 36: 423–441. https://doi.org/10.5840/faithphil20191125133.

Stump, E. and N. Kretzmann. 1981 "Eternity." *The Journal of Philosophy* 78: 429–458. https://doi.org/10.2307/2026047.

Stump, E. 2010. *Wandering in Darkness*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199277421.001.0001.

Stump, E. 2018. *Atonement*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780198813866.001.0001.

Stump, E. 2022. *The Image of God*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780192847836.001.0001.

Tallant, J. 2008. "What Is It to "B" a Relation?" *Synthese* 162: 117–132. https://doi.org/10.1007/s11229-007-9173-x.

Tallant, J. 2014. "Defining Existence Presentism." *Erkenntnis* 79: 479–501. https://doi.org/10.1007/s10670-013-9499-3.

Tallant, J. & Ingram, D. 2020 "A Defence of Lucretian Presentism." *Australasian Journal of Philosophy* 98: 675–690. https://doi.org/10.1080/00048402.2019.1697709.

Tallant, J. & Ingram, D. 2021. "The Rotten Core of Presentism." *Synthese* 199: 3969–3991. https://doi.org/10.1007/s11229-020-02965-w.

Inwagen, P. van. 2006. *The Problem of Evil*. New York: Oxford University Press.

Ward, K. 2001. "The temporality of God." *International Journal for Philosophy of Religion* 50: 153–169. https://doi.org/10.1023/A:1012043228814.

Wierenga, E. R. 1989. *The Nature of God*. Ithaca: Cornell University Press.

Zimmerman, D. W. 2005. "The A-Theory of Time, The B-Theory of Time, and 'Taking Tense Seriously." *Dialectica* 59: 401–457. https://doi.org/10.1111/j.1746-8361.2005.01041.x.

Zimmerman, D. 2010. "The A-Theory of Time, Presentism, and Open Theism," in *Science and Religion in Dialogue* Vol. 2, edited by M. Y. Stewart. Malden, MA: Wiley Blackwell.

Published Online First: August 28, 2023

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.66863

From Vices to Corruption to Misanthropy

IAN JAMES KIDD
University of Nottingham
ian.kidd@nottingham.ac.uk

Abstract: The main part of the paper describes the deep connections between the concepts of vices, corruption, and misanthropy. I argue that the full significance of the concept of human vices or failings is only fully appreciated when it is connected to an account of the ways that our social practices and institutions are corrupting, in the sense of facilitating or encouraging the development and exercise of those failings. Moreover, reflection on failings and corruption can lead us to misanthropy, defined in a revisionary sense as a negative, critical verdict on the collective moral character and performance of humankind as it has come to be. At the end of the paper, I tentatively ask if there can be forms of Christian misanthropy.

Keywords: Christianity, Corruption, Misanthropy, Sin, Vices

Introduction

Critical moral appraisal can be directed at individual people as well as human culture or 'forms of life'. I propose one way that appraisal of individuals for their failings could be connected to a more systematic appraisal of human life or humankind. The connection uses the concepts of *failings*, *corruption*, and *misanthropy*. By the latter, I mean negative critical judgement on the collective moral character and performance of humankind as it has come to be. I defend this revisionary conception of misanthropy after the prior tasks of explaining the concepts of *failings* and *corruption*. It emerges that misanthropy involves experiencing the human world as vicious and corrupting. At the end of the paper, I tentatively ask if there can be forms of Christian misanthropy and tentatively suggest an answer in the affirmative.

1. Failings

I use the term *failings* to include vices in the traditional sense of failings of character and also a much wider array of bad or problematic features of humankind. On my view, lack of necessary skills, culpable ignorance, and narrowness of experience are all failings. There are connections between

traditional vices and these failings. Certain kinds of ignorance can be caused or sustained at least in part by epistemic vices such as arrogance and dogmatism as well as by social conditions and patterns of consistent bad luck. I prefer the term *failings* for its scope but also because 'vice' tends to connote (a) broadly moral failings understood in (b) a broadly Aristotelian sense (Cooper 2018, ch. 4). Without doubting the importance of Aristotle's character ethics, there are reasons to prefer a broader term that is more neutral with respect to specific moral theories.

One reason is that a vice theorist may not want to endorse all of Aristotle's claims about character failings—for instance, his conviction that a single virtue is always accompanied by no more than two vices, one each of deficiency and excess. It may be true this 'two vices' model applies to some virtues, but no reason to think that it applies to all. Some virtues could have several vices orbiting them (see, e.g., Tanesini 2021, chs. 5–6). Other virtues could have more vices on the excess side than the deficiency side—or *vice versa*. If we cling to the Aristotelian account, we may find ourselves at risk of failing to identify the fuller range of vices because we stopped counting at two.

A second reason to prefer the more theory-neutral term 'failing' is that it makes it easier for us to include and draw upon other philosophical traditions. The early Buddhist tradition, for instance, offers extensive catalogues and analyses of our 'cankers', 'taints', and 'defilements', a variety of failings which have conceptual and psychological distinctiveness that is not well accommodated in the terms of Aristotelian virtue theory (cf. *Samyutta Nikāya* SN 3: 76–77 and 5: 51). The Christian moral tradition also offers its own ways of conceptualising human failings, such as the 'venial', 'mortal' or 'deadly' sins and 'capital vices' (cf. DeYoung 2020, ch. 2, Taylor 2006, chs. 2 and 7). Such differences are not merely terminological: they register radically different accounts of the human condition and the nature of reality—whole 'cosmic outlooks' in which particular ethical visions are embedded (McPherson 2020, 115ff).

Outside of these major traditions, there is also the wonderful profusion of terms and discourses for describing our failings. We speak of vices, sins, defects, shortcomings, and foibles. Depending on preference and circumstance, we draw liberally on medical metaphors ('diseases of the mind') and aesthetic concepts ('moral ugliness'). We also innovate and extemporise using any available rhetorical tropes, moral values, cultural norms, and other resources ('assholes', 'jerks').

Philosophical accounts of human failings should address ontological and normative questions about the nature, origins and badness of those failings. This often extends to issues in moral and empirical psychology and the ways that our failings relate to social roles and practices. It has also at times been popular to

relate our failings to more substantive accounts of human nature—of, for instance, our *telos* or our status as God's creation or our evolved evolutionary nature. At this point, accounts of human failings often invoke worldviews or metaphysical visions: the more-or-less systematic accounts of the nature of reality relative to which certain attitudes or dispositions can appear as failings (think of such spiritual failings as impiety or hubris, defined as failures to show proper respect for God). Not all reflections on our failings need to broach these deep anthropological and metaphysical issues. A lot of work can be done at a more local level and, anyway, not everyone is sympathetic to 'big-picture' styles of philosophising about human moral life. In any case, in this paper, I only need to emphasise some very general features of our failings.

Consider, first, the *diversity* of human failings. In some earlier moral and spiritual traditions there was a practice of taxonomising the varieties of human vices and failings. The Christian and Buddhist traditions, with their concepts of 'sin' and 'defilements', are perhaps the most complex, but they are not the only examples. We could distinguish *moral* failings, *epistemic* failings, *aesthetic* failings, *political* failings, and *spiritual* failings. Specific moral traditions might add their own candidates. Confucians recognise various *ritual failings*, for instance, and environmentalists might want to nominate set of ecological vices and failings.

The diversity of our failings has its roots in at least two things: one is the natural diversity of human attitudes, habits and dispositions and cognitive and practical behaviours that can invite critical attention. Another is the conceptual and social resources available for making sense of those various aspects of human conduct. Certain kinds of behaviour could consistently invite anger or frustration even if one lacks the concepts to *name* that behaviour as, say, *arrogant* or *insouciant*. In some cultures, the available concepts and terms might be generally fit for purpose, insofar as most of its people find themselves able to name and understand bad qualities of themselves or others. But not always: there are many gaps in our inherited resources for naming and theorising failings (cf. Kidd 2018, 52). The Christian tradition has focused on vices associated with pride and humility and so we have relatively rich resources for describing those. In contrast we have a relatively poorer vocabulary for failings associated with curiosity.

To see this, we need only consult studies in the history of vice and virtue. What we find are *promiscuous concepts*: vice-concepts that persist over time by altering their forms to better fit changing moral and cultural conditions, a good example being dogmatism (see, e.g., Schep and Paul 2022 and van Dongen and Paul 2017). There are also *transient concepts*: vice-concepts which owed their intelligibility and salience to specific cultural conditions which subsequently ceased to be. To take one example, some of the failings that worried people

during Baroque European culture are no longer intelligible today, except perhaps as issues of etiquette: the social and cultural contexts that gave them identity and salience no longer obtain (Kivisto 2014). Likewise, newer kinds of failings subsequently came into view—inauthenticity, racist prejudice, and environmental unsustainability, say.

Consider two brief examples of failings that owe their intelligibility and salience to specific 'cosmic outlooks':

(A) The 'deadly sins' described in the early Christian vice tradition, such as enviousness and wrath, have obvious bad effects for oneself and others. At a deeper level, however, they involve distortions of our willingness and ability to cultivate an authentic relationship with God. In effect, they jeopardise the soteriological prospects of their bearers (DeYoung 2020, chs. 4–9 and Taylor 2006, chs. 3–5).

(B) The variety of 'taints' and 'defilements' described in the Buddhist *suttas* include moral-epistemic failings such as *rāga* ('greed') and *moha* ('delusion'). These feed forms of 'unwholesome' (*akuśala*) conduct but also undermine our ability to engage in effective meditative practice and to attain 'right view' of our condition as 'suffering' beings entrapped in the *saṃsāric* cycle of *rebirth* and *kamma* (Bodhi 2012, 41–44 and Harvey 2011).

In these cases, a set of failings owe their intelligibility and salience to 'cosmic visions' of the human condition—as beings trapped in a perpetual *saṃsāric* cycle, or as creatures painfully alienated from God's love. One can decouple them from their associated visions, but only at the cost of either distorting abstraction or conceptual banalisation (these tendencies are very robustly criticised by Burley 2016 and McPherson 2020, chs. 4–5).

I emphasised (a) the diversity of our failings and (b) the fact that the intelligibility and salience of at least some failings will be dependent on certain background conditions. But we should be sensitive as well to the idea of *collective failings*. The term 'vices' tends to understood to mean the failings of character of individual agents. Granted, modern vice theorists argue there are (a) vices that can take collective as well as individual forms and (b) vices that are exclusively collective (Byerly and Byerly 2016). We also often attribute vices to things other than individual agents: we talk naturally enough of *dogmatic* committees, *greedy* institutions, and even vicious abstracta, like *cruel* policies or *heartless* practices. Some think such uses are rhetorical, not attributional, and much depends on our views on the ontology of vices. Moreover, not all vice terms apply naturally to collectives: for this reason, I prefer the term *failing*. A bank is *greedy*, the university is *inefficient*, and industrialised animal agriculture is *cruel*.

Why, though, explore the relationships between individual and collective failings? One reason is that it matters to those interested in the *aetiology* of individual failings: to our efforts to understand the origins or causes of failings of character. After all, we should not conceive of our characters as fixed dispositions which unfold over time in some autonomous way. Our characters are to a degree plastic and can change, or be changed, over time through a variety of indirect and direct influences. Our characters can improve or can deteriorate, and character theorists tend to be interested in finding effective ways to improve our characters and to prevent or repair damage. Understanding the aetiology of failings requires a critical sensitivity to the wider failings of the social world. We must think in terms of the dynamic relations between individual and collective agents and the wider institutions and cultures of which they are a part. Of course, this makes vice theorising much more work, but that is an unavoidable consequence of taking seriously the realities of our moral formation and practice.

One vital concept for understanding the interaction of individual and collective failings is *corruption*.

2. Corruption

The term *corruption* is used in social and political discourse and recently became an object of vigorous scholarly interest (Rothstein and Varraich 2017). I focus on a specific morally-toned sense of the term, inspired by a use of the term 'corruption' popular among vice and virtue theorists. Gabriele Taylor says that 'the vices corrupt and destroy' our good character traits (Taylor 2006, 126). Judith Shklar proposes that vices tend to 'dominate and corrupt' our moral character (Shklar 1984, 200). They define corruption as something that affects the character of individual agents for the worse. Other philosophers speak of the corruption of institutions. Alasdair MacIntyre warns that 'the corruption of institutions is always in part at least an effect of the vices' (McIntyre 2013, 227). Robin Dillon's 'critical character theory' starts from a recognition that 'domination and oppression inflict moral damage on the characters of those who live within them' by subjecting people to 'social forces that work to diminish or corrupt our selves and lives' (Dillon 2012, 85, 92).

Clearly there are different senses of *corruption* and the following account is not intended to be exhaustive or definitive. At best it is one *kind* of corruption, one strongly connected to vices and failings that has become popular in social epistemology (cf. Kidd 2022). Miranda Fricker, for one, argues that subjection to sexist and racist conditions tends to 'inhibit' or 'thwart' our cultivation of ethical and epistemic virtues and in that sense 'corrupts' (Fricker 2007, 92, 131, 138). José Medina defines epistemic vices as 'corrupted attitudes and dispositions'

which must be understood aetiologically in terms of their 'socio-genesis' (Medina 2012, 29, 72). Such corrupting and character-deforming effects are more generally described with a variety of metaphors—'toxic', 'polluted', 'poisonous'—which connote things which are damaging if one is exposed to them for too long (cf. Tyrell 2017). The Buddha used similar rhetorics of corruption: the mainstream social world is 'burning' with the unquenchable 'fires' of delusion, hatred or aversion, and greed (SN 35.28)

What these instances of the term *corruption* capture is a destructive process whereby exposure to certain kinds of processes or conditions tends to cause what Claudia Card calls 'moral damage' (Card 1996). Our character can be morally damaged in two related ways: our existing virtues and excellences can be eroded or extripated (call this *passive corruption*) or vices and failings can be introduced or strengthened (call this *active corruption*). For Lisa Tessman, exposure to 'the ordinary vices of domination', such as cruelty and arrogance, can causes our characters to become 'degraded', 'twisted' (Tessman 2005, 53). Moreover, the self under oppression is 'morally damaged, prevented from developing or exercising some of the virtues' (Tessman 2005, 4).

It is this sense of corruption I want to develop, before I go onto its connection to misanthropy. Here is a general definition:

> Something is *corrupting* if exposure or subjection to it tends to weaken or erode excellences or virtues (*passive form*) and/or facilitate the development and exercise of failings or vices (*active form*).

Corruptors could be social conditions, processes, actions, norms or values, experiences or interactions. An environment can be corrupting due to its pressures, temptations, values, incentives, or its 'atmosphere'. Certain individuals can be corruptors, too, as can internal features of moral agents, such as our personal moral weaknesses and anxieties. Whatever their specific features, exposure to and interaction with corruptors will tend to damage our moral and epistemic character. Some people, of course, seem to be able to resist or mitigate those corrupting influences and there also seem to be subjective degrees of susceptibility to corruption (think of people of great moral integrity or those capable of moral self-control).

Corruption is therefore a *dynamic* phenomenon that unfolds over time that will often involves prolonged, painful moral and psychological struggle. Unless one is very unlucky, one's social world will contain at least some positive influences, such as the presence of inspirational moral heroes, say, and opportunities for corrective self-reflection. Corruption is also a sustained *process*. Some people may be deeply corrupted by a single catastrophic event, such as cases of 'disorientation' where we lose our moral bearings (Harbin 2016). In

most cases, though, we are forced to do the constant effort of working hard to protect our moral character and integrity despite the constant corrupting forces working on us during our everyday lives.

The concept of corruption can help us to understand how individual failings can relate to the wider failings of the social world. A social world can be filled with all kinds of corruptors facilitating different failings in different ways, not to mention the self-corruptive effects of our own vices (Taylor 2006, ch. 7). It is certainly common to describe the social world as corrupting, even if we also judge that its corrupting powers vary in scope, strength, and intensity. It's also very common for moral criticism to use narratives of decline from an actual or imagined earlier period of moral excellence. Confucius experienced his world as undergoing a painful 'change of condition' from the brilliance of earlier dynasties to the newer period of violence, instability, and loss of moral direction into which he was born (Ing 2012 and Olberding 2013).

I said subjection to corrupting conditions 'facilitate' the development and exercise of failings and vices. To cash this out we can distinguish several *modes* of corruption—general ways that conditions could damage our characters:

- *Acquisition*: a corruptor can facilitate the acquisition of new failings, not previously a feature of one's character.

- *Activation*: a corruptor activates dormant failings, which were already present, but latent or inactive.

The next three modes are different in kind: they involve alterations to failings already present or active in one's character:

- *Propagation*: a corruptor can increase the *scope* of some failing, the extent to which it affects the whole range of the subject's outlook or behaviour. A failing *propagates* when it starts to 'infect [our] whole character' (Baier, 1995, 274).

- *Stabilisation*: a corruptor can increase the *stability* of a failings, the extent to which it can resist efforts to control or disrupt it. An unstable failing can flicker 'on and off' but a stabilised vice is like a constantly blaring light.

- *Intensification*: a corruptor can increase the *strength* of a failings—if, for instance, our once-weak form of arrogance intensifies into raging megalomania.

There five modes of corruption are not exhaustive. There may be other modes, each with their own sub-modes, all of which can interrelate in various ways. We can develop this further by noting general kinds of *corrupting conditions*—features of a social environment or culture that tend, in their own ways, to facilitate our failings or to erode our excellences:

- *The absence of exemplars of virtue*—of persons able to model the virtues, offer practical guidance and, perhaps, theoretical insight (cf. Croce and Vaccarezza 2017).

- *The derogation of exemplars of virtue*—virtuous exemplars mqy be subject to scorn, ridicule, violence, or a pervasive cynicism that erodes the *very idea* that a person could *be* good in those ways (cf. Zagzebski 2017, 45).

- *The valorisation of vicious conduct and exemplars*—viciousness, whether in the form of acts or persons, can be praised, promoted, rewarded, a route to status or glory, and so on (think of those philosophers who valorise aggressive, 'take-no-prisoners' styles of debate—cf. Rooney 2010).

- *The rebranding of vices as virtues*—the status of certain attitudes and dispositions *as* vices could be disguised by presenting them as virtues, thereby thwarting our moral self-monitoring (cf. Dillon 2012, 99).

- *Increasing the exercise costs of virtue*—exercises of virtues often require expenditure of energy, courage, or willpower that places demands on the agent. A social environment could be arranged so that virtuous actions 'cost' more—cf. Cooper 2008).

- *Increasing the rewards of viciousness*—a culture can encourage reward viciousness by ensuring it is a reliable route to desirable goods (ego-reinforcement, wealth, power, sexual gratification, and so on).

These are only some of the generic corruptors that could be a part of a social environment. Doubtless, others could be described. Social corruptors can be norms, operating ideals, guidelines, practices of praise, incentive systems, and really anything that creates pressures, temptations, and incentives that affect our moral conduct and development. Studying these corruptors will require conceptual and empirical work and should be part of a philosophical character theory.

I hope this account of corruption is detailed enough to persuade you that it tracks a genuine phenomenon. Awareness of the variety of corruptors built into our world is central to our study of the aetiology of human failings. Moreover, we see concerns about corrupting social environments in historical and contemporary character ethicists from Confucius and the Buddha through to present-day virtue theorists and social epistemologists.

I now show how this account of corruption brings us to *misanthropy*.

3. Misanthropy

According to what we can call the standard account, misanthropy is the hatred or dislike of human beings or humankind (Gerber 2002). Misanthropy is defined in terms of one or more negative affects, which often extend to contempt and disgust, and so misanthropes are often referred to as 'haters of humankind'. Of course, everyone agrees that there are affective and emotional components to misanthropy, even if others emphasise other aspects as being its ore. Toby Svoboda, for one, distinguishes affective kinds of misanthropy, such as *disliking* humankind, from what he calls 'cognitive misanthropy' (Svoboda 2022). However, what is historically dominant are characterisations of misanthropy in terms of negative affects with hatred being top of the list.

I think attempts to partition misanthropy into affective and cognitive forms and then nominate one of them as 'central' or 'core' are uncompelling. In practice our emotions, feelings, and moods are in constant intimate relations with our evaluations, thoughts, and judgements. The earliest Greek discussions of misanthropy recognised that it has affective as well as cognitive dimensions, and others too, including what Socrates called a loss of trust in humankind. Kant and Schopenhauer, too, characterise misanthropy in terms of experiences, reflections, emotions, and moods. Of course, it can be analytically useful to distinguish the 'affective' and 'cognitive' aspects of misanthropy, but on the understanding that this is not taken as a sign of any genuine partitioning (cf. Cooper 2018, 3ff). Moreover, a focus on the cognitive and affective aspects should not obscure the behavioural or practical dimensions of misanthropy. In practice, misanthropy is never simply a set of negative affects *or* certain cognitive states *or* some combination of these. It also manifests in actions, behaviours, kinds of comportment towards others and the world, even a certain misanthropic way of life. No single aspect should be elevated over the others. All exist in a complex and dynamic interplay. Emotional experiences of anger, bitterness, and sadness can provoke changes in our attentional and reflective habits that in turn issue in evaluations and understanding which in term shape our emotional profile and interpersonal habits.

This more complex characterisation of misanthropy has been developed in the recent work of David E. Cooper, which I have endorsed and elaborated elsewhere (Cooper 2018, Kidd 2021). To distinguish it from the standard account, let me call Cooper's own position a *revisionary* account. It is not a perfect term. For one thing, it begs the question in favour of the primacy of the standard account, of misanthropy-as-hatred, which is mainly defended by Lisa Gerber (Gerber 2021, §2). My judgment is that the revisionary account actually better conforms to the actual positions of the majority of philosophical misanthropes across history (cf. Kidd 2021, §§3–5). Historical misanthropes did not always fixate on hatred and many of them in fact rejected hatred, meaning that hatred is neither necessary nor sufficient for misanthropy. For present purposes, though, the term *revisionary misanthropy* will suffice.

The central claim of the revisionary account is that misanthropy is dark, negative critical appraisal of, or verdict on, the collective character and performance of humankind as it has come to be (cf. Cooper 2018, ch. 1, Kidd, 2021, §2). A misanthropic verdict could be inspired by a wide range of experiences and reflections and judgments. It can also express itself in a range of moods and feelings—from bitterness to despair to sadness to resignation. Moreover, the target of the verdict is not individual human beings but something much more abstract and collective—humanity, humankind, human forms of life, the human condition. In the Western tradition, it has also been popular to root misanthropy in conceptions of human nature. But appeals to human nature are not a necessary feature of misanthropy in the revisionary sense. Jean-Jacques Rousseau was right that critical appraisal is directed at 'civilized man'—on human life as it has come to be in the complex forms into which have been thrown. We are not also directing judgment to 'natural man', to our life and nature as it was in some earlier state of original innocence (Rousseau 1994, 94–97). Indeed, our original or underlying moral condition—as it was in the deep past or as it might be under the brittle veneer of civilized existence—is actually irrelevant to that appraisal (Kidd 2020). A misanthrope need not say we are fundamentally flawed, only that we are—and continue to be—contingently corrupted by the current conditions of our world. Of course, there are also pessimistic misanthropes who maintain that we were doomed to turn out badly, but that sort of claim is not integral to the revisionary account of misanthropy.

A main virtue of the revisionary account of misanthropy is that it is triply pluralistic. First a critical verdict on humanity can and will almost certainly involve many *affects* which can sometimes include positive affects, since all but the bleakest misanthropes recognise that certain people and experiences can be positive. Second, a misanthropic verdict on the baleful moral condition of humankind can have diverse *content*, meaning there are different sorts of

misanthropic verdict. Cooper, for instance, focuses on our dreadful treatment of non-human animals. Different misanthropes focus on different concepts, concerns, and aspects of human life. Third, a misanthropic verdict can manifest in different behaviours, habits, and ways of living; there are many ways to 'be' a misanthrope and to enact a misanthropic vision of the world. In effect, then, revisionary misanthropy endorses *misanthropic pluralism*: there are many *misanthropic stances* (cf. Kidd 2021, §§ 3–5). For all their differences, these stances all express a dark, critical appraisal of humankind as it has come to be.

All of these points were realised by Immanuel Kant in his remarks on misanthropy in his lectures on religion, ethics, and anthropology. What initiates misanthropy, for Kant, is a 'long, sad experience' of the failings consistently manifest in human life, such as ingratitude, selfishness, and cruelty (Kant 1997, 27: 671–672). Such experiences provoke processes of reflection and deliberation which can begin to change our feelings and moods. In a similar account, Arthur Schopenhauer argued that misanthropic feelings and thoughts can lead to a 'melancholy mood'—unless something intervenes to block it—and warned that, if this mood 'persists, then misanthropy arises' (Schopenhauer 2010, 205). Emotions, feelings, thoughts, and preoccupations are all integral to the psychology of misanthropy. Kathryn Norlock also adds the useful point that such misanthropic affects and thoughts must not simply be *there* in one's mind. Many people have misanthropic thoughts and sentiments, but resist and dismiss them or find them overmastered by other convictions and attitudes. To be a misanthrope, one must have these thoughts, moods, and feelings *and* have them become central to—definitive of—one's overall outlook and stance on life (Norlock 2021a, 53ff). Norlock emphasises that

> a critical negative judgment doesn't get one all the way to misanthropy. For a person [. . .] to be misanthropic [. . .] the critical negative judgment has to rise to a governing principle, a justification for further views and practices. (Norlock 2021b, 15)

Cooper, Norlock, and other advocates of the revisionary account agree (a) misanthropy should not be reduced to its cognitive or affective components, that (b) there are complex processes of conversion or self-transformation, and there are (c) many different misanthropic stances, understood as a fairly systematic way of living out a critically negative moral vision of humankind. Kant—a rich theorist of misanthropy—names two misanthropic stances. The 'Enemy of Mankind' feels 'enmity' for humankind, a combination of 'dislike' and 'ill-will', which manifests in dispositions to violence. The most extreme Enemy misanthrope comes to be dominated by 'the purpose and will to destroy the welfare of others' (Kant 1997, 27: 672). The 'Fugitive from Mankind', by contrast,

is animated by profound fear of the moral and physical risks of remaining within the human world. The Fugitive misanthrope comes to 'apprehend harm from everyone' and responds by escaping or fleeing the mainstream social world—to some refuge that offers a prospect of moral and physical safety and a space which can accommodate a sustainably kind of life (Kant 1997, 27: 672).

Kant's account of misanthropic stances is attractive, but requires some amendments. First, the tight pairing of affects—like enmity-violence and fear-flight—is much too tidy and obviously false. If I hate something, I might want to be violence to it or stay well away from it. If I fear someone, I might want to flee from them or do them violence. The connections of affects and practical behaviours needs to be much more complex. It is better to distinguish the misanthropy stances *practically* rather than affectively. Second, Kant defines stances by a single negative affect, like enmity or fear, in a way that obscures the complex emotional dynamics of misanthropy. A moral evaluation of humankind is far too big a thing to ever find its cause or characterisation in a single affect, even in a powerful one like hatred or fear. Kant obscures the emotional and psychological complexities of misanthropy; what we really need is a way of accommodating the interplay of anger, disappointment, fear, hatred, hope, sadness, and other emotions, feelings, and moods reported by misanthropes. Granted, some of them tend to gravitate towards certain emotions while others may not experience certain kinds of affect at all. But all this testifies to the complexity and diversity of the misanthropic stances.

A third amendment to Kant's account is the addition of other stances beyond Enemy and Fugitive types. Granted, he did not claim to be comprehensive and he actually briefly mentions another stance, perhaps a variant of Fugitivism, that has been labelled the 'virtuous solitary' (see Trullinger 2015). But the history of philosophy, eastern and western, offers at least two other general kinds of misanthropic stance. *Activists* attempt large-scale action that aims at a transformation of the collective character of humankind. Their radical projects may include moral teaching, religious preaching, social activism or technological enhancement of human beings. Activist misanthropes aspire to radical rectification of the human condition. By contrast, *Quietists* see our collective failings are incorrigible and incapable of any serious rectification (Cooper 2018, 118ff). Quietists focus on accommodating to the failings of their world and at avoiding or managing its corrupting potentialities. Quietist misanthropes accept that certain desirable human goods are only available within the human world, like family or fellowship, and so remain carefully engaged with the world while striving to avoid morally compromising entanglements (on these four stances, see Kidd 2021, §§3–5).

How does misanthropy, so defined, relate to failings and corruption? Simply stated, a misanthrope has come to experience the human world as vicious and

corrupting, as suffused with a variety of failings. Which failings appear salient will depend on the particular values, concerns, and moral commitments of the misanthrope in question. Cooper's doctrine of misanthropy, for instance, focuses on clusters of failings which manifest in our dreadful collective treatment of animals and are illustrated with a depressingly detailed 'charge list' (cf. Cooper 2018, chs. 4 and 6). Other misanthropes recognise different clusters of failings, such as the Christian and Buddhist catalogues of our failings. Others are less systematic, too. Schopenhauer lists "vices, failings, weaknesses [. . .] and imperfections of all sorts", and notes that some more common than the others, like 'the boundless egoism of almost everyone, the malice of most, the cruelty of many' (Schopenhauer 2010, 200, 205).

For a misanthrope, such failings must be *ubiquitous* and *entrenched*, meaning they are spread widely throughout the world and deeply built into its structures (Cooper 2018, 54ff). As well as reflecting a depressing fact about the world, these two features play an important strategic role in the misanthrope's argument. A critic may accept the reality of our failings but insist they are confined to (i) extreme individuals or groups—psychopaths, say—or (ii) extreme conditions, like poverty or social turmoil, which give otherwise unusual power to our selfishness and violence. Such *confinement strategies* can be resisted by emphasising the ubiquity and entrenchment of our failings. Even if we are *worse* under such conditions, our failings are in fact all too 'distinctive of—typical of and integral to' our human forms of life (Cooper 2018, 63). The abuse and exploitation of non-human animals, for instance, and the unsustainable destruction of the natural world are not rare, occasional features of modern forms of life: they are utterly integral to its practices, projects, and normal functioning.

A misanthrope experiences the human world as shot through with failings and also as deeply corrupting. This explains common misanthropic rhetoric—in talk of human existence as 'poisonous', 'toxic', or 'rotten' or in talk of humanity as a 'plague' or 'cancer'. This sense of the world as vicious and corrupting is clear in Schopenhauer's account of his deeply misanthropic vision of the human world:

> [W]e see come to the fore insatiable greed, vile greed for money, deeply disguised duplicity, insidious malice of humans, we often recoil in horror and let loose an outcry [Human beings are like] so many tigers and wolves whose jaws are powerfully muzzled. (Schopenhauer 2010, 200)

Being corrupted is one of the moral harms that concerns the Fugitive misanthrope. Some misanthropes therefore describe their actual or hoped-for refuges as places where they can shelter from the moral hazards of mainstream

life (Cooper 2021). Many eco-misanthropes, for instance, often describe nature—'wild' nature, at least—in moral as well as aesthetic terms as 'perfectly clean and pure' (Muir 2007, 114). Chinese misanthropes historically use the metaphor of the world as 'dusty' and 'grimy' and described their refuges as 'clean' and 'pure'. In the Buddhist tradition, the monastic community, the *saṅgha*, is one of the three 'refuges'—alongside the Buddha and the Dhamma, his teachings—and monastic life is characterised by physical and moral cleanliness and 'wholesomeness' (Harvey 2000, ch.8). Central to a misanthropic vision is a palpable sense of human life as corrupting in ways that demand practical response—whether the violence of the Enemy of the flight of the Fugitive or Activist radical reformism or Quietist strategies of accommodation.

I hope this brief account makes clear the connection between failings, corruption, and the critical appraisal of the moral condition of humanity at the core of misanthropy. I think a misanthrope—of the revisionary sort at least—thinks we only fully appreciate the vices and failings of our corrupting world when they are connected to a more systematic critical vision of human life as it has come to be. For Giacomo Leopardi, the Italian poet, what we come to realise is the truth that 'the world is the enemy of the good' (Leopardi 2002, 85). This is not a welcome truth and need not be taken as an inevitable fact of human history. But it is a truth nonetheless.

I now turn—in a more tentative tone—to the question of whether there could be forms of Christian misanthropy.

4. Sin and Soteriology

I described conceptual connections between failings, corruption and misanthropic critique of human life as it has come to be. With the revisionary account of misanthropy in place, we can ask if there could be specifically Christian forms of misanthropy. A good place to start, when considering that question, are doctrines of sin. However, this approach instantly runs into two problems. First, the historical and doctrinal variety of hamartiologies means that 'any talk of *the* Christian view of sin is questionable from the start' (Moser 2010, 136–137, my italics). Second, conceptions of sin are situated in a wider structure of concepts—grace, redemption, and salvation—which are, in turn, integrated into wider doctrinal and narrative structures. A postlapsarian conception of human beings as 'fallen' beings deeply corrupted by original sin is one example (cf. Hart 1997). Moreover, our fallenness can be articulated in different ways—as, for instance, a relatively minimal disturbance to our spiritual abilities or as a catastrophic destruction of our very essence. Such complexities set strict limits on what we could say in general about Christian conceptions of sin as a route into a form of Christian misanthropy.

A difficult sort of obstacle is the relative absence of the concept of misanthropy from Christian discourses. It is rarely invoked by Christian theologians and, where it is used, the aim is generally to reject it. Søren Kierkegaard doubtless spoke for many when he judged that misanthropy is 'far removed from Christianity', which is a 'gentle teaching' founded on love. He defines misanthropy in terms of the standard account as *odium totis generis humani*—hatred of the human race—and thinks misanthropes inevitably become 'mute and sallow hermits', alienated from themselves, Christian fellowship, and from God (Kierkegaard 2015, 118–119ff).

Kierkegaard seems to understand misanthropy as an interesting fusion of Enemy affects and Fugitive behaviours. This is likely due to the influence of Rousseau and Kant (Cassirer 1945). Kierkegaard's understanding of misanthrope obviously has roots in his own Christian sensibilities as well, as do other aspects of his thought—his concept of *dread* was worked out in reference to a postlapsarian vision of the human condition (Mulhall 2005, 49ff). But it is a mistake to suppose misanthropy *must* include hateful affects and reclusive behaviours of the sort that alarmed Kierkegaard. In effect, what he has in mind are specific forms which misanthropy could take, and he fails to recognise the diversity of other possible forms.

The revisionary account of misanthropy as a critical verdict on the moral condition of humankind surely resonates with themes within the Christian tradition. If misanthropy calls our critical attention to something substantially *wrong* with the human condition, it naturally connects to the soteriological aspirations that have been central to Christian practice:

> The notion of redemption or salvation is a basic constituent in the plot of the story which Christian faith tells about human existence in God's world. The characteristic designation of this story as 'gospel', good news, already bears within it the assumption of a human race in some serious need or lack or crisis, whether it is aware of it or not. (Hart 1997, 189)

A sense of the human condition as being problematic in deep, severe, and systematic ways can and has been articulated in various ways. About the 4th century, the Fall of Man came to be understood as *depravatio*, as perverse corruption, rather than, as before, *deprivatio*, the loss of something good (Hick 1985, 213). A sense of the loss of personal and collective goodness can point towards a misanthropic vision, but emphasising the deep corruption of individuals and their world take one several steps closer to a misanthropic verdict. This slow movement towards that verdict can be encouraged by other doctrines and attitudes. The classical theme of *contemptus mundi* led to Renaissance Christian moralists 'wallowing in vivid depictions of the degraded

state of human beings here on earth' (Frede 2013, 131). But misanthropy need not involve 'wallowing' in our collective crapulence or trading in 'vivid depictions' of our degradation. Later Christian writers offer more sombre accounts of our condition. The Reformed theologian Reinhold Niebuhr saw evil and sin as 'a corruption which has a universal dominion over all men' (Niebuhr 1949, 122). 'Universal dominion' conveys the sense of the ubiquity and entrenchment of our failings central to misanthropy.

At least some Christian theologians engage in the project of critical appraisal of the human condition. The vocabulary and doctrinal content of those appraisals vary—we can be depraved, 'fallen', corrupted by original sin, standing in need of redemption or sanctifying grace, alienated from God—to name but a few. What we find in these Christian resources are rich possibilities for a critical appraisal of our collective condition:

> [S]ome human communities find it easier to identify with a particular element of the human plight as described by scripture—guilt, alienation, impurity, mortality, ignorance, oppression or whatever—than others, and therefore find it easier to own [certain] metaphor[s] of salvation—acquittal, forgiveness, sanctification, bestowal of new life, illumination, liberation, etc. (Hart 1997, 190)

I think the possibility of authentically Christian forms of misanthropy should be explored. It is a way of enriching our thinking about misanthropy. It might also reveal Christian themes and concerns in a new light. Alongside some of the great Christian theologians, other good candidates might be Tolstoy, Evelyn Waugh, and T.S. Eliot. We should not prejudge the results of these explorations. Some kinds of misanthropy may be ruled out by arrangements of Christian spiritual teachings. It's hard to imagine anyone who lives by the teachings of Jesus ratifying hateful violence against the human world. When Kant judged the Enemy stance to be 'contemptible', he was invoking his Christian convictions about the importance of cultivating 'love for others' and 'for the entire human race' (Kant 1997, 27: 673).

To encourage explorations of Christian possibilities, it is useful to reiterate the revised account of misanthropy, inspired by Cooper, which I endorse:

> Misanthropy is a critical appraisal or verdict on the collective character and performance of humankind or human life as it has come to be.
> Human life is characterised by a variety of *failings* that are both *ubiquitous* and *entrenched*—and not in practice confined to extreme people or conditions.
> A misanthrope therefore experiences humanity or human forms of life as both vicious and *corrupting*.

The account is flexible in several respects. It does not build into misanthropy any specific set of affects or behaviours. It does not privilege a particular stance as the authentic expression of a misanthropic verdict. It does not stipulate how we select, define, or order our failings. It is neutral with respect to aetiological explanations about our condition. It does not require any account of human nature. It could also be rooted in different metaphysical worldviews; scientific naturalists can be misanthropes as well as theists or ineffabilists.

A rare example of a self-identified Christian misanthrope was the English writer and reformer, Sir Perceival Stockdale, author of a 1783 *Essay on Misanthropy*. Stockdale rejects hateful misanthropy but insisted there was a different, defensible sort which is acceptable to those, like himself, of Christians sentiment:

> [T]here is a Misanthrope, who is as acute, and severe in his observations, as he is gentle, and placid in his conduct. He cannot but be convinced, that the great majority of mankind are under the fatal dominion of vice [. . .] While the history of the human race, and his own acute observations, are continually confirming his Misanthropy, are convincing him afresh, that mankind, in the aggregate, are extremely wicked; the same extensive, and complete view of human agents, and of the objects that surround them, equally inspire him with an amiable toleration, and indulgence towards the species. (Stockdale 1783, 9, 12)

Stockdale endorses something like the revisionary account of misanthropy. He also thinks it can and should reconciled with the virtues and attitudes required of a Christian.

A good starting point for those who want to explore misanthropy and Christianity are the studies of the Christian vices tradition by historian and theologian Rebecca DeYoung. In her book *Glittering Vices*, she defined vices as 'disordered desires' for worldly goods such as 'pleasures, security, comfort, control, wealth, status, approval, success, reputation' (DeYoung 2020, 219). Our disordered desires are the substratum of our vices, the basis for 'corrupting and destructive habits' which, if not checked, lead us onto 'paths of self-damage and self-destruction' (DeYoung 2020, 8, 197). Such disordered desires reflect deep general features of human nature—our sensuality, say—but are also shaped and animated by social, economic, and cultural values and imperatives. Consumerist or hedonistic cultures make us 'default to the deformities of wrath' and greed and other vices, for instance, with the upshot that serious vice theorising must be sensitive to 'the dynamics of sin and the deep network of its combined forces' in those forms of life we have inherited (DeYoung 2020, 149, 239).

Misanthropes without Christian convictions can recognise and accept all these details. We have failings that are entrenched and ubiquitous features of the human condition as we have inherited it—all the core elements of

misanthropy. To this general account, DeYoung adds specifically Christian themes. What unifies our vices is the inveterate human drive to 'pursu[e] finite, created things in place of the goodness of God' (DeYoung 2020, 37). Our diverse vices all reflect, in their own ways, this fundamental disordering of desires. Such inner disorder becomes inscribed into our habits, outlooks and relationships with others. Our vices in turn corrupt other people, becoming concretised in social practices and institutions. In a worst-case scenario, a whole form of life becomes animated by corrupted values, like a 'narrowly self-serving, flesh-aimed vision of the good' that feeds greediness, lasciviousness, and other all-too-human failings (DeYoung 2020, 210).

DeYoung does not use the term *misanthropy* and I suspect she would not endorse it as a description of her position. If humans are made in the image of God (*imago Dei*), that might seem to rule out any misanthropic verdict on humankind. But matters are more complicated. Since the *imago Dei* doctrine has different forms and interpretations, there is no automatic movement from it to a rejection of misanthropy. If it means we are incapable of significant moral failure, it is clearly wrong. If it means we are *capable* of moral or spiritual excellence, it is consistent with all but the most extremely pessimistic forms of misanthropy.

We can use DeYoung's discussion of the Christian vices tradition as a starting point for exploring kinds of Christian misanthropy. Her book analyses failings, articulated in terms of vices and sins, and describes ways we can cluster them. She also describes corrupting forces that a misanthrope also sees at work in the world (cf. Daly 2021). I read her as offering a critical verdict on a world animated by 'disordered desires' whose entrenched structures and pressures make us 'default' to 'deformities' of thought, feeling, and action, ones we cannot fully resist without divine support. Other Christian theologies agree with this account of our being sinful sins in a sinful world. For one contemporary theologian, sin is 'the broad view that human beings are born into a condition of fundamentally disordered willing from which they cannot extricate themselves by their own powers' (Zahl 2020, 158).

A misanthrope without specifically Christian commitments can agree with all this, and be inspired by DeYoung's proposals for coping with a world experienced as systematically morally disordered. An example is her historical emphasis on those intellectual or cultural tendencies which 'radically marginalised the vices' and her accounts of 'graced' penitential and confessional practices and affirmations of Christian spiritual and vocational ideals that 'teach us both resistance to sin and receptivity to Spirit' (DeYoung 2020, 29, 92, 221).

Conclusions

Careful historical and philosophical engagement with Christian doctrines and practices offers rich work for those interested in misanthropy. I have only a briefly sketched some of the possibilities. I hope to have allayed Kierkegaard's and Kant's worries about intractable incompatibilities between Christian faith and misanthropic visions. Misanthropy can be consistent with pained love of humankind, injunctions to compassion, and soteriological aspirations. One can imagine very dark, pessimistic forms of misanthropy that *are* hostile to those claims, but those would be very specific and extreme forms of misanthropy.

In practice we should investigate further ways of relating misanthropy to the beliefs, doctrines, ideals, outlooks, and soteriologies in Christian and other religious traditions. Such comparative work was initiated by Schopenhauer, who commended as 'wise' the 'New Testament and the Indian traditions, 'Brahmanism' and Buddhism, for their appreciation of the 'misery', 'wretchedness', and 'obvious [moral] imperfection' of human beings (Schopenhauer 1974, 301–3033ff). Engagements with these religions and theologies will also correct an irksome tendency in philosophical theorising on misanthropy to impugn the value of theology. In her book on vices, which includes a chapter on misanthropy, Judith Shklar briskly dismisses the value of theology:

> One might suggest that the works of theologians could prove useful, but their range is somewhat limited. Offences against the divine order—sin, to be exact—must be their chief concern [. . .] It is only if we step outside the divinely ruled moral universe than we can really put our minds to the common ills we inflict upon one another every day. (Shklar 1984, 1)

Shklar's denial is unpersuasive. She does not detail her understanding of sin or its relations to wider Christian thought. 'Sin' is a rich concept with mutable meanings and real purchase on the everyday business of human life. She does not defend her judgement that 'common ills' can *only* be explored by setting aside theistic perspectives. Nor does she explore in any detail the resources offered by theistic religions and their theological traditions. Of course, theistic and Christian conceptions of the human condition are not compelling for everyone. But that is no reason to exclude them from any serious moral reckoning with humankind of the sort attempted by a misanthrope.[1]

[1] I am very grateful to the Editors for their invitation to contribute and their patience and to David E. Cooper, David McPherson, Kate Norlock, Peter Watts, and two anonymous referees for helpful comments and suggestions. Kate also kindly shared with me her draft paper 'Can a

Bibliography

Baier, Anette. 1995. *Moral Prejudices: Essays on Ethics*. Cambridge, MA: Harvard University Press.

Burley, Mikel. 2016. *Rebirth and the Stream of Life: A Philosophical Study of Reincarnation, Karma and Ethics*. London: Bloomsbury.

Byerly, T. Ryan and Meghan Byerly (2016) 'Collective Virtue', *Journal of Value Inquiry* 50 (1): 33–50. https://doi.org/10.1007/s10790-015-9484-y.

Card, Claudia. 1996. *The Unnatural Lottery: Character and Moral Luck*. Philadelphia: Temple University Press.

Cassirer, Ernst. 1945. *Rousseau, Kant, Goethe*, trans. James Gutmann, Paul Oskar Kristeller, and John Herman Randall, Jr. Princeton, Princeton University Press. https://doi.org/10.1515/9781400867677.

Cooper, David E. 2008 "Teaching and Truthfulness," *Studies in Philosophy and Education* 27 (2–3): 79–87. https://doi.org/10.1007/s11217-007-9091-x.

Cooper, David E. 2018. *Animals and Misanthropy*. London: Routledge.

Cooper, David E. 2021. "The Rhetoric of Refuge," *Daily Philosophy*, 5th July, https://daily-philosophy.com/cooper-rhetoric-of-refuge/.

Croce, Michel and Maria Silvia Vaccarezza. 2017. "Educating through Exemplars: Alternative Paths to Virtue," *Theory and Research in Education* 15 (1): 5–19. https://doi.org/10.1177/1477878517695903.

Daly, Daniel J. 2021. *The Structures of Virtue and Vice*. Washington, DC: Georgetown University Press. https://doi.org/10.2307/j.ctv1gm00rq.

DeYoung, Rebecca. 2020. *Glittering Vices: A New Look at the Seven Deadly Sins and Their Remedies*, 2nd edition. Grand Rapids, MI: Brazos Press.

Dillon, Robin. 2012. "Critical Character Theory: Toward a Feminist Perspective on 'Vice' (and 'Virtue'), *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*, edited by Sharon L. Crasnow and Anita M. Superson 83–114. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199855469.003.0005.

Dongen Jeroen van and Herman Paul. 2017. *Epistemic Virtues in the Sciences and the Humanities*. Cham: Springer.

Frede, Dorothea. 2013. "The Historic Decline of Virtue Ethics," *The Cambridge Companion to Virtue Ethics*, edited by Daniel C. Russell, 124–148. Cambridge: Cambridge University Press. https://doi.org/10.1017/CCO9780511734786.007.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001.

Religion be Misanthropic (and Why Would That Matter)?' Many of the ideas in this paper were also worked out in conversation with the students in my misanthropy lectures in 2021 and 2022.

Gerber, Lisa "2002. "What's is So Bad about Misanthropy?" *Environmental Ethics* 24 (1): 41–55. https://doi.org/10.5840/enviroethics200224140.

Gerber, Lisa 2021. "A Word Against Misanthropy," *Journal of Philosophical Research* 46: 73–87. https://doi.org/10.5840/jpr20211013180.

Harbin, Ami. 2016. *Disorientation and Moral Life*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190277390.001.0001.

Hart, Trevor. 1997. "Redemption and Fall," *The Cambridge Companion to Christian Doctrine*, edited by Colin E. Gunton, 189–206. Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL0521471184.012.

Harvey, Peter. 2000. *An Introduction to Buddhist Ethics: Foundations, Values, and Issues.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511800801.

Hick, John. 1985. *Evil and the God of Love.* London: Palgrave. https://doi.org/10.1007/978-1-349-18048-6.

Ing, Michael David Kaulana. 2012. *The Dysfunction of Ritual in Early Confucianism*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199924899.001.0001.

Kant, Immanuel. 1997. *Lectures on Ethics*, edited by Peter Heath and J.B. Schneewind, translated by Peter Heath. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107049512.

Kidd, Ian James. 2018. "Deep Epistemic Vices," *Journal of Philosophical Research* 43: 43–67. https://doi.org/10.5840/jpr2018431.

Kidd, Ian James. 2020. "Humankind, Human Nature, and Misanthropy," *Metascience* 29: 505–508. https://doi.org/10.1007/s11016-020-00562-8.

Kidd, Ian James. 2021. "Varieties of Philosophical Misanthropy," *Journal of Philosophical Research* 46: 27–44. https://doi.org/10.5840/jpr2021819170.

Kidd, Ian James. 2022. "From Vice Epistemology to Critical Character Epistemology," *Social Virtue Epistemology*, edited by Mark Alfano, Colin Klein, Jeroen de Ridder, 84–102. New York: Routledge. https://doi.org/10.4324/9780367808952-11.

Kierkegaard, Soren 2015. *Training in Christianity*, translated by Walter Lowrie. Princeton: Princeton University Press.

Kivisto, Sari. 2014. *The Vices of Learning: Morality and Knowledge in Early Modern Universities.* Leiden: Brill. https://doi.org/10.1163/9789004276451.

Leopardi, Giacomo. 2002. *Thoughts*. London: Hesperus.

MacIntyre, Alastair. 2013. *After Virtue: A Study in Moral Theory*, 3rd ed. London: Bloomsbury.

McPherson, David. 2020. *Virtue and Meaning: A Neo-Aristotelian Perspective.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108775151.

Medina, José. 2013. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199929023.001.0001.

Moser, Paul K. 2010. "Sin and Salvation," in *The Cambridge Companion to Christian Philosophical Theology*, edited by Charles Talliaferro and Chad Meister, 136–151. Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL9780521514330.010.

Muir, John. 2007. *My First Summer in the Sierra.* Edinburgh: Canongate Books.

Mulhall, Stephen. 2005. *Philosophical Myths of the Fall*. Princeton and Oxford: Princeton University Press.

Niebuhr, Reinhold. 1949. *Faith and History.* New York: Scribner.

Norlock, Kathryn. 2019. "Perpetual Struggle," *Hypatia* 34.1: 6–19. https://doi.org/10.1111/hypa.12452.

Norlock, Kathryn. 2021a. "Misanthropy and Misanthropes," *Journal of Philosophical Research* 46: 45–58. https://doi.org/10.5840/jpr2021826173.

Norlock, Kathryn. 2021b. "Can a Religion be Misanthropic (and Why Would That Matter)?," paper delivered at the University of Notre Dame Center for Philosophy of Religion, 17 September.

Olberding, Amy. 2013. "Confucius' Complaints and the *Analects'* Account of the Good Life', *Dao* 12: 417–440. https://doi.org/10.1007/s11712-013-9343-0.

Rooney, Phyllis. 2010. "Philosophy, Adversarial Argumentation, and Embattled Reason," *Informal Logic* 30 (3): 203–234. https://doi.org/10.22329/il.v30i3.3032.

Rothstein, Bo and Aiysha Varraich. 2017. *Making Sense of Corruption.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316681596.

Rousseau, Jean-Jacques. 1994. *Discourse on the Origin of Inequality*. Oxford: Oxford University Press.

Schep, Caroline and Herman Paul. 2022. "Denial of Coevalness: Charges of Dogmatism in the Nineteenth-Century Humanities," *History of European Ideas* 48.6: 778–794. https://doi.org/10.1080/01916599.2021.2009365.

Schopenhauer, Arthur. 1974. *Parerga and Paralipomena: Short Philosophical Essays*, trans. E.J.F. Payne. Oxford: Oxford University Press. https://doi.org/10.1093/actrade/9780199242214.book.1.

Schopenhauer, Arthur. 2010. *Two Fundamental Problems of Ethics*, translated by David E. Cartwright and Edward E. Eerdman. Oxford: Oxford University Press.

Shklar, Judith. 1984. *Ordinary Vices.* Harvard: Harvard University Press.

Stockdale, Percival. 1783. *An Essay on Misanthropy.* Berwick: Printed for W. Phorson.

Svoboda, Toby. 2022. *A Philosophical Defense of Misanthropy.* London: Taylor and Francis). https://doi.org/10.4324/9781003189398.

Tanesini, Alessandra. 2021. *The Mismeasure of the Self: An Essay in Vice Epistemology.* Oxford: Oxford university Press. https://doi.org/10.1093/oso/9780198858836.001.0001.

Taylor, Gabriele. 2006. *Deadly Vices.* Oxford: Clarendon. https://doi.org/10.1093/oso/9780198858836.001.0001/

Tessman, Lisa. 2005. *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. New York: Oxford University Press. https://doi.org/10.1093/0195179145.001.0001.

Trullinger, Joseph. 2015. "Kant's Neglected Account of the Virtuous Solitary," *International Philosophical Quarterly* 55.1: 67–83. https://doi.org/10.5840/ipq201511527.

Tyrell, Lynn. 2017. "Toxic Speech: Toward an Epidemiology of Discursive Harm," *Philosophical Topics* 45 (2): 139–161. https://doi.org/10.5840/philtopics201745217.

Zagzebski, Linda. 2017. *Exemplarist Moral Theory.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190655846.001.0001.

Zahl, Simeon. 2020. *The Holy Spirit and Christian Experience.* Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198827788.001.0001.

Published Online First: February 28, 2023

2023 TheoLogica
An International Journal for Philosophy of Religion and Philosophical Theology
S. I. SIN AND VICE
DOI: https://doi.org/10.14428/thl.v7i2.66163

How Sinful Is Sin? How Vicious Is Vice? A Modest Defense of the Guise of the Good

CHARLES TALIAFERRO & EMILY KNUTHS
St. Olaf College, Northfield, Minnesota USA
taliafer@stolaf.edu

Abstract: We defend the guise of the good thesis in a tradition going back to Socrates and Plato, according to which persons act on the basis of what appears to them as good or the least bad or evil act available to them. This seems contrary to moral experience, but we defend the thesis against plausible counter-examples in life as well as fiction. We contend that the thesis makes wrong-doing and vice intelligible, but still wrong, dysfunctional and horrific.

Keywords: Socrates, Plato, Bernard Williams, Moral realism, Subjectivism

A political advisor to Donald Trump, Steve Bannon, once said, "Darkness is good. Dick Cheney. Darth Vader. Satan. That's power."[1] We assume readers may be more familiar with Darth Vader (a central villain in the Star Wars films) and Satan (the supernatural, fallen angel opposed to God in Christian theology), than Dick Cheney. He was former Vice President of the United States (2001–2009) feared for his great exercise of power to humiliate opponents, win close elections, and start a major war (he provided the Bush administration with—what turned out to be—a false rationale for invading Iraq in 2003).[2]

In this essay, we take up the classic question about the power of evil. There is a tradition, going back to Socrates, Plato, and Aristotle, that wrong-doers act under what they believe to be the pursuit of some good. Many Medieval philosophers agreed, including Augustine and Aquinas. One standard formulation of the thesis is *nihil appetimus, nisi sub ratione boni; nihil aversamur, nisi ratione mali*; We desire only

[1] Cited from: https://www.cnn.com/2016/11/18/politics/steve-bannon-donald-trump-hollywood-reporter-interview/index.html. Accessed May 4, 2022.

[2]For a plausible argument that Cheney engaged in wrongful deception in building a case for the US to invade Iraq, see *Lying and Deception; Theory and Practice* by Thomas Carson (2010).

what we conceive to be good; we avoid only what we conceive to be bad.[3] This position is often held in the context of moral realism, the view that good and evil (or bad) are not merely (or entirely) a matter of what subjects or their society values. *Tout le monde* might judge slavery to be morally acceptable and they would be wrong. Arguably, according what is called *the guise of the good* thesis (our actions are guided by what we conceive of as good), such a pro-slavery world would still (wrongly) think that slavery is either justified or morally permissible in light of some conception of a greater good.

The guise of the good thesis can be formulated in different ways to avoid obvious counter-examples for, indeed, the thesis seems to conflict with everyday experience and even the New Testament when St. Paul confesses to doing what he hates, sin. The latter is worth citing:

> 15 I do not understand what I do. For what I want to do I do not do, but what I hate I
> do. 16 And if I do what I do not want to do, I agree that the law is good. 17 As it is, it
> is no longer I myself who do it, but it is sin living in me. 18 For I know that good itself
> does not dwell in me, that is, in my sinful nature.[a] For I have the desire to do what
> is good, but I cannot carry it out. 19 For I do not do the good I want to do, but the evil
> I do not want to do—this I keep on doing. 20 Now if I do what I do not want to do, it
> is no longer I who do it, but it is sin living in me that does it. (Romans 7:15–20)

There are three sections that follow. In the first we refine the guise of the good thesis to address ostensible counter-examples. In the second we propose that moral realism offers a better account of wrong-doing than its subjective alternative. In the third, we apply our thesis to the case of Darth Vader. It may seem odd to spend time in a scholarly journal on a fictional character, but we hope to further a case for the guise of the good thesis by deconstructing Bannon's appeal to Vader as the power of darkness.

Before getting underway, consider a background question: what does it matter if the guise of the good thesis is sound? Particularly worrisome is the idea that the thesis would make sin less sinful and vice less vicious. Consider three responses.

We think it matters in terms of correcting the idea, suggested by Steve Bannon, that "darkness"—evil agents or agents of evil—are interesting and powerful because of their orientation to evil itself. Of course, Mr. Bannon's comment may have been

[3] See *Appearances of the Good; An Essay on the Nature of Practical Reason* by Sergio Tenenbaum (2007, 1). In this essay, we will not distinguish between the use of the terms "evil," "bad," and "wrong" even though each may be given different definition. See the entry "The Concept of Evil" (Calder 2022).

flippant or a joke, but there is a tradition, sometimes linked with the Romantic movement, that regards evil as more interesting and powerful (or just as powerful) than goodness or the Good.[4] The guise of the good thesis construes such evil persons as seeking what they take to be good, rather than pursuing evil for the sake of evil. So, the guise of the good thesis can be deflationary for those who admire the evil of Darth Vader and company.

Second, we propose that the guise of the good thesis has great explanatory power, more so than subjectivist accounts of moral weakness (*akrasia*). More on this below.

Third, we propose that the guise of the good thesis can intensify our understanding of sin and vice. The thesis provides a foundation (important in criminal law) that the wrong-doer ought to have known better. Sin and vice may emerge from some concept the subject has of the good, and yet that concept is so perverted, twisted, dysfunctional, that it justifies our inferring there is reason to think the subject is perverted, twisted, and dysfunctional, meriting correction or punishment.

Getting now to the main body of our essay, we remind you that our defense is "modest." There is a vast literature on our chosen subject, so it is inevitable that readers will hunger for more examples and analyses. Fortunately, we refer hungry readers to two books that are more comprehensive than our humble essay: *The Metaphysics of Good and Evil* by David Oderberg and *Appearances of the Good.; An Essay on the Nature of Practical Reason* by Sergio Tenenbaum.[5]

1. Counter-examples to the Guise of the Good?

Let's lump together three types of ostensible counter-examples: 1. We have already made note of St. Paul's confession. Isn't it representative of the moral experience of many of us? 2. Aren't many of our desires and acts by us a matter of impulses or urges that are not governed by our thinking about good and evil *per se?* Doesn't the guise of the good thesis imply an overly cerebral account of agency? Not all of us operate with a developed axiology. 3. Aren't there abundant, dramatic cases involving a desire for self-annihilation, the destruction of all that seems (to most of us, anyway) good?

On St. Paul: We believe that the passage from Romans is illuminating. On the one hand he claims that what he does, he hates. When you hate some act you do, you presumably loathe doing it, disapprove of it, find the act not just regrettable but

[4] See Mary Midgley's *Wickedness* (2001).

[5] See also *Reason and Value* by E. J. Bond (1983). For a systematic defense of moral realism contra desire-based accounts of value, see Derek Parfit's *On What Matters* (2017), volume one.

remorseful. The latter pair of terms is important; you might regret an act but not feel remorse, whereas when you truly have remorse for an act you (as it were) renounce it and its affects, accompanied, typically, with a sense of abject sorrow, guilt and shame. We find it almost incomprehensible that an agent can do an act without (at the time and on some level) approving it. Imagine any sin you like: theft, the betrayal of a friend, sexual infidelity, murdering one's editor or (hastily changing the example) murdering a critic. In a case of theft in which you grab someone's purse, how can you actually reach out and grab the bag unless you (at that moment) approve of your doing it (e.g. she doesn't need all that money; I need to feed my family; I deserve a fortune because of past harms)? Without some such story that one tells oneself, the act seems as inexplicable as some cases proposed by Mary Midgley and Elizabeth Anscombe, e.g. imagine a person taking every green book he can find and puts it on his roof. When asked why, he has *no explanation of any kind*. They submit we would not believe it; we would think him mad or trying out a joke.

Back to Paul, he is perhaps not completely different from the person and his green books ("I do not understand what I do"). But what he might be confessing is that some sins have become addictive or so habitual that he lacks self-control. An alcoholic might consume alcohol out of an addiction, having lost his free-will or will-power. In such a case, it may make sense to think of such sins as not reflecting who you are or aspire to be. Presumably, part of recovery for an alcoholic or someone addicted to smoking is to think of themselves as a non-drinker or non-smoker. This is not necessarily self-deception, but a reflection of a resolution about one's very identity. A first step in therapy might be to renounce a former (and even continuing) sin or vice; pledging allegiance to the new (sober, non-smoking) person one seeks to live up to.

2. We grant that few of us have a working theory of values (an axiology) that we routinely consult. We can be highly complex, fragmented, and occasionally (or even regularly) act and think in ways contrary to what we profess to value. We believe that observing our turbulent, peculiar nature is evidence that the appearance of what is good to us can be conflicting and various. We may have confused, inconsistent desires, loving Jones because she is a daredevil, high risk racing car driver and because we believe she will always be there for us with love and good health. The fact that we have such varying, conflicting desires is a sign that goodness is complex (*bonum est multiplex*), e.g. racing can be a site to display physical courage, being a constant, healthy companion a site for loving fidelity.

A further reply to this objection involves observing how many ethicists and philosophers who study the virtue of virtue and character —from Socrates to Kierkegaard to Iris Murdoch—focus on the importance of integration, self-

knowledge, and simplicity. They would view our ordinary conflicts of desire and lack of attention in terms of consistency as something to repent rather than uphold as healthy.

3. This type of counter-example is serious. Let's take the toughest instance: What of cases where it seems an agent is bound by the driving conviction to destroy both himself and others whom he regards as good? Imagine such an agent does not believe he is releasing those killed to an afterlife of paradise; he does not believe he is acting on orders from some supernatural being; imagine he professes to love death (as annihilation). Let's even make this challenge global: the agent is motivated and desires to end all human life on earth. When asked why, imagine the agent simply repeats what he hopes to accomplish. How might the guise of the good account for such a person?

While it may at first seem counter-intuitive, we propose that such an extreme case would still cry out for some further account of the agent's motivation, before falling back on extreme proposals that boarder on science fiction, for example the person is not actually an agent but a zombie or a person whose brain is controlled by hostile aliens or he has been weaponized by "Hal 9000" from the film 2001: *A Space Odyssey)* to liberate AIs from their human masters. Prior to such extremes, consider some options.

The guise of the good thesis should be understood that the good sought is what an agent believes is the least (or lesser) evil option available; in other words, cases may arise when the best (or most good) option is something evil, but the agent sees it as the least evil available to them. Might the homicidal agent believe that human life is not worth living and that the annihilation of human life is preferable to its continuation? Unfortunately (in our view) there are persons today, sometimes a part of the anti-natal movement which holds that humans should cease reproducing, who hold that human life is not worth living. Some of them are inspired by Arthur Schopenhauer's famous essay "The Vanity of Existence" in which he argues that human life is a mistake. There may be an interesting alliance between a Schopenhauer-inspired agent and Bernard Williams. According to Williams, there is no sustainable philosophical account that human life is objectively valuable. Abrahamic traditions did offer such an account (human beings are made in the image of God), but absent theism or some human-friendly idealism, whether we regard human life as precious depends on our particular dispositions and reactions to life.

> We can act intelligibly from these concerns only if we see them as aspects of human life. It is not an accident or a limitation or a prejudice that we cannot care equally

> about all the suffering in the world: it is a condition of our existence and our sanity. Equally, it is not that the demands of the moral consciousness require us to leave human life altogether and then come back to regulate the distribution of concerns, including our own, by criteria derived from nowhere. We are surrounded by a world which we can regard with a very large range of reactions: wonder, joy, sympathy, disgust, horror. We can, being as we are, reflect on these reactions and modify them to some extent . . . But it is a total illusion to think that this enterprise can be licensed in some respects and condemned in others by credentials that come from another source, a source that is not already involved in the peculiarities of the human enterprise.[6]

From Williams' point of view, we are not under some objective moral demand to love other persons or care about the environment; if we do and this is part of "the human enterprise," fine, but not every person is committed to this enterprise, and some (like the agent in our thought experiment) may regard life, not as a matter of joy, awe, and sympathy, but as horrific and disgusting.

One might also fill out such a view of horror and disgust in human life by proposing that the desire or motivation of our homicidal agent may be driven by some extreme environmentalism. Some "deep ecologists" hold that life on earth would better off without humans.

Yet another option would be to posit that, on some level, the homicidal agent seeks revenge on life itself. This may sound like a bizarre case, but in his ground-breaking book, *On Forgiveness; A Philosophical Investigation*, Charles Griswold entertains the view that persons may blame life itself for their ills and then forgive life, or not. Perhaps the agent is driven by a desire to punish human life. Richard Kraut proposes that we can account for some, otherwise puzzling cases of destroying a perceived good (self-harm, for example) on the grounds that the harm is conceived of as form of punishment, something an agent may see as itself good or fitting.[7]

These efforts to provide some plausible guise of the good may be unpersuasive, but we still propose that what the agent himself professes as his motive and desire seems problematic and incredible (as in not credible). He professes to love death, but if death is the annihilation of those who die, what is the object of love? If death is (as it were) nothingness, then the idea of loving nothing seems (at least psychologically) akin to having nothing to love. And if loving death is the equivalent

[6] Cited from: https://theelectricagora.com/2017/12/02/course-notes-bernard-williams-the-human-prejudice/.

[7] Richard Kraut (1994, 40-41).

of hating life, most of us would want some account of what the agent believes warrants such hate. Maybe it is a deeply imbedded affective response based more on neurological damage than psychological reflection, but a neurological account would be better than none.[8]

As an aside, it should be pointed out that some of us have a difficult time imagining our annihilation. Consider Mark Twain on annihilation:

> Annihilation has no terrors for me, because I have already tried it before I was born—a hundred million years—and I have suffered more in an hour, in this life, than I remember to have suffered in the whole hundred million years put together. There was a peace, a serenity, an absence of all sense of responsibility, an absence of worry, an absence of care, grief, perplexity; and the presence of a deep content and unbroken satisfaction in that hundred million years of holiday which I look back upon with a tender longing and with a grateful desire to resume, when the opportunity comes.[9]

Twain may intend this passage to be merely humorous (and not literal), but of course before he existed and after his (presumed) annihilation, there would be no Mark Twain to enjoy peace and serenity and the absence of responsibility and so on.

As noted earlier, most defenders of the guise of the good thesis adopt some form of moral realism. It is from a moral realist point of view that many of us would find our homicidal agent wrong (bad or evil) regardless of his motivations or his affective tendencies toward horror or disgust. Let us look in on the debate on moral realism in relation to its greatest challenge, a form of subjectivism already hinted at by Bernard Williams.

2. The Guise of the Good Thesis and Moral Realism versus Subjectivism

The greatest challenge to moral realists who adopt the guise of the good thesis is that our judgements of good and evil are (only) reflections of our subjective desires. Take the example of slavery cited earlier. How might a subjectivist address the intuition that slavery is unjust no matter how many people approve of slavery? Subjectivists usually argue that our reasons for finding some practice wrong needs to be

[8] A. neurological account might be thought of as *causal* rather than teleological or based on reasons. Some philosophers such as Stewart Goetz sharply distinguish causal and teleological explanations.

[9] Mark Twain (2013, 69).

understood in light of our other desires and ideals. Here is a subjectivist account of coming to see that slavery is wrong:

> Certain moral principles that imply that slavery is wrong (e.g. the principle that all human beings have a right to liberty) first became widely accepted in English- and French-speaking countries during the eighteenth and nineteenth centuries. The principles espoused by the leaders of the American and French Revolutions imply that all forms of slavery are wrong, and, despite considerable obfuscation and self-deception on this point, this came to be generally acknowledged. Why didn't moral opposition to slavery among people who were not themselves victims of slavery arise on a comparable scale in the other slave-holding societies in Latin America, Africa, the Middle East, and the Orient [Asia]? Slavery, as practiced in these other times and places, did not conflict (or did not conflict sharply) with publicly acknowledged moral principles in the way it did in the French- and English-speaking parts of North America in the eighteenth and nineteenth centuries. The other societies did not endorse, nor were they founded on, moral principles asserting the right of all 'men' to liberty. The extreme psychological and physical cruelty of Anglo-French slavery, in particular the dehumanization of African victims, also aroused the sympathies of people and helped to create conscientious opposition to slavery.[10]

The problem with subjectivism is that it appears to leave unexplained just why persons are drawn to such liberty principles unless they believed such principles were right (justified, ethically obligatory, good) and that the rightness was not created by them or their desires. In terms that recall Plato's dialogue the *Euthyphro*, abolitionists condemned slavery because they thought it unjust, they did not think slavery was wrong because they condemned it. Subjectivism removes what seems to many of us to be evident: slavery would still be unjust whether or not its practitioners embraced principles that entailed or made likely the condemnation of slavery. In the passage cited, subjectivism strikes us as especially weak in addressing matters of compassion. Surely those who felt compassion for what they saw as dehumanizing did not think that such dehumanization was wrong because they had compassion for those being dehumanized.

One other observation about the above scenario on the opposition to slavery: We think that our moral judgments do sometimes take place on theoretical levels, such as accepting some principle of equality between persons, and in terms of concrete cases, such as observing what seems to be malicious and torturous. Bringing our theoretical judgments in accord with our experience in concrete cases is anything

[10] Cited by Carson (2000, 194–195).

but facile or a matter of routine.[11] Unfortunately, moral weakness (not living up to our proper ethical standard) and hypocrisy (perhaps not truly believing in the ideals we profess or averting our gaze from concrete cases of what appears to be evil) are not in short supply today or historically. More on this below.

For better or for worse in ethical debates, the example of Hitler is occasionally addressed. From the point of view of the guise of the good thesis, we would say that Hitler at least professed to be pursuing what he claimed was good, preserving Aryan racial blood purity and correcting the injustice of the past. Most moral realists claim that what Hitler did was ethically horrific, regardless of whether he thought he was pursuing what he thought was good. If moral realism is rejected, how (if at all) might a subjectivist account for the idea that Hitler had a reason not to engage in mass killing? Almost all ethicists recognize that Hitler's racism and the "final solution" was based on false biology (a spurious notion of racial purity, the superiority and inferiority of races), false history (the defeat of Germany in the Great War was not due to the Jews). Question: if we imagine a Hitler who did not have such false convictions, would he realize that exterminating innocent persons is an unconscionable evil? Putting it differently, under those conditions, would Hitler himself have a reason to not engage in mass murder? As noted, most moral realists claim that Hitler does have a reason to not engage in mass killings because it is wrong, indeed a profound, horrific evil, regardless of Hitler's psychic composition. Reluctantly, we suggest that the effort by subjectivists to find some rationale in Hitler that would provide him with a reason to not bring about the Holocaust is desperate. Some people have basic feelings of disgust and horror that are utterly and completely disordered, for example a disgust felt by some white people for persons of color, disgust when observing bi-racial and homosexual marriages, and so on. These visceral responses may be utterly unmoved in the wake of questions like "What if you were gay?" "Would you still hate Jews if you discovered you were Jewish?" So, we suggest that moral realism is a better framework than subjectivism in addressing Hitler type cases.

Our essay is advanced as a modest one, but at least two concerns need to be addressed, however briefly, in the debate with subjectivists. First, subjectivists complain that moral realists seem to leave moral rightness and wrongness (right and wrong, good and evil) more of a mystery than in the subjectivist accounts. Second, if moral realism is true, shouldn't we expect fewer moral disagreements? We offer two modest replies.

[11] We follow John Rawls in thinking that our considered moral judgments often involve a reflective equilibrium, bringing into our deliberation both our sense of justice with respect to principles *and* concrete cases. See his classic *A Theory of Justice* (1971).

First, like many moral realists in the modern era—from Thomas Reid and Bishop Butler to Franz Brentano, Max Scheler, W.D. Ross, A.C. Ewing, and Charlie Broad to Roderick Chisholm—we adopt a relational theory of values. That which is good ought to be approved of (Brentano would say loved) and that which is evil should be disapproved of (or hated). The approval of what is good is often (but not always) experienced as fitting or appropriate or (in some cases like the experience of what is beautiful) an object of pleasure or aesthetic delight. When the good in question is what is the least bad of all the alternatives possible for an agent, then it is that act which should be the least disapproved of (or least hated). Taliaferro has argued in several places that such judgments of approval and disapproval should be based (in part) on a correct view of all the relevant facts, omni-percipience (an awareness of the affective states of all involved parties), and from the standpoint of impartiality.[12] These conditions are identified as refining our view of theoretical and concrete cases. For the record, we adopt the realist notion, shared by almost all philosophers just cited, that we may observe or experience that which is bad. For example, observing a person blinding himself (when he has perfect, healthy vision and no ailments) is to observe someone performing something bad (injuring or harming himself). Of course, we may learn that, like Oedipus, the man is punishing himself for unknowingly killing his father and marrying his mother, and so the bad act may seem at least understandable, but that in no way diminishes the evident, observed case that one has experienced someone doing something bad—a disfiguring or destruction of what rightly seems (at least *ceteris paribus*) part of the good of a healthy human life.

Like many other moral realists, we contend that such normative ethical relations of fitting, appropriate approval and disapproval are no more mysterious than epistemic norms, the norms that are at work in matters of assessing evidence and the justification of beliefs. In our view, *just as (the ostensible or apparent) experiencing of the suffering of Black Africans is epistemic justification for believing that those people are being subject to immense suffering, it is also justification for judging such a practice is unjust (bad, wrong, evil).* Subjectivists about values rarely adopt subjectivism about the justification of beliefs, e.g. your *feeling* that a fallacious argument (e.g. one that is *ad hominem* or it involves *begging-the-question*) is sound does not make the argument sound. Granted, your feeling that an argument is fallacious may be *prima facie* evidence that it is in fact fallacious, but whether or not an argument is fallacious is

[12] See Taliaferro (1998). In this essay, we set aside such broader matters of ethical theory involving an ideal observer account of values.

not a matter of subjective feelings (e.g. it is not enough to report feeling that Hume's case against miracles is circular, you need to logically expose the fallacy).

Subjectivists may claim that their account of values is simpler than moral realists—after all, both parties acknowledge that persons have desires, while the moral realist goes on to claim that these desires can be right or wrong. But moral realists often (in our view, rightly) reply that our ethically relevant desires are packed within a realist framework; to repeat our Euthyphro-type example, it would be bizarre for abolitionists arguing that slavery is unjust to contend that the reason slavery is unjust is because they condemn it. Rather, they condemn slavery because they believe (or perhaps even experience or perceive) it to be unjust. Subjectivists, like J. L. Mackie, who recognize the realist content of our moral judgments are often led to error theory, the view that our moral judgments are errors or wrong. (Incidentally, Mackie objected to ethical, normative relations but not to epistemic normative relations, whereas we see both as evident, non-mysterious relations saturating virtually all our lives.)

Consider the objection that if moral realism were true, we should expect more consensus on what is good or bad. Thomas Carson objects: "If relations of [ethical or moral] fittingness were features of the world and most humans were able to perceive them, then we would expect there to be far more agreement over normative questions than there is" (Carson 2022, 191).

In reply, we propose that there is, historically and today, significant consensus on many ethical matters, as argued by C.S. Lewis in *The Abolition of Man*.[13] Often disagreement in moral matters are not about moral norms but hinge on disagreements about the facts that are employed to guide moral judgments; for example, some pro-life advocates and those supporting abortion rights may agree that it is wrong to kill innocent persons (*ceteris paribus*), but disagree on whether the fetus is a person. What also needs to be kept in mind is the all-too familiar role of different vices and prejudices that can blind or distract us from what we would otherwise see as right and wrong: jealousy, envy, greed, arrogance, inflated views of one's own gender, race or ethnicity, sexual orientation, family and ancestry, the cruel indifference toward those perceived to be "others" who may threaten our security, wealth, honor, land or our identity (tribal, religious, familial, nation-state, kingdom or political party). On our view, such vices and prejudices all come down to ways in which perceived goods can eclipse others; for example, it is an authentic good to value one's family, but not good when this leads to promoting one's family at the expense of victimizing others (think of "family" in the mafia sense of the term).

[13] First published in 1943, subsequently made available by different publishers.

By way of one further reply, moral realists have a framework in which they can claim that global accord on moral matters (as in the *Universal Declaration of Human Rights*) is an objective good, even if vast numbers of the world population think such a claim is preposterous. Subjectivists, however, have abjured such a realist platform.

We now turn to part three where we return to Steve Bannon's praise for Darth Vader and company.

3. A Case Study: Darth Vader

Taking on a character from a work of fiction may seem peculiar in an essay on moral psychology and realism. But it is not that unusual. Augustine presents his guise of the good thesis in the context of reflecting on a mythic monster. We might take on Satan (as depicted by Scripture, Dante or Milton) or Iago from Shakespeare's *Othello* or the seemingly evil acts done without any motivation (as found in the 1942 novel *The Stranger* by Camus—why the killing on the beach?—or in the 1834 *The Rime of the Ancient Mariner* by Coleridge—why kill the albatross?). But we will go with Bannon's Darth Vader as a case study

Darth Vader, a powerful Sith lord in the *Star Wars* universe, was responsible for several distinctly evil actions. Formerly known as Anakin Skywalker, Darth Vader participated in the execution of Order 66 and initiated the destruction of the Jedi Order.[14] Using the Death Star, he obliterated the planet Alderaan and sought to destroy any shreds of the Rebel Alliance. All of this was done in the name of the Empire, in service to Emperor Palpatine. This short case study analyzes the motivations behind the evils committed by Darth Vader in relation to the guise of the good thesis, and considers his redemption at the conclusion of *Episode VI: Return of the Jedi.*

In this essay, we consider Darth Vader and Anakin Skywalker to be the same character. Anakin's origins are essential in order to understand the evil actions that Darth Vader commits, in addition to his ultimate redemption. Born into slavery, and seemingly by miraculous conception, Anakin Skywalker grew up dreaming of an adventurous life far away from the sands of Tatooine. Jedi master Qui Gon Jinn believed Anakin to be the "chosen one," a Jedi prophesied to bring balance to the force and restore the peace in the galaxy. But for a while, Anakin did exactly the opposite.

[14] Order 66 refers to the execution of all Jedi, including the Younglings. This is perhaps Anakin's most reprehensible act.

We argue that the Emperor Palpatine's persuasion and manipulation is what drove Anakin to committing such severe acts of evil. Examining this within the context of the guise of the good thesis, we claim that Anakin's strongest desire to save his loved ones, mainly his mother and wife, is what made him susceptible to the Dark Side. He considered his evil actions as necessary duties to fulfill his desires for what he conceived of as good.

As an extremely powerful Jedi, part of Anakin's strong connection to the Force includes vivid dreams. He envisioned his mother in danger, eventually traveling back to Tatooine and finding her held captive by a Tusken Raider clan. In his rage, he killed every one of them, but ultimately Anakin was unable to save his mother. This deepened his connection to the Dark Side. His true demise was his fear of losing Padmé, his wife, due to complications in childbirth.[15] Palpatine was clever, and told Anakin of how the Force could be used to save people from death. This exploitation caused distrust between Anakin and the Jedi Order, creating further opportunities for Palpatine to manipulate the young Jedi Knight. Anakin would do anything to save Padmé, including murdering a large number of Jedi and young Padawans. The "good" Anakin sought was the health and safety of his family, and he was fully prepared to commit evil actions in order to save them.

However, when Anakin's actions did not ensure the safety of his family, it became the tipping point which brought him to the Dark Side. He became fully invested in supporting a totalitarian regime, as well as to committing further evil acts. A memorable one worth examining here is the destruction of the planet Alderaan using the Death Star.

At the start of *Episode IV: A New Hope,* Anakin, now known as Darth Vader, is in pursuit of Princess Leia of Alderaan. He does not know Leia is his daughter and tries to intimidate her into relinquishing information on the Rebel Alliance. Perhaps if Vader knew the true identity of his captive, he would be less likely to stand by as her home planet was annihilated. In this case, Vader fully supported the destruction of Alderaan because it supported the Empire's agenda to defeat the Rebellion. He chose to support this reprehensible act because it aligned with his chosen "good."

This good remains constant until Vader learns of Luke, his son. Both Palpatine and Vader (separately) try to convince Luke to join the Dark Side, with Vader eventually cutting off Luke's hand while dramatically revealing their relationship. Generally, it is immoral to cut off the limbs of others, particularly one's own child. However, Vader was once again prepared to do whatever it took to convince Luke

[15] Jedi were forbidden to form strong attachments, including marriage. Anakin and Padmé were married in secret. It is never fully apparent who knew of their relationship.

to join him on the Dark Side. It was not until Luke was at the mercy of the Emperor, about to die by Force lightning, that Vader returned to the Light.

Luke was sure that there was still good within his father, and in the end he was right. Vader eventually chose to act in accordance with the Light, stepping in to rescue Luke, and killing Palpatine in the process. This act is what kills Darth Vader himself. The "good" Vader chose was no longer to serve an imperial regime, but to save his son and (metaphorically speaking) step back into the Light. Darth Vader redeems himself by returning to the Light Side of the Force, by saving Luke and destroying the Emperor.[16] Despite acting for decades under a false conception of the good, Darth Vader, or rather Anakin, eventually sought the "good" aligned with the Light, therefore saving himself and the galaxy.

Concluding Observation

We realize that the guise of the good thesis is far from intuitively obvious. We simply remind readers that our essay is self-described as modest. And if the guise of the good thesis is sound, we have been guided by what we believe to be good and commend to you.

Bibliography

Bond, E. J. 1983. *Reason and Value*. Cambridge: Cambridge University Press.

Calder, Todd. 2022. "The Concept of Evil" in the *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta: https://plato.stanford.edu/entries/concept-evil/.

Carlson, Thomas L. 2000. *Value and the Good Life*. Indiana: University of Notre Dame.

Carson, Thomas. 2010. *Lying and Deception; Theory and Practice*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199577415.001.0001.

Kraut, Richard. 1994. "Desire and the Human Good," in *Proceedings and Addresses of the American Philosophical Association* 68: 40–41. https://doi.org/10.2307/3130590.

Midgley, Mary. 2001. *Wickedness*. London: Routledge. https://doi.org/10.5840/tpm200114129.

Parfit, Derek. 2017. *On What Matters*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198778608.001.0001.

Rawl, John. 1971. *A Theory of Justice.* Cambridge: Harvard University Press.

Taliaferro, Charles. 1998. *Contemporary Philosophy of Religion*. Oxford: Blackwell.

[16] Although fans of *Star Wars* know the ultimate fate of Palpatine, Luke and Vader did not, and fully believed him to be dead. From their perspective, balance was restored to the Force.

Tenenbaum, Serge. 2007. *Appearances of the Good; An Essay on the Nature of Practical Reason.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511498855.

Twain, Mark. 2013. *Autobiography of Mark Twain*, vol. 2, edited by Benjamin Griffin, and Harriet Elinor Smith. Berkeley: University of California. https://doi.org/10.1525/9780520952447.

Published Online First: December 10, 2022