

Moral Dualism and the Problem of Evil

Isaac Nevo

Department of Philosophy, Ben-Gurion University of the Negev, Beer Sheva, Israel

Email: yanni@bgu.ac.il

How to cite this paper: Nevo, I. (2024). Moral Dualism and the Problem of Evil. *Open Journal of Philosophy*, 14, 404-423. <https://doi.org/10.4236/ojpp.2024.142026>

Received: March 28, 2024

Accepted: May 8, 2024

Published: May 11, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The aim of this paper is to argue against moral dualism in the understanding of the nature of evil, namely the conception of evil as an independent source of guidance, in opposition to the good, rather than a failure in pursuit of an apparent good. Focusing on moral evil as the intentional infliction of gratuitous pain and suffering by one human being on another, i.e., pain and suffering that are not required by a morally acceptable purpose, I argue against two forms of such dualism. Value dualism divides moral value into antithetical normative principles, good and evil, each with its own guiding power. On this view, evil can intelligibly be pursued for its own sake, rather than a failure of some kind in acting “under the guise of the good.” Agent dualism divides human agents, based on character and disposition, into followers of good and followers of evil. On this view, the pursuit of evil can be accounted for in terms of basic character traits and dispositions, not related to more fundamental motives, reasons, or choices. In both versions humanity is divided into two moral classes. I argue that the two forms of moral dualism discussed in this paper fail to render evil perpetrators intelligible in terms of reasons for action. While suggesting, in line with accounts by [Arendt \(1951, 1963, 1978\)](#), [Anscombe \(1963\)](#) and [Neiman \(2002\)](#), a non-dualist account of evil-doing as a dysfunction in the pursuit of intelligible goals, I will go on to criticize dualistic views of both kinds in the work of philosophers such as [Velleman \(1992\)](#), [Silber \(2012\)](#), [Bernstein \(2002\)](#), [Hacker \(2021\)](#), and [Kekes \(1990\)](#).

Keywords

Evil, Moral Dualism, The Guise of the Good, Moral Egalitarianism

1. Introduction

In experiencing evil, what is it that we are experiencing? The antithesis or the failure of goodness, a principle of moral darkness or a shadow where the light is

blocked? In what follows, I wish to revive this question, the question of moral dualism (Manicheism), as it pertains to moral, rather than natural or other forms of evil, but independently of the theological concerns with which the problem of evil had been traditionally interwoven. I shall focus on moral evil as the intentional and deliberate infliction, by one human being on another, and possibly on other sentient organisms, of gratuitous pain and suffering, i.e., pain and suffering that are not required by a morally acceptable purpose, in the grim forms of death and destruction, exploitation and humiliation, torture, abuse and neglect. The question is, how should this sub-category of the immoral, namely evil, be understood? Moral dualism suggests itself as one answer to this question, albeit one that does not render human perpetrators fully intelligible even as it satisfies a felt need for the strongest possible condemnation. I shall argue against two versions of moral dualism, in which evil is conceived as a principle or motive inherently opposed to what is good, in defense of a non-dualist view of evil as a human failure in the pursuit of an apparent good.

Famously, [Arendt \(1951\)](#) had changed her mind on this question, admitting in a letter to G. Scholem that her earlier view on radical evil, as described in her *Origins of Totalitarianism*, suggests a demonic dimension of depth regarding what was “only extreme” in political murderousness ([Arendt, 1978: 251](#)). While in her early book, Arendt took Kant to task, for having rationalized radical evil as merely “perverted ill-will” which could be understood in terms of “comprehensible motives” ([Arendt, 1951: 459](#)), in her later imagery, rather than having a “truly radical nature” (1951: viii-ix) evil was seen as more like a “fungus” which can “lay waste to the whole world,” while itself being revealed as “nothing.” (1978: 251). Thus, she moved from a dualist axiology in which good and (absolute) evil were thought of as opposing and independent principles, each with its own intrinsic quality, to a non-dualist value-theory, for which only the good, or lack thereof, has any normative pertinence (“depth”), but she faced criticism for her related notion of the banality of evil, namely, evil as thoughtlessly perpetrated with entirely petty motives, as failing to do justice to the complexity of the phenomenon in question, portraying the perpetrator (Eichmann) as too much of a “cog in the machine” with little complexity of his own. But whether or not banality is an appropriate description of evil as perpetrated in pursuit of comprehensible motives, the status of evil, either as its own “principle” or as a failure in pursuit of the (apparent) good, remains puzzling.

There are both value-theoretic and moral-psychology questions underlying this issue. The conceptual, value-theoretic question is whether judgments of evil require a bifurcated scheme of values, a dualism of both good and evil as antithetical sources of normative guidance. Alternatively, moral evaluation could be seen as the application of value-ideals upon a recalcitrant reality, which may not measure up to such ideals but is not itself intrinsically evil. The latter view replaces the duality of good and evil with an assessment of the presence or absence of some “positive” value. On this view, evil is due to some failure in the applica-

tion of value, rooted in the world of human affairs not in any “principle” of its own.

The psychological question concerns the implications of value dualism and value idealism regarding our understanding of human conduct. On the dualist conception, the account of human behavior must make sense of intrinsically evil motives or dispositions; on the idealist side, some account must be given of the many possibilities of perpetrating evil while pursuing an apparent good. In so far as human action, being intentional and deliberate, is answerable to reason, i.e., can be rationalized in terms of ends and means, positive values (the good) appear to be fundamental, for negative ones (evil) do not serve the purpose of such rationalization. If there are intrinsically evil motives or dispositions, reason-based accounts of human conduct must be set aside. In favor of some alternative, dispositional or libertarian, moral psychology. Examples of both these alternatives will be discussed below, with some arguing that even perpetrators of great evil, indeed, even Satan himself, could only be understood as acting under the guise of some (intelligible) good, while others insisting on the psychological possibility of acting out of pure malice or in the service of inherently evil character dispositions.

2. Value-Privation and Dysfunction

Taking a clue from traditional privation theories of evil, I shall argue against various forms of moral dualism that still surface in contemporary accounts.¹ Going back to Augustine, privation theories of evil aimed both at rejecting Manichean dualism and supporting the benevolence of an omnipotent God.² Here, I shall focus on the former issue, suggesting a limited, evaluative rather than ontological, privation theory of evil, or rather of the grounds of evaluating (or judging) evil. The crux of this limited privation theory is that although evil resides in positive, rather than privative conditions, namely, actions, dispositions, intentions, decisions, or other psychological and societal facts that are not themselves “absences” (or “corruptions”) of other things, the evaluation of these non-privative conditions as evil points to the absence of value. Privation is judged relative to a task of evaluation. To evaluate something, morally or otherwise, is to assess the presence or absence of some value in given circumstances. The values may be projected by us, but once projected we have the task of assessing their application. Cruelty, for example, resides in a character disposition, which is very much existent, to enjoy the suffering of others; but whether it is

¹Originating with Augustine, the privation theory of evil became the predominant Christian theological doctrine. See, for example, the Confessions of Augustine (1952). Bk. 7, ch. 11-13, Enchiridion (Augustine, 2012), Ch. 3-4. See also St. Thomas Aquinas (1995), *On Evil*. For contemporary discussions of the Privation theory, see G. Stanley Kane (1980), Anglin & Goetz (1982), Todd C. Calder (2007).

²Manicheism, a religion founded by Mani, a Mesopotamian figure, in the 3rd century AD, gained influence in the Roman Empire but survived mainly in Asia for Millennium thereafter. Its followers, including for some time St. Augustine before his conversion to Christianity, subscribed to belief in a cosmic dualism of light and darkness. For more on the history and doctrine of the Manichean religion, see Burkitt (1922).

evil in implementation is not intrinsic to the disposition itself. The cruelty of a sadist need not be evil if it is satisfied in agreement with the masochist. It becomes evil only when such agreement is lacking, i.e., when respect for the dignity and humanity of the recipient is not in place. Evil is thus evaluatively, not ontologically privative. Similarly, inequality resides in social conditions that are very real, but as Rawls taught us it is not intrinsically unjust, for it may be to the benefit of the least advantaged (the difference principle). The question is not why the bad exists, but what makes it evil, i.e., the absence of positive value or the presence of an inexplicable “negative value.”

However, as has been traditionally recognized, such an account is not complete. Absence of value is a necessary but not a sufficient condition for negative evaluation since value may be absent where no negative evaluation is warranted. What negative evaluation requires in addition is some grounds on which the value in question is to be expected, some grounds on which we can view the absence of value to be dysfunctional.³ In the moral sphere such dysfunction may amount to a failure of empathy, altruism, or respect for humanity, assuming these to constitute normal or proper functioning.

More controversially, identifying an absence of value may be thought of as not even necessary for negative evaluation, since negative evaluation proceeds by identifying a presence rather than an absence, the presence of an antithetical value of some kind. On this supposition, evil is not the absence of goodness but the presence of an active force of malice, a guiding principle through which, in Milton’s phrase, evil becomes the perpetrator’s “good.” I shall accept the insufficiency of value-absence for judgments of evil but insist on its necessity. Rejecting any suggestion to the objective reality of an intrinsically evil force, I shall supplement value-privation with (psycho-social) dysfunction as grounds for judgments of evil. Evil arises from human failings. It is not an accomplishment of a Manichean force lurking within us. The hard question is to identify the dysfunctions, personal or societal, that take us from pursuing what is morally comprehensible to perpetrating what is not.

My aim here, though, is not to supply such an account, only to indicate the need for it by questioning its dualistic rivals. Evil-inducing dysfunctions have been discussed at various levels, empirical, historical, and philosophical. Stanly [Milgram’s \(1963\)](#) obedience-to-Authority studies found, for example, that 65% of participants in a carefully designed experimental setting were willing to expose a “victim” to increasing electrical shocks, up to the lethal dose of 450 volts, at the experimenter’s orders to do so. Milgram connected his findings to the facts of collective obedience during the Holocaust, which still serves as a paradigm of evil. Similarly, historians and philosophers spoke of the erosion of moral

³The point was clearly recognized by traditional privation theorists. As [Aquinas \(1975\)](#) has it: “evil is in a substance because something which it was originally to have, and which it ought to have, is lacking in it.” (*Summa Contra Gentiles*, III, Part I, 44). The example given is that the absence of wings in a man is no evil to him, since wings are no part of his nature.

inhibitions resulting from various social conditions. Zygmunt Bauman (1989) describes the dysfunctions required for participation in mass murder. These include legal authorization of violence, routinization by division of labor, and ideological dehumanization of the victims (Bauman, 1989: p. 21).

Accounts such as these presuppose moral attitudes such as empathy, humaneness, or altruism to be components of proper psychological functioning, perhaps rooted, as inhibitions, in human evolution, but ones that are fragile and limited, vulnerable to disturbances from other components of the human psyche or social setting—self-interest, group membership, prejudice, and the like. Social behavior normally balances such conflicting tendencies, realizing some sort of equilibrium. But there are circumstances which force serious imbalances, e.g., erosions of inhibitions, to an extent that makes mass murder possible. These are the cases that call for judgment in terms of evil, namely, the absence of goodness through psycho-social dysfunction.

3. Contemporary Moral Dualism

These considerations raise a modern humanist version of the problem of evil. While the traditional problem involved the possibility of evil given the benevolence of an omnipotent God, the modern problem is how moral evil is to be accounted for given the answerability of motivated human action to reason. Susan Neiman (2002), for example, describes the problem as one about the “intelligibility of the world as a whole” (2002: p. 8). As she points out: “Two kinds of standpoint can be traced... The one, from Rousseau to Arendt, insists that morality demands that we make evil intelligible. The other, from Voltaire to Jean Améry, insists that morality demands that we don’t.” (ibid). Contemporary moral dualism consists in rejecting the answerability of human action to the demands of intelligibility, allowing it, instead, to be guided by intrinsically evil purposes or dispositions.

Kant is perhaps the greatest authority on the rationalist side of the issue, insisting that incorporating “evil qua evil for incentive into one’s maxim” is “diabolical,” and hence, not humanly possible, and that evil is rooted in the “subordination” of the moral law to the incentives of self-love (Kant, 1996 [1793]: 84). Critics such as Bernstein (2002) (following Silber, 2012) have argued that such an approach cannot do justice to the “diabolical” nature of 20th century atrocities, and that a richer moral psychology should allow for such possibilities as choosing to defy the moral law. As Bernstein puts it, “there is no free choice (*wilkür*) unless there is free choice to be morally evil, and even devilish.” (Bernstein, 2002: 42). Kant’s distinction between human depravity and the diabolical is a form of the privation-dysfunction account of evaluation. However, the moral dualist implications of Bernstein’s (and Silber’s) view, allowing evil to be a value pursued, in radical freedom, for its own sake, are yet to be accounted for. Absolute freedom, so I shall argue, is hardly satisfactory as an account of absolute evil; it is to explain the obscure by the obscurer, leaving the explanandum unmotivated and unintelligible.

Contemporary dualism takes two forms, a stronger and a weaker form, that I shall refer to as value-dualism and agent dualism, respectively. The stronger version bifurcates value into antithetical normative principles, each with its own guiding power, that are freely chosen. The weaker version bifurcates human agents, based on character and disposition, into followers of good and followers of evil. In both versions humanity is divided into two moral classes, as against egalitarian moral humanism, namely the view that humanity is unified, morally speaking, in being responsive to reason. Common to both these versions is the view that a moral dualism is necessary, both intellectually and morally, if evil is to be “seen” or “faced,” i.e., if it is to be recognized for what it is, without hiding its true nature under some guise of goodness or rationality, and if it is to be confronted, morally speaking, without offering undue “forgiveness” or “legitimacy,” as the price of understandability—without, that is, bringing evil perpetrators back “into the fold” where a more appropriate reaction would be to cast them out as standing “beyond the pale.” In what follows, I shall consider versions of both these forms of moral dualism.

4. Value Dualism

A contemporary example of value-dualism appears in a curious discussion regarding the moral psychology of Satan—Milton’s (1984) character in *Paradise Lost* who calls: “Evil be though my good” (*Paradise Lost* bk. 4, l. 110 (1667))—on the part of philosophers debating the relation between motivation and evaluation, namely, whether the object of desire is always given under the guise of some good. Milton’s characterization of Satan’s motive is ambiguous. It could be read as suggesting that evil is to be reinterpreted as a good (thereby becoming Satan’s intelligible motive), or that evil is to replace the good as a motive (without any reinterpretation, i.e., evil as such). Let me call these the reinterpretation vs. the replacement readings of Satan’s motives, and by implication of evil motives generally.

Elisabeth Anscombe (1963) dismisses as unintelligible the suggestion that anybody’s motive could be “the good of it is that it’s bad,” namely, that the bad itself could count as an ultimate motive—a desirability characterization that answers the question “why?” regarding the point of an action. If the bad is cited as a motive, a further question immediately arises “and what’s the good of its being bad?”. Regarding Satan, the answer to that question might be “condemnation of good as impotent, slavish, and inglorious.” Hence, Satan’s motive can be made intelligible: “The good of making evil my good is my intact liberty in the unsubmitiveness of my will.” (Anscombe, 1963: 75). Thus, evil is reinterpreted as an apparent good, a certain kind of liberty, and only under that description does it make Satan’s motive intelligible. In other words, even Satan does not constitute a counter example to the conceptual connection between motivation and (positive) evaluation—the guise of the good doctrine—which is Anscombe’s way of saying that even the figure of Satan need not incline us to project evil as a guid-

ing motive, and that the sphere of moral evaluation need not be bifurcated into antithetical sources of value.

By way of criticizing Anscombe's account, Velleman (1992) offers a replacement interpretation of Satan's words, thereby, exemplifying value-dualism. Anscombe's Satan, he argues, is "a rather sappy Satan," whose "horns" have been "blunted," in so far as he "can desire evil only by judging it to *be* good, and so he remains at heart a lover of the good and the desirable....." (Velleman, 1992: 19). For Velleman, Satan does not lose sight of the evil he desires and does not turn into "another well-intentioned fool." (ibid: 18). Rejecting the guise of the good as a necessary feature of desire, Velleman liberates not only Satan from the burden of the good, accepting the possibility that the ultimate object of one's desires need not be good under any interpretation. Human agents, too, he argues, need not be so conceived. Rather they may be "disaffected, refractory, silly, satanic, or punk," and he "hopes for a moral psychology that has room for the whole motley crew." (ibid: 3). So, while Anscombe brings Satan down to earth, assimilating his motives to those of human agents, thereby tacitly rejecting a dualist understanding of evil as an independent source of immorality, Velleman raises (or degrades) human agents to the level of the Satanic, concluding that they too can desire evil for its own sake, without any redeeming guise, thereby, tacitly projecting evil as more than a mere failure in pursuit of the good.

Velleman's Satan does not lose sight of the evil he seeks, and by implication satanic human agents seek evil for its own sake. For this to be possible, evil must replace the good as their source of normative guidance and be itself their "good." But what does it mean to allow evil to become an independent source of normative guidance? What makes this idea conceptually troublesome is that it severs normative evaluation from considerations of value, and ties it, instead, to some principle of "disvalue," or to a force that is intrinsically antithetical to all value, whatever that may mean. After all, to evaluate is to apply value, namely, to measure the presence or absence of value in what is evaluated; when value is found to be lacking in what is evaluated negative terms of evaluation are coherently employed. But to evaluate negatively is not positively to assert the presence of a "negative value," and it is entirely unclear what that would amount to, i.e., what the human followers of Velleman's Satan would be aspiring to.

The debate over Satan's motives, in application to historical events, has also been touched upon by Silber (2012), in his account of "devilishness" as a mode of freedom. In criticizing Kant's view, Silber accepts devilishness, namely, actively rejecting "the moral law," thereby pursuing evil for its own sake, as a human possibility, realized in the figure of Hitler, who unlike Arendt's Eichmann could not be dismissed as banal, or weak of will in any respect. What may appear as a rational impossibility, i.e., intentionally pursuing "evil qua evil," is better understood as a possible irrationality, one that falls within the repertoire of human freedom. Silber relates this to the figure of Milton's Satan, who goes well beyond the Kantian limits of evil. As Silber puts it:

The Satan of Milton's *Paradise Lost*... (Milton, 1984) exemplifies the transcendent sort of devilishness Kant rejects as a romantic illusion of the heroic grandeur of wickedness. But Milton... presents a compelling example of the genuinely demonic. Although Milton's portrayal is a work of imagination, it describes accurately the factual evil we confront in Auschwitz – evil that far transcends the conceptual limits of Kant's theory. In Auschwitz and in Hitler we confront not the wickedness that results from impotence but the demonic evil of a powerful though irrational exercise of freedom. (Silber, 2012: p. 333).

Throughout his account, Silber assumes an “intentionalist” understanding of the Holocaust and the history of Nazi Germany. On this view, the Holocaust was planned and executed “top-down” as an exercise of Hitler's will: “when Hitler willed the final solution to the Jewish question, he also willed with full consistency the means to this end,” (ibid. 315) which fell to his subordinates to assemble and implement. Silber concludes that “in Hitler we confront not an absence of self-directed will but, together with Stalin, one of history's ultimate examples of focused, malevolent volition.” Hence, “the power of Hitler's personality imposes itself on ethical theorists as an historical reality to be accounted for in any ethics relevant to human life.” (ibid. p. 332). Kant's account of evil in terms of the failure to make the moral law one's sole incentive, subordinating it to incentives of self-love, is found to be inadequate in relation to such single-minded determination to reject morality and pursue evil.

Silber's alternative account is given in terms of an expanded conception of freedom. In Kant, freedom of the will is understood as the subjection of the will to the moral, rather than the natural law. Human beings are free as noumenal entities, and their freedom consists in the determination of a higher law (the moral law) which operates beyond the level of phenomena. Otherwise, Kant held, freedom could only amount to lawlessness. According to Silber, however, the historical devilishness of Hitler forces us to accept a more radical conception of free choice as totally irrational.

So, we are to imagine historical, though devilish, figures such as Hitler or Stalin, choosing to reject morality altogether, and doing so irrationally, just as an exercise of their free will, and not for reason of any other “good” that they may imagine as gaining. Their strength of character indicates that like Velleman's Satan, they do not lose sight of the evil they are after by subordinating it to some other interest. Nevertheless, they are irrational; their choice cannot be “rationalized” in terms of any motivating good.

Notice, however, the discrepancy between the clear-sightedness of Velleman's Satan and the irrationality of Silber's devilish characters. Unlike Satan, whose power does not depend on disclosing his grounds, the absence of rationalizing grounds, misjudged as they might be, would leave a historical perpetrator incapacitated, locked in his own inscrutable “choice” without being able, as his role surely requires, to communicate it to others. To generate his power, a human

“devil” must communicate some grounds for his choices, thereby “losing sight” of the evil he seeks to an extent Satan would not need to.

The whole picture is entirely implausible. Perpetrators of great evil are flawed human beings, not healthy devils. Hitler chose to exterminate the Jews for a misguided reason that he could perfectly well communicate to his fellow Germans, who in turn could share his prejudices and misjudgments and support him politically. Much the same holds for Stalin in his choice to collectivize Soviet economy at any cost. They both chose evil freely, and to some extent irrationally, but they didn’t choose evil as such, nor the complete eradication of morality. Rather, they chose great evil in the service of a misguided cause, racial purity, or the classless society, which they themselves, and their supporters, judged to be beneficial in some respect. They were not fully rational in their choices, suffering from prejudices and biases that obscured the self-destructiveness of their aims, but they were not as irrational as Satan’s choice would render them if they were to make it. Their choices surely destroyed morality for their collaborators and victims, but that was inspired by a perverted morality, not the destruction of morality as such. We come back to the distinction between replacing the good in Satan’s motives and reinterpreting it. Historical figures such as Hitler and Stalin can only make sense as engaged in gross reinterpretations of the (morally) good in their motives; if we try to describe them as replacing their “good” with evil (as such) we lose the possibility of understanding not only them but also the sources of their power.

5. Agent Dualism

The problem with value-dualism is that by severing the evaluation of evil from any understandable value it fails to make sense of the agents’ motives. By contrast, agent-dualism does not proceed by attributing an intrinsically evil motive to the agent but goes on to describe evil agents as intrinsically disposed to evil and vice, beneath the level of rational accountability. Thus, [Hacker \(2021\)](#) argues that “seeing” evil for what it is requires an acknowledgement of what is intrinsically evil in the tendencies of perpetrators. Hacker presents the following alternative:

...if there really are evil people and if evil deeds are indeed done, then the power of reason is much more limited than the liberal Enlightenment and their heirs envisaged. For evil people appear to be beyond the reach of reasoned argument with respect to the most fundamental of moral considerations: the infliction of death without warrant and cruelty without constraint. ...This challenges the ideals of liberal Enlightenment concerning the brotherhood of Mankind under the aegis of reason. It is therefore not surprising that the medicalization of evil is commonly favored to show that those who do such deeds are themselves victims of forces beyond their control and are not responsible, or not wholly responsible, for their crimes. For then they cannot be viewed as being ‘beyond the pale’ or as being irre-

deemable. No human being, it may be held, can be evicted from the brotherhood of man. (Hacker 2021: 88-89)

To which Hacker retorts: “Maybe no one can be evicted, but... perhaps those who do evil cast themselves out.” (ibid). Hacker dismisses the alternative of viewing evil as some dysfunctional privation of goodness as not only flawed in overemphasizing the explanatory role of reason, thereby, assimilating the presence of evil to an absence of rational accountability, but also as morally offensive in “medicalizing” evil, i.e., in viewing perpetrators of evil as “victims” of some dysfunction. Conceding that evil itself cannot serve as a motive, for it lacks any justifying power, Hacker still allows that reference to evil as an explanation of action is intelligible in another way. It is, he points out acceptable as a dispositional explanation:

Nevertheless, ‘because he is evil’, like ‘because he is good’, is an acceptable answer to the question ‘why did he (does he) do that?’... it is not an explanation in terms of the agent’s reasons, but rather a dispositional explanation: that’s the sort of person he is. People of that bent tend to do such terrible things. (Ibid, p. 120-121).

While rejecting privation theories of evil, Hacker offers an agent dualism in terms of a fundamental disposition to cause harm and inflict pain. He does not offer an account of how evil might serve as a “value” antithetical to goodness, or how evil could serve as a motive. His argument is that the rational-humanist account of human conduct in terms of reasons-for-action is too weak to account for the reality of evil, and it leads to both intellectual and practical failures regarding our confrontation with evil, the failures of assimilating evil to weakness and excusing the perpetrators through “medicalization.”

On Hacker’s argument, an enlightenment-based conception of evil leads to moral blindness. Those who uphold that view, so he argues, cannot see evil for what it is. They cannot see it at its source, as a fundamental disposition on the part of its perpetrators that is entirely insensitive to rational accountability. Consequently, they underestimate evil as they attempt to view it, like all human conduct, “under the guise of the good,” failing to see its ultimate source in what is inherently, though inscrutably, corrupt. Thus, they convict the action (as evil) but exculpate the motive (as benign), thereby “medicalizing” the agent (as victim of weaknesses), excusing his excesses as failures, and offer treatment (or limited punishment) in place of banishment to somewhere “beyond the pale.” Such is the intellectual price of Enlightenment’s emphasis on the answerability of human conduct to reason as a fundamental human trait.

The basic disposition to commit evil is not analyzable in terms of intelligible reasons. We are thus back with the alternative presented by Neiman between a morality of refusing to understand the perpetrators of evil, for fear of forgiving or exculpating them, and a reason-based account of human conduct that insists on understanding its motives while humanizing even the worst perpetrators. The

former view offers banishment and moral segregation as the appropriate response to evil; the latter offers diagnosis of dysfunction in relation to comprehensible motives, and punishment within the limits of human solidarity. One conceptual issue between these two fundamental views is the alleged connection between understanding and forgiveness that seems to motivate the refusal to understand, or the worry that understanding evil acts in terms of their motives and failures would lead to excusing the agent as merely a victim of her own dysfunctions (“to understand all is to forgive all”). If that connection is false, i.e., if understanding does not amount to any sort of legitimization, then at least one argument in favor of agent dualism can be removed.

However, understanding does not entail forgiveness. While forgiveness may require an understanding of motive, and of the failings that led from motive to offense, it requires a further step beyond understanding. That step is a willingness to accept an apology, or even just a visible change in the offending agent, and a decision on the part of the offended party, to open her heart and to share a communal life with the reformed offender. Mere understanding is certainly not sufficient for that, and it is compatible with an insistence on punishment. Hence, it is entirely possible to understand without forgiveness. What understanding does achieve, morally speaking, is a view of the offender as a human being, despite the evil involved, and it could serve to constrain the punishment accordingly. For punishment is, indeed, not permanent banishment from human fellowship, and it is to be conditioned by the equal dignity (and human rights) that that fellowship confers. There are, of course, hard cases in this regard, e.g., psychopaths whose dysfunction places them beyond the reach of reasoning, or people whose monstrous deeds defy any feeling of fellowship. But monstrous deeds do not require actual monsters, only human beings; and these can, ultimately be diagnosed, analyzed, and understood. Indeed, the refusal to understand evil conduct in terms of its motives and reasons, along with the failures and dysfunctions revealed through them, is not just an intellectual failure. It is also a moral failing, or may lead to one, in so far as it obscures the roots of evil in what is common, and human, and so it leaves us without sufficient intellectual tools to confront it at its source.

Kekes (1990) offers a rich version of agent dualism, derived from what he calls “the tragic view of life,” and a moral theory to suit – character morality – which finds evil to be rooted in character, rather than free choice. Kekes goes beyond Hacker’s “dispositional explanation,” accounting for the dominance of such dispositions, or vices, in an agent’s character as an inevitable outcome of “the essential conditions of life.” It is not that evil is a failure in the pursuit of the good, but rather that good and evil, virtue and vice, compete for dominance over a person’s character as forces beyond her control, and come to the fore in the agent’s habitual, but “unchosen” behavior.

The tragic view depicts human motivation as the arena in which our virtues and vices wage their endless battles, and it forces us to recognize that the

issue remains undecided. Thus, tragedy prompts us to see human character as fundamentally flawed (Kekes, 1990: 5).

Kekes's account touches on the question of the relation between tragedy and morality. Traditionally, scholarship on the tragedy/morality relation had been divided into two rough camps. As Sebastian Gardner (2003) sums it up, one camp "assumes a fundamental compatibility between tragedy and morality..." while the other believes that "there is... a mutual antagonism between tragedy and morality." (Gardner: p. 321). The question is whether the point of view of tragedy can be assimilated to that of morality or is it, rather, an a-moral view, one emphasizing fate over autonomy, and the irrelevance of virtue or merit in preventing catastrophic outcomes. While moralists read moral purposes, or lessons, into the tragic outcome, seeing it as consequent upon preventable flaws, a-moralists object to such interpretations as didactic moralizing that obscures the terror of human existence by surrounding it with blame and reproach. On the latter view, the tragic outcome cannot be rationalized in terms of the protagonist's guilt given his flaws or errors.

Against this background, Kekes offers a third view, the assimilation of morality to the tragic point of view, rather than the other way around, thereby incorporating the a-moralist's view of tragic existence into the fabric of morality itself, accepting (as Kekes quotes from Schopenhauer) "a significant hint as to the nature of the world and existence" (Kekes, 1990: p. 5; Schopenhauer, *The World as Will and Representation*, 1: 252-53). On this reading, morality incorporates what Schopenhauer called "the terrible side of life... the triumph of wickedness" (ibid) as a rival to goodness. Hence, a moral dualism appears, in which human motivation is no longer the sphere of rational accountability but rather "the arena" in which vice and virtue endlessly oppose one another. Moral dualism results from assimilating morality to the tragic perspective, allowing its "hint" regarding the nature of reality to relegate the pursuit of moral ends—goodness, happiness, duty, justice—to just one side of the normative sphere. The central task of morality is no longer merely to diminish the gap between the is and the ought, or the real and the ideal, applying value in a valueless world, but rather to stand in competition with another source of guidance—the wicked—as an inescapable aspect of human motivation.

From this perspective, traditional morality, with the pursuit of goodness at its core, appears very much like the moralist's interpretations of tragedy, namely as an attempt to moralize "the terrible side," the terror at the heart of tragic existence, using morality as a kind of world-bound theodicy, thereby hiding the reality of evil under a veneer of goodness. Kekes describes this moralizing tendency as a "soft reaction" to evil, resting on what he calls "choice morality"—a morality that confines moral judgment to the sphere of human choice, regards human beings as moral equals, and takes goodness to constitute the underlying order of reality. By contrast, he offers a "hard reaction" to evil, the better to "face" and confront it, which he articulates by means of "character morality." Thus,

Kekes does not simply reject morality as an inadequate response to the tragic situation of humanity but offers to reform it dualistically, to incorporate the tragic reality and objectivity of evil into moral theory.

Kekes defines evil as “undeserved harm,” where harm is undeserved when it is not matched by its victim’s demerits. The definition is “externalist” in the sense that it makes no reference to the motivations of the perpetrator. Undeserved harm afflicts human beings due to what Kekes calls “the essential conditions of life,” which are the contingency of human existence, the indifference of nature to merit, and the destructiveness of human action. Such evil is, of course, not always moral evil since the essential conditions of life often hit us without the involvement of other human beings. Death, disease, earthquakes, and similar undeserved catastrophes occur entirely outside the sphere of morality. However, undeserved harm may be incurred by human causes, in which case they constitute moral evil. But though there are, according to Kekes, “moral monsters” who cause evil by choice, most moral evil is caused without such “diabolical” features. For the most part, moral evil is “characteristic, but unchosen.” Being rooted in character, such evil justifies an inference, which lies at the heart of character morality, from the moral status of an action (as evil) to that of the perpetrator.

Character morality emphasizes character rather than choice, the dependence of desert on merit rather than equality of moral status amongst human beings, and a mixed view of human nature, torn between good and evil dispositions, in place of any view which prioritizes goodness as underlying the order of nature or the requirements of reason. Kekes articulates this view by means of nine theses, portraying evil as objective, independent of choice, resulting from the dominance of vice as the essential conditions of life. Character morality emphasizes the power to cause evil, independently of choice, and sanctions an inference from evil actions to the moral standing of perpetrators as inherently evil. Emphasizing the centrality of moral desert and the dependence of desert on moral merit, character morality rejects moral egalitarianism and bifurcates humanity into the deserving and the undeserving. Finally, much as Manicheans viewed cosmic reality as mixed between darkness and light, and sought to re-separate these elements, character morality views human nature as mixed between virtue and vice, locked in endless battle, and seeks to separate the elements in a direction favorable to human welfare.

Kekes argues extensively for all this, primarily by subjecting the soft reaction to evil and the morality of choice that underlies it to a sustained critique. The three principles of choice morality that he considers are the co-extensiveness of the sphere of morality with the sphere of choice, which makes “unchosen (moral) evil” a contradiction in terms, the moral equality of human beings (moral egalitarianism) as possessors of a fundamental moral status and dignity, and the fundamental goodness of human nature. Regarding the first principle, Kekes points out that it turns a blind eye to most (moral) evil, which is unchosen, and thus renders morality incapable of “facing” that evil. Regarding moral egalitarianism, the charge is that it obscures the inequality of merit between people and

thus violates the principle of proportionality between merit and desert, again failing to face evil, or deal with it appropriately. The idea that there is a moral status we all share, regardless of merit, which defines our rights and constrains what can be done to us is, so Kekes finds, groundless. It is “a false belief that pervades the sensibility of an age” (ibid. 108), much like the belief of prior ages in the divine rights of kings, or the sinfulness of premarital sex. Kekes emphasizes that his opposition to egalitarianism is restricted to the question of moral status and does not extend to other forms of equality (class, race, etc.), but that still leaves those other claims to equality without their most important moral basis. Finally, the view that human nature is basically good, and that there is a rational and moral order, in the world or in humanity, in relation to which evil is a departure, involves metaphysical assumption that cannot be sustained.

Contrary to Kekes’s account, choice morality is only one way of grounding the attitude to evil which he calls the soft reaction. An alternative basis for it, and for the fundamental principle of moral egalitarianism, would emphasize not choice, with its Kantian suggestion of noumenal free will, but reason, namely, the openness of human action to rational account, or failure thereof. Human action is subject to rational criticism not in the sense that it is always decided freely among alternatives but in the more fundamental sense that it is intentional, i.e., that it can be accounted for in terms of reasons. As such, it can be justified, or not. There need not be any moment of choice, where alternatives are clearly presented and an undetermined act of free will picks one of them. Still, the characteristic human action, caused (as Davidson (2001) would put it) by the events that constitute the desire and belief in question, is answerable, being intentional, to moral demands. It is this feature of answerability to reasons, rather than choice, that underlies moral egalitarianism, since it is shared by human beings, as well as the priority of goodness, not as a fundamental feature of reality but as a basic requirement of rationality.

The hallmark of moral dualism is the account of good and evil as antagonistic forces whose principle of operation is not made clear. In Kekes’s account, evil operates in us as forces expressing essential conditions of human existence. Kekes does not deny that the actions accounted for in terms of the forces he describes are also intentional and deliberate, and that in so far as they are to be understood in terms of their agents’ intentions, they can be rationalized in terms of the apparent good intended. He just thinks that the kind of explanation that can be given in those terms, prioritizing reason and goodness, is shallow; the tragic view, with the forces it diagnoses, goes beneath the surface of the reasons specified by deluded agents. But whether such an account is shallow depends on the further step it requires, namely an account of the failure, or dysfunction, that takes us from the apparent good intended to the evil perpetrated. Such an account may be given in terms of the tragic view and the forces it recognizes, but other explanations are possible which may go even deeper. Kekes supplies a rich framework for a dualist account, but ultimately, the forces he invokes—the essential conditions of life in their appearance as internal forces (vices) influencing

motivation and even gaining dominance in a person's character—are themselves inexplicable. What he offers is a (tragic) perspective on human life, which informs an ethics that could be debated, but not a satisfactory explanation.

6. Acting *Sub Specie Boni*

The rejection of moral dualism, namely, of evil as an intrinsic property of motives or dispositions on the part of evil agents, chimes with the doctrine known as “the guise of the good” doctrine. According to that doctrine, human motivation is tacitly evaluative. In so far as we act intentionally, i.e., from desire and relevant belief, we act *sub specie boni*; otherwise, our actions and our intentions do not cohere. The point of this paper, though, is more limited. It is that human action cannot be conceived as being performed *sub specie mali*, i.e., under the guise of evil. Action may fail to be fully intentional, leaving room for a variety of less than comprehensible outcomes. But that is different from being fully directed by evil with no admixture of an apparent good. Philosophers have disputed this connection between motivation and evaluation, so let me return to the arguments advanced by Velleman against the doctrine, based on which he opens the door to Satan's diabolical motives.

Velleman offers a critique of the traditional view, or rather a contemporary version (Davidson's) of that view, as well as an alternative account which concedes much to the traditional view but not enough, so he argues, to sustain it. His critique of the traditional account points to the transition it makes from the notion of desiring an object to that of holding an object desirable, i.e., having an evaluative attitude towards it. That transition comes at the point of bringing together a causal (motivational) view of action and a reason-based account. The causal and the rational account merge at the point of viewing the cause in question, namely the element of desire, as inherently involved in a value judgment that justifies the action in addition to describing its causation.

The latter part of the argument offers a different analysis of desire (and belief) in terms of their respective directions of fit and finds a disanalogy between belief and desire that renders the evaluative conception of desire too strong a conception of its direction of fit. While the direction of fit analysis serves to account for an “objective” sense in which the attitude of desire is evaluative, it does not suffice for what is needed, namely a “subjective” account of the value pursued in desire as mentally accessible. Desire's direction of fit does involve an evaluation of its object as *faciendum* rather than *factum*, i.e., as “to be brought about” rather than as already being there. But that does not supply any reason why it should be brought about, and desire is compatible with many different grounds for favouring its object, including (in the case of perversity) that there is no good reason for bringing the thing about, and (in the case of evildoing) that it is bad rather than good.

So let us look more carefully at these arguments. First, regarding Davidson's approach. Davidson reconciles the causal and the rational accounts of agency by

holding desire to be inherently evaluative. The difficulties with this approach arise in generally attributing evaluation to conation. On this account, to desire is to apply a concept of value to the object desired, but this makes it impossible to attribute desires to animals, or pre-linguistic children, who lack the conceptual apparatus required. Thus, if one makes the evaluative element of desire part of a graspable content of the desire in question, then one makes desire too intellectual to be satisfied by the non-linguistic. If, by contrast, one removes the evaluative component from desired propositional content, then the evaluative component is no longer accessible to the agent, no longer capable of subjectively justifying the action on top of describing its causality. The causal and the rational accounts are again decoupled.

But Velleman's critique may be too strong at this point. The agent may be justified in his action, given his desire and belief, without necessarily being able to justify himself explicitly. For example, a pre-linguistic child may desire her bottle of milk, and indeed give behavioural expression of her desire. While the baby does not grasp the milk bottle as desirable, there are other ways in which desirability in an object might manifest itself in the child's behaviour. The parent might say: "she hasn't had anything to eat for three hours, so she's right to make her demands." We need not view pre-linguistic babies as robots, answering only to a causal-motivational account. A reason-based account may fit not merely the fully articulate but also those of lesser capabilities. Similar descriptions could be given to animal behavior. A dog expects his master to take him on his daily walk and exhibits the expectation by jumping and barking excitedly. Relative to past interactions between dog and master, the dog is right to expect his daily walk and to communicate that to the master. The dog's desire is naturally expressed, as Davidson might put it, in an attitude that could be interpreted as expecting the daily exercise and giving expression to its desirability. The dog can't express it verbally. But others surely can.

Let me now move on to Velleman's positive account. As noted, Velleman proceeds by characterizing the attitude of desire's direction of fit as constituting an "objective" sense in which it could be seen as evaluative. The direction of fit of an attitude is the way it is supposed to fit its propositional object. In Velleman's words: "The propositional object of desire is regarded not as fact – not, that is, as *factum*, having been brought about – but rather as *faciendum*, to be brought about; it is regarded not as true but as to be made true" (Velleman, 1992: 8).

The propositional object of desire is, thus, to be brought about, or to be made true. But what is this status? Is it not at least one respect in which the propositional object is held to be good? Velleman concedes that it is but argues that the status of *faciendum* belongs not to the propositional object grasped by the agent but to the structure of the attitude. An attitude requires a qualifying predicate to account for its orientation to the object. To believe is to hold an attitude to the proposition in question "as true" (or the propositional object as *factum*); whereas to desire is to hold an attitude towards the proposition in question "as

good” (or the propositional object as *faciendum*), but that does not amount to a judgment, mentally accessible to the agent, regarding the propositional object as good. The evaluative component has been shifted from the propositional content of the attitude to its “objective” status as the kind of attitude it is. As such it is insufficient for the purposes of justifying the action so motivated.

The point is made by Velleman in the following way:

I concede that desire can justify action objectively, by making true a proposition that could guide one’s actions if one gained appropriate access to it; but I deny that desire justifies action subjectively, by constituting an evaluative attitude whose justificatory force is already available to guide one’s actions (Velleman, 1992: 10).

One might think that the objective sense of justification, rooted as it is in the attitude’s direction of fit, goes a long way to account for the intuition behind the evaluative account of desire. Maybe the subjective action-guiding force expected by Velleman is not necessary. It could be said that grounding the evaluative component in the orientation of the attitude rather than its propositional content gives the cognitivist all that is needed. Against this, Velleman expands his analysis and points to relevant differences between different cognitive and conative attitudes, not just belief and desire. Thus, on the side of cognitive attitudes, belief is to be distinguished from other attitudes such as accepting or fantasizing. On these attitudes, the direction of fit is still the same. The attitude consists of regarding the proposition in question as true, but unlike belief that does not involve regarding the proposition as true for the right reason. In belief we regard the attitudinal proposition as true aiming thereby to track the truth, while we may “accept” a proposition as true, e.g., for the sake of argument, regardless of its actual truth. In Velleman’s language: “Belief not only has truth as its constitutive predicate but also has correctness in matters of truth-or, as we might put it the ‘real truth’—as its constitutive aim.” (1992: 14).

Similarly, other conative attitudes also regard their propositional object as *facienda*, wishing, hoping etc. And it might be argued that desire is analogous to belief, in its mode of approval, in so far as desire involves not only regarding its propositional object as to be brought about, but also with the aim of tracking what is really to be brought about. Such an account, Velleman concedes, would indeed have justificatory potential, as described in Davidson’s theory of rational guidance. But Velleman does not think the account is merited, and it is here that his account seems most arbitrary. Velleman argues that desire, properly speaking, does not track the good (the real *faciendum*) but only the attainable. As grounds for this, he claims that we cannot desire the unattainable (or the already attained). And since desire tracks the attainable it does not track the real *faciendum* and cannot serve the purposes of justification claimed by the cognitivists. But the point is debatable. Is it an a priori truth that desire tracks the attainable, or that any attitude that tracks the unattainable does not qualify as a desire? Unattainability clearly frustrates desire, but it does not eliminate it.

Of course, tracking the “real” *faciendum* should be taken with a grain of salt. Desire tracks desirability, and in this way it is an evaluative attitude, only *ceteris paribus*, not all things considered. One may, of course, desire things that are, overall, bad, or unworthy. But in desiring them there must be something appealing in them, otherwise the desire in question remains unintelligible, and so does the agent. That is the point of the phrase “the guise of the good,” namely, that appearances may be misleading. Indeed, the sense in which desire is evaluative is that it tracks *facienda* that are, in one way or another, attractive, or desirable, not that they are so all things considered. And the tracking of such *facienda* does not amount to a full justification of the ensuing action on the part of the agent, but the much weaker condition of making the agent intelligible, even if ultimately unjustified.

Velleman takes as a point in favour of his account that if true, the account may incorporate the possibility of desires with no justifying force. He says:

That desire doesn’t aim at correctness explains why desire can be perverse. ...one can often desire things conceived as worthless, or even bad, and desire them precisely under those descriptions. (1992: 17).

But that is the point at issue. One can, indeed, desire what is worthless, evil, or perverse. But can one desire them “precisely under those descriptions”? For it might be argued that under those descriptions, the desired things in question would not constitute *facienda* any more than they would constitute subjectively recognizable goods. They would not intelligibly be considered as “things to be brought about.” To the extent that they are of that status, they must be sought after under some different description, some “guise” which would render the agent minimally intelligible.

Assuming that there are perverse desires, namely, desires formed in such moods as despair or self-destructiveness, Velleman concludes that such desires cannot be accounted for in terms of some justifying guise. If the object is desired because it is undesirable the desire in question cannot be ultimately aiming at the desirable. As he puts it: “correctness of approval simply cannot be one’s aim when one approves of something under the description that it is unworthy of approval.” (1992: 18). It follows that desire cannot be accounted for in terms of the good, or correctness of approval regarding a *faciendum*, an object which is to be brought about. But the question remains open whether such desires are indeed conceivable. Can one desire an object under the description of it as undesirable? Can such an object constitute any person’s *faciendum*? The mood of self-destruction generates desires for harm, and that surely is a familiar phenomenon. But does the self-destructive person desire harm just because it is harm, i.e., for its own sake and precisely under that description? That could be questioned. The self-destructive need not be conceived as so thoroughly irrational as to defy understanding. A more plausible description would be that they want what is harmful under some (mis) description of it as desirable, and the issue

that calls for account is how the split in their motives occurs, how they come to desire what in fact is undesirable.

In short, Velleman does provide a sense in which desire is evaluative, though it need not be subjectively available, and fails to provide a sense in which the undesirable may become the object of desire, the *faciendum*, under that description. A moral dualism cannot be supported on this basis. The followers of Satan, if they are at all coherent, cannot pursue evil “precisely under that description,” and like the rest of us they tacitly evaluate the object of their desire, that which is to be brought about, as good, though it may in fact be extremely evil.

Let me sum up by saying that the two forms of moral dualism discussed in this paper fail to support a morally based refusal to render evil perpetrators intelligible. The idea that evil cannot be “seen” for what it is, or that it cannot be adequately confronted, unless its principled, though inexplicable character is acknowledged has not been established, and it leads to a moral failing of its own. In both its forms, moral dualism bifurcates humanity into two moral classes, undermining the moral principle of human equality (moral egalitarianism) in matters of dignity and responsibility. The moral risk involved in such dualistic attitudes is the risk of failing to identify moral evil, as well as goodness, as rooted in what is common to all of us as human beings. On the contrary, Evil is rooted in human failings that we all share; it is not confined to those of us who are devilishly irrational or ill-disposed. In confronting evil, we should not look to others; it lies within our shared humanity.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Anglin, B., & Goetz, S. (1982). Evil Is Privation. *International Journal for Philosophy of Religion*, 13, 3-12. <https://doi.org/10.1007/BF00148934>
- Anscombe, G. E. M. (1963). *Intention*. Cornell University Press.
- Aquinas, T. (1975). *Summa Contra Gentiles*. University of Notre Dame Press. <https://doi.org/10.2307/j.ctvpj7fpk>
- Aquinas, T. (1995). *On Evil. Question 5, Article 3* (Translated by John Oesterle). Notre Dame. <https://doi.org/10.2307/j.ctvpj75ww>
- Arendt, H. (1951). *The Origins of Totalitarianism*. Meridian Books.
- Arendt, H. (1963). *Eichmann in Jerusalem*. The Viking Press.
- Arendt, H. (1978). *The Jew as Pariah*. The Grove Press.
- Augustine (1952). *The Confessions*. William Benton, Publisher, Encyclopedia Britannica, Inc.
- Augustine (2012). *Enchiridion*. Simon and Schuster.
- Bauman, Z. (1989). *Modernity and the Holocaust*. Polity Press.
- Bernstein, R. J. (2002). *Radical Evil*. Polity Press in Association with Blackwell Publishers.

- Burkitt, F. C. (1922). The Religion of the Manichees. *The Journal of Religion*, 2, 263-276. <https://doi.org/10.1086/480284>
- Calder, T. C. (2007). Is the Privation Theory of Evil Dead? *American Philosophical Quarterly*, 44, 371-381.
- Davidson, D. (2001). Mental Events. In *Essays on Actions and Events* (pp. 207-225). Clarendon Press. <https://doi.org/10.1093/0199246270.003.0011>
- Gardner, S. (2003). Tragedy, Morality and Metaphysics. In J. L. Bermudez, & S. Gardner (Eds.), *Art and Morality* (pp. 218-259). Routledge.
- Hacker, P. M. S. (2021). *The Moral Powers: A Study of Human Nature*. Wiley Blackwell. <https://doi.org/10.1002/9781119657828>
- Kane, G. S. (1980). Evil and Privation. *International Journal for Philosophy of Religion*, 11, 43-58. <https://doi.org/10.1007/BF00138764>
- Kant, I. (1996). Religion within the Boundaries of Mere Reason. In A. W. Wood, & G. Di Giovanni (Eds.), *Religion and Rational Theology* (pp. 39-215). Cambridge University Press.
- Kekes, J. (1990). *Facing Evil*. Princeton University Press. <https://doi.org/10.1515/9780691217963>
- Milgram, S. (1963). Behavioral Study of Obedience. *The Journal of Abnormal and Social Psychology*, 67, 371-378. <https://doi.org/10.1037/h0040525>
- Milton, J. (1984). *Paradise Lost*. Lerner Publishing Group.
- Neiman, S. (2002). *Evil in Modern Thought*. Princeton University Press.
- Silber, J. R. (2012). Appendix: Kant at Auschwitz. In J. Silber, (Ed.), *Kant's Ethics: The Good, Freedom, and the Will* (pp. 314-345). Walter de Gruyter. <https://doi.org/10.1515/9781614510741.314>
- Velleman, D. J. (1992). The Guise of the Good. *Nous*, 26, 3-26. <https://doi.org/10.2307/2215684>