C36

CHAPTER 36

KANT ON EVIL

MELISSA MERRITT

C36S1

1 INTRODUCTION

C36P1 The titular aim of this chapter is expressed in a conventional, but nevertheless somewhat misleading, way. So we should begin by heading off misguided expectations about our topic. One source of potential confusion stems from the fact that the English word *evil* is cognate with the German *Übel*, which Kant pointedly distinguishes from *Böse* (CPrR 5: 59–60)—*Übel* is something disagreeable in our physical condition, whereas *Böse* is something bad in our person or moral condition.[1] So, while some philosophical discussions about 'the problem of evil' might express consternation about how 'bad things'—earthquakes, cancer—can happen to 'good people', for Kant the topic is exclusively to do with imputable badness, i.e. *Böse*, not *Übel*. The other source of potential confusion is that 'evil' in contemporary English generally carries the sense of extreme moral badness or depravity. But Kant is not especially fascinated by such extremity, and famously denies that a human being could be 'diabolically' evil, bent on pursuing evil for evil's sake (Rel. 6: 35). For Kant, evil is perfectly ordinary, since it is endemic in human nature.[2]

C36P2 Our topic, then, is the 'radical evil in human nature'—Kant's thesis about the human moral condition that he develops primarily in his 1793 *Religion within the Boundaries of Mere Reason*.[3] According to this thesis, the human moral condition is corrupt by default and yet by our own deed; and this corruption is the origin (root, *radix*) of human badness in all its variety, banality, and ubiquity. Since this corruption is 'detectable as early as the first manifestation of the exercise of freedom in the human being' (Rel. 6: 38), it is understood to be coeval with coming into the use of one's reason, and acquiring the resources of a genuinely practical point of view. The implication is jolting and bleak: just as soon as one counts as the imputable source of one's own actions, one brings this corruption upon oneself, inevitably but culpably.

C36P3 Now, for Kant, intentional actions are expressions of commitment to practical principles or 'maxims'—roughly, views about what is a reason for doing what—and

rational agency is a kind of causality. In the *Religion*, as elsewhere, Kant supposes that we are endowed with basic incentives both of self-love and morality; and once we come into the use of our reason, we form principles to act on these incentives. But the metaphysics of agential causality—that a person is the imputable source of his own actions—requires there to be some overarching commitment that characterizes the causality of his will.[4] Kant then takes radical evil to be our default commitment to treat the principle of self-love as a condition of compliance with the moral law, rather than the other way around (Rel. 6: 36). This commitment is itself imputable and thus must be understood as something we *do*: it is the *peccatum originarium* as opposed to any of the myriad manifestations of badness that draw from it, *peccata derivativa* (Rel. 6: 31).

C36P4 *Peccatum originarium* is Augustine's term for original sin; and Kant's one use of this phrase is the closest he comes to referring to that doctrine explicitly in this context. As with any topic in Kant scholarship, controversy reigns—even on a point so fundamental as whether radical evil is equivalent to, or some critical rehabilitation of, the Augustinian doctrine. To give an idea of the range: where Allen Wood takes Kant's use of the phrase *peccatum originarium* to show that he 'does explicitly equate' radical evil with original sin (2014: 32n.1), Paul Guyer regards the *Religion* as 'a work of the radical Enlightenment, a deconstruction of Christianity', and treats radical evil as an alternative to—but certainly no rehabilitation of—the doctrine of original sin (2009: 144–45). Moreover, Kant indicates his rejection of central points of the Augustinian doctrine, such as that we have passively inherited the sin of Adam (Rel. 6: 40). Nevertheless, commentators since Goethe have been inclined to interpret radical evil in an Augustinian light, and then wrestle with the awkwardness of how this thesis could be integrated into Kant's ethics of autonomy.[5] Continuing in this vein, Kant is then charged with deep ambivalence about his own philosophical project that he is deemed powerless to resolve.[6]

C36P5 I am not going to pursue the question of Kant's relation to Augustine here;[7] but I do want to draw an initial lesson from our sidelong glance at this swath of the secondary literature—particularly on the expression of puzzlement about the coherence of radical evil with the 'critical' works in ethical theory, such as the *Groundwork* and *Critique of Practical Reason*. It is certainly appropriate for commentators to seek interpretive coherence across texts; nevertheless, I think that a considerable share of confusion arises from the assumption that Kant's account of radical evil is a direct outgrowth of the ethical arguments of the *Groundwork* and the second *Critique*.[8] This supposition should be rejected, since—as I will argue—the account of radical evil belongs to an argument that works in a different way, for different ends. Among Kant's chief aims in the critical ethical works is to make a case for the unconditional bindingness of the moral law, which he achieves by proceeding from an examination of the nature of rational agency as such. Only once this case is made does Kant look to incorporate creaturely facts about the human being into his argument, such as our endowment for moral feeling.[9] We could begin to register that matters might be quite different in the *Religion* by tracking the constant refrain about 'the human being' running throughout. Radical evil is a thesis about the *human* moral condition—a condition that is not strictly necessary, since in some important sense it can and should be otherwise.[10]

C36P6 One way to understand this last point is to recognize that Kant's thesis about radical evil belongs to a larger argument about human development. This teleological argument, I will argue, draws on distinctively Stoic ways of thinking about 'the human being'. Commentators have not recognized this Stoic background, and as a result have misunderstood the nature of Kant's thesis about radical evil and its argumentative context. Thus I begin by showing how the interpretive landscape has formed along two fault lines in recent decades: along one, commentators dispute the nature and justification of Kant's thesis about radical evil; along the other, whether radical evil has fundamentally to do with human psychology or human sociality. My own position, which is developed in the remainder, takes sides on the first contest, and reveals the second to be spurious: the Stoic background helps us understand that radical evil has to do with human psychology and human sociality in equal measure, which casts fresh light on its ethical significance for Kant.

C36S2

2 TWO FAULT LINES

C36P7 Kant explicates radical evil as 'the propensity to evil in human nature' ('*der Hang zum Bösen*', 6: 28) in the second section (§II) of Part One of the *Religion*.[11] As a result, commentators have tended to fix their attention on that part of Kant's account, treating the real beginning of the argument as something of an afterthought. But Kant announces his premise in the first section (§I), 'the original predisposition to good in human nature' (*ursprüngliche Anlage zum Guten*, 6: 26). With this, Kant signals that he assumes a conception of providence: namely, we are endowed with everything that we need to make ourselves good.[12] We will examine this conception of the good and this invocation of providence later. For now we just need to fix on an essential point of contrast: the predisposition to good is *original—laid* in us in the way we are created—whereas the propensity to evil, Kant claims, is acquired.

C36P8 Commentators widely recognize this point of contrast, but sometimes suspect it to be a distinction without a difference.[13] Kant explains that by 'propensity' he means 'the subjective ground of the possibility of an inclination' (Rel. 6: 29). So a propensity is some kind of underlying readiness to be inclined or attracted along certain lines, rather than the inclination itself. Kant recognizes that a predisposition might be thought of in similar ways, and struggles a bit to mark the distinction that he wants. In a footnote, he offers an example along the following lines: someone who has had no exposure to or knowledge of the effects of alcohol could not be said to have any inclination for alcoholic drinks; but such a person could still have a *propensity* for this inclination, so that she accordingly develops the inclination once she is so exposed.[14] Conceptually, he is trying to distinguish the underlying readiness to be attracted along certain lines from the myriad ways a person might go on to be actually attracted: perhaps to gin, but not beer. That much fits the idea of a propensity to evil, which Kant conceives as a kind of underlying readiness to be attracted to badness—in effect, to attach oneself to one's own faults or vices. This propensity can be distinguished from whatever may turn out to be

the particular faults or vices at issue, which are likewise found in such flourishing variety among us. But in another way the example does not suit the idea of a propensity to evil: for the propensity it describes is naturally understood to be a given fact about a person's make-up—*not* the result of anything she does, and thus not anything imputable.[15] Kant elaborates that he takes a propensity to be the sort of thing that can either be given or acquired; and the propensity to evil is certainly acquired—namely, '*brought* by the human being *upon* himself' (Rel. 6: 29).

C36P9 Of course, if we are freely self-determining creatures, goodness or badness alike could only ever be by our own doing. The question, then, is how Kant could justify the idea that we make ourselves bad *by default—that* this is the inevitable point from which we set out as persons, and from which we try to become good. Kant's remarks about how to understand this are among the most difficult in his whole account.

C36P10 The remarks I have in mind are found in the third section (§III), titled: 'The human being is by nature evil' (Rel. 6: 32). Call this *the old saying*. Kant endorses it as an expression of the thesis of radical evil, provided that it is correctly understood. The remarks at issue concern how it should and should not be understood. Kant first explains that the old saying 'simply means that being evil applies to him considered in his species' (Rel. 6: 32). I interpret this to mean that radical evil is a species characteristic, the significance of which we will consider further in Section 3. For now it is enough to recognize that such a claim should in principle accommodate the possibility that not all human beings have this propensity—perhaps from somewhat similar considerations as it can be admitted that not all human beings are bipedal. Most commentators, however, take Kant to be asserting the universal proposition that all human beings have the propensity to evil. Since universality, necessity, and a priority stand together in Kantian thought, these commentators assume that Kant thinks this propensity is *necessary* in human beings: thus Morgan (2005: 64), for example, follows this line of thought, citing *Religion* (6: 32). But his assumption that the propensity is necessary should seem strange when we look further at the text of *Religion* (6: 32). Kant next tells us what the old saying *does not* mean: namely, that 'this quality may be inferred from the concept of his species . . . for then the quality would be necessary' (6: 32). Rather than speaking as if this propensity is obviously necessary in human beings, as Morgan presents it, Kant instead speaks as if this propensity is *clearly not* necessary. Kant does not elaborate on his thinking here: but in this context, at the very least, it means that the old saying is not an analytic proposition. Commentators who take Kant to be asserting a universal proposition—all human beings have the propensity to evil—then suppose it to be a synthetic a priori claim that is 'crying out for a transcendental deduction that he does not provide' (Morgan 2005: 65). Their work is then cut out to reconstruct the missing transcendental justification on Kant's behalf—thus I label these 'transcendental' approaches to radical evil.[16]

C36P11 However, when Kant denies that the old saying is analytic, he does not say that it is synthetic a priori instead. The old saying, Kant continues, 'rather' means

C36P12

> that the human being, as one is acquainted with him through experience, cannot be judged otherwise, or [that] one can presuppose it [sc. the propensity to evil] as

> subjectively necessary in every human being, even the best. (Rel. 6: 32, translation modified)

C36P13 While Kant first rules out taking the old saying to mean that radical evil is necessary in human beings, he follows by saying that it may be presupposed as '*subjectively* necessary in every human being' (my emphasis). A natural way to take the upshot of this remark is that, from some 'subjective' perspective of human choice, we unavoidably corrupt ourselves in this way. However, it should be obvious that this corruption could not arise as a matter of natural-causal necessity: for we freely bring it upon ourselves. It should also be obvious that we do not corrupt ourselves on grounds of practical necessity: for this corruption must be some deviation from the normative standard on which all claims of practical necessity rest by Kant's lights. Yet there is scope for the thought that the human being inevitably corrupts himself in the course of his development: this is a regularity belonging to a natural progression, and (for all we know) may not have ever admitted any exception. But it is not strictly necessary—thus a fortiori the old saying cries out for no transcendental deduction.

C36P14 These considerations form the basis of a competing approach, which I label 'teleological': its governing principle is that Kant's thesis about radical evil belongs to an account of human development. Developing this approach, David Sussman suggests that '[t]he peculiar necessity that Kant associates with radical evil is to be found in the special process that human beings must undergo if they are to acquire their characteristic powers of practical reason' (2005: 154). Sussman's thought is that the human being must do certain things owing to the developmental imperative of completing (i.e. fully realizing or perfecting) its essentially rational nature in virtue; and *as human beings*, we can do these things only by working through radical evil. But this does not make radical evil a necessary quality of human beings (see again Rel. 6: 32). Rather, Kant supposes, a certain amount can be said about what human beings are simply 'like'—enough to explain how we inevitably corrupt ourselves at a certain point in our development. While we have warrant to suppose that the human being is *like that* (i.e. *such as* to inevitably corrupt itself at a certain point in its development), we have no warrant to suppose that the human being is *necessarily* like that, rather than some other way. Moreover, as I noted, Kant's argument takes as its premise the 'original predisposition to good': hence there must be a developmental imperative to overcome radical evil. From these considerations, the difficult passage about the old saying looks far less puzzling: its upshot is that our inevitably self-corrupted moral condition is not necessary, since it can and should be otherwise. We in principle have everything we need to overcome it.

C36P15 The divide between 'transcendental' and 'teleological' approaches to radical evil stems from competing views about the modality of the old saying. Another division concerns whether Kant's thesis about radical evil is chiefly concerned with human psychology or human sociality. One example of this difference can be marked within teleological approaches: while Sussman, as noted, considers radical evil in the context of the purposive development of individuals, Sharon Anderson-Gold (2001) argues that radical evil principally belongs to Kant's account of the purposive development of

humanity through history—and thus chiefly concerns the development of the *species*. Although these are not mutually exclusive approaches, they involve importantly different commitments about what is central versus penumbral to Kant's account of radical evil. Notably, in *Religion* 1 Kant explains radical evil in resolutely psychological terms as a corruption that *each* invariably brings upon *himself*; but at the outset of *Religion* 3, he speaks of radical evil as a corruption that we collectively inflict upon *each other* in society. Claiming the centrality of the latter perspective, Anderson-Gold (2001) argues that radical evil has an essentially social *context*, while Wood (1999) makes the stronger claim that it has an essentially social *origin*. Commentators working in this vein (also Rossi 2005) find a powerful corrective to assumptions about 'individualism' in Kant's ethics. The rejoinder from moral-psychologically oriented commentators is that the imputability of radical evil requires an explanation in terms of individual psychology.[17]

C36P16 The remainder of this chapter develops a particular variant of the teleological approach to radical evil. For Kant, radical evil belongs to an account of human development that draws on distinctively *Stoic* ways of thinking about teleology and 'the human being' (Section 3). Among other things, this perspective shows why there is no real trade-off between psychological and social accounts of radical evil (Section 4)—and why this result is significant for Kant (Section 5).

C36S3

3 RADICAL EVIL AND HUMAN DEVELOPMENT IN *RELIGION* 1

C36P17 Radical evil is a thesis about the human moral condition, which for Kant belongs to a teleological way of thinking about 'the human being'. More particularly, it belongs to an account of human development that draws from a Stoic precedent. My aim here is to outline the key features of this Stoic precedent, and then explain what it tells us about Kant's argument in *Religion* 1.

C36P18 The Stoic precedent is the theory of *oikeiōsis*, described by one prominent commentator as nothing short of 'their novel theory of the basic state of affairs which grounds all human and animal action' (Inwood 1985: 184).[18] While we are interested in the specification of this theory for the human being, to understand it correctly we need to start with the general picture. And since we are effectively dealing with the Stoic view of what it is to be an animal, we can make our approach by observing first how Stoics distinguish animals from plants: while both animals and plants have natures, only animals are ensouled, and thereby alive.[19] As a result, an animal is most fundamentally conceived as an agent: an animal is the sort of creature that has to *do* certain things—act appropriately—if it is to complete (i.e. fully realize or perfect) its nature. This will prove an important point for the specification of the thesis: the Stoic idea of appropriate action (*kathēkon*, *officium*) is fundamentally teleological. To act appropriately is to act in

completion-promoting ways, i.e. in ways that enable it to fully realize its nature as a creature of a certain kind.

C36P19 The particulars of the theory of *oikeiōsis* concern how such action is possible.[20] Seneca argues that it is only possible if the animal has some kind of sense of its 'own constitution', which of course is not any articulate, theoretical conception of it; rather, it is a felt self-awareness that is manifest in the concrete practical know-how of appropriate action, situation by situation and over the course of a life. Appropriate action is possible because the animal has an affinity for, or sense of kinship with, the objects and actions that preserve it in its own constitution.[21] Hence the term *oikeiōsis* borrows a standard term for expressing relations of individuals and property within a household, and extends it to express the thought that animal action is oriented by a sense of what is properly one's own. Finally, since teleology and theology both belong to the Stoic study of nature or 'physics', the animal is conceived as a *creature* that is looked after in being set up this way: 'Nature brings forth her offspring, she does not toss them aside', as Seneca memorably puts it (Epistulae 121.18). Animals are providentially endowed with somatic and psychological predispositions that compel appropriate action.

C36P20 In this way, on the Stoic theory, divine nature (or Nature) leads animals to complete (fully realise or perfect) their natures. This is also true for the human being, going back to Zeno: 'nature leads us' to the telos of virtue (Diogenes Laertius, 7.87). Two basic points about the specifics of human *oikeiōsis* need to be made here: one concerns the conception of this telos as *good*, the other the complications of our being led to it. The first point is relatively straightforward, provided we take on board the Stoic conception of the cosmos as a rationally governed whole. While non-rational animals can only be complete according to their own kind, Seneca explains, the rational animal is called to be 'complete in accordance with the nature of the cosmos' (Epistulae 124.14); on this basis, the perfection of our essentially rational nature is taken to be genuinely, or unqualifiedly, *good* (Epistulae 124.23). We are not created good, but rather endowed with everything we need in order to make ourselves so.

C36P21 The idea that Nature leads us to complete our essentially rational nature is complicated by the Stoic understanding of what it is involved with the acquisition of reason.[22] To acquire the resources of reason is to come into a rational point of view, so that one's actions are then expressions of assent to evaluative propositions. Once this happens, we can no longer be directly compelled by our predispositions to act in the appropriate, nature-completing ways: now we can only act on our own initiative.[23] Nature may lead us to virtue, but only by endowing us with uncorrupted starting points (Diogenes Laertius 7.89). In the words of Seneca, 'nature does not predispose us to any fault' (*nulli vitio conciliat*, Epistulae 94.56)—but predisposes us to the good. If we do become good, it can only be by our own doing. By the same providential dispensation, any badness can only be own work, as well.

C36P22 Moreover, according to a report of Alexander of Aphrodisias (1991: 122.3–5), Stoics suppose that we invariably make ourselves bad at the very moment we acquire the resources of reason, and so are no longer children.[24] It is easy to see that we wouldn't invariably make ourselves good at this point: goodness is the goal of our development,

which we can only reach if we acquire a conception of the good that we would not yet have.[25] But why should we invariably make ourselves bad, by Stoic lights? The answer lies in the considerations just adduced. As we acquire the resources of reason, our actions are now expressions of assent to evaluative propositions, to views about what is good and bad; and these views are bound to be mistaken before we have acquired a genuine understanding of the good. To be sure, it has been debated since antiquity how Stoics can really explain human badness, given their assumptions about providence. While we cannot take up those debates here,[26] we can draw attention to a simpler fact: Stoic accounts of human *oikeiōsis—such* as Cato's in Cicero's *De Finibus—do* not consider human corruption.[27] Cato indirectly justifies this omission when he explains that he is concerned only with what follows from 'natural principles' (Finibus 3.20). Human corruption is the perversion of those very principles, and thus lies somewhere outside the story of how nature leads us to virtue.

C36P23 The evidence of this Stoic background is apparent from the beginning of *Religion* 1, starting with the preamble to the main argument (6: 19–20). Kant considers an optimism about human moral progress that he associates with Enlightenment pedagogues such as Rousseau, and traces to Seneca: the world is moving 'from bad to better'—or, if it should be allowed that this is not immediately apparent, then 'at least there is in the human being the predisposition to move in this direction' (Rel. 6:20). Kant does not reject the very idea that we are providentially endowed with such a predisposition, but rather that it provides sufficient ground for the optimism of this tradition. We can begin to appreciate this point if we take a look at the line of Seneca with which Kant concludes this passage:

C36P24

> Sanabilibus aegrotamus malis, nosque *in rectum genitos* natura, si sanari velimus, adiuvat. [We are sick with curable ills, and if we are willing to be restored to health, nature lends a hand, for we are *born to be upright*]. (Seneca, *De Ira* 2.13.1, slightly misquoted at Rel. 6: 20)[28]

C36P25 The providential assumption is expressed in the image of nature lending us a hand, guiding us towards our telos of virtue, or 'uprightness'.[29] Progress is ours, Seneca intimates, if we are only 'willing to be restored to health'. But Seneca has overlooked a genuine difficulty, Kant thinks: the human being is not normally so willing.

C36P26 When we turn to Kant's own argument, we see that he never asks *whether* there is a 'predisposition to good in human nature'; he rather assumes that there is, and immediately sets out an explication of this idea. The predisposition has three aspects which can be seen as 'elements of the determination of the human being [*Elemente der Bestimmung des Menschen*]' (Rel. 6: 26). Kant alludes here to an eighteenth-century debate about the vocation or destiny (*Bestimmung*) of the species, which he takes to consist in the perfection of our essentially rational nature (Anth. 7: 321).[30] From this we can see that 'good' is understood teleologically, just as it is in Seneca's account of human *oikeiōsis*. To say that the predisposition is 'original' is to say that it is laid in us in the way we are created: we are providentially endowed to feel affinity for those objects and actions that preserve us

in our constitution, and are thereby oriented towards the completion of our essentially rational nature. Two further points about Kant's conception of the human *Bestimmung* are important for his account of human development in the *Religion*. One draws from the Stoic recognition that our appropriate action cannot be directly compelled by our predispositions: as Kant puts it, 'the human being' can complete his essentially rational nature only by acting on his own initiative, 'according to ends he himself adopts' (Anth. 7: 321). The second point does not have a Stoic antecedent, however. Kant contends that this completion can be reached only by the species, not the individual (Anth. 7: 333; UH 8: 19).

C36P27 At this point we should revisit the aim, announced in the title of *Religion* 1, to give an account of 'the radical evil in human nature' (6: 18). Why should the 'original predisposition to good' come first in this account? The answer is that radical evil is the corruption *of this predisposition* (Rel. 6: 43). Now, the Stoic background reveals that the original predisposition is conceived as the basis of appropriate action: it concerns our affinity for the objects and actions that, as a Stoic would say, preserve us in our constitution. Kant breaks this down into the affinities that guide us to act appropriately (1) in our animality, as living beings; and then (2) in our humanity, as creatures capable of freely setting ends and acting on them;[31] and finally (3) in our personality, as morally imputable agents (Rel. 6: 26). The idea of affinity or kinship is reflected in the details of these elements. The first is an affinity for what preserves one as a living being, and which does not as such require the resources of rationality. While Kant puts this under the heading of 'physical … self-love' (Rel. 6: 26), it is not a narrowly conceived self-care: it includes not only the sexual drive, but also to care for our offspring, and to join in community with other human beings (Rel. 6: 26). The second is also presented as a 'physical' self-love, but one that does require the resources of rationality, since it consists in taking a developmentally appropriate interest in one's own agency. It can thus be understood as an affinity for those ways of acting that develop the powers of our own agency, which Kant takes to require comparison and competition in a social context (Rel. 6: 27). The third element is our inborn readiness to be moved simply by our recognition of what morality requires (Rel. 6: 27–28), or moral feeling, and can thus be understood as an affinity for such actions as are appropriate to one's status as an imputable person.

C36P28 Radical evil, then, is the corruption of this entire package. To consider what this means, we first need to take account of a further feature of Kant's account of the original predisposition to good. Kant devotes at least as much space to the kinds of vices that 'can be grafted' onto the first two elements of this package (Rel. 6: 26, 27) as he does to those elements themselves.[32] Now, a graft takes the energy of the rootstock and turns it to some other end—as, for example, the developmentally appropriate interest in the powers of one's own agency may find twisted expression in the envy and Schadenfreude that Kant mentions in this context. The grafting analogy in this context can be spelled out as follows. Radical evil is the preparation of the rootstock (the original predisposition to good): it is a corruption of that original predisposition, so that it can take certain grafts, i.e. unnatural or inappropriate attachments (particular vices). Radical evil is thereby conceived as a fundamental corruption of our sense of what is appropriate to us

as human beings; and particular vices (which include 'the *passions*' understood as 'malignant inclinations', see Rel. 6: 93) are the myriad ways this perversion can take shape.

C36P29 Moreover, the fact that Kant considers grafted vices in this context—i.e. while expositing the original predisposition to good—is significant against the background of prominent Stoic accounts of *oikeiōsis*, such as that of Cicero's Cato, that omit consideration of our corruption. This omission can be justified, I noted, on the ground that our corruption does not flow from 'natural principles'. Indeed, the metaphor of the graft is Kant's way of nodding to this point, as is his distinction between the *predisposition* to good (an endowment laid in us in the way we are created) and the *propensity* to evil (which we bring upon ourselves). However, Kant also speaks freely of both as lying 'in human nature' (Rel. 6: 26, 28); and we may be likewise puzzled by his repeatedly speaking of radical evil as 'innate' (Rel. 6: 32, 38, 42–43). But if radical evil can be traced to 'the first manifestation of the exercise of freedom in the human being' (Rel. 6: 38), then we cannot be evil from the cradle, just as we do not passively inherit the sin of Adam (Rel. 6: 40). In all these passages, Kant indicates that 'innate' is just shorthand for 'lies in human nature'—and *that* claim will be understood in terms of its being inevitably brought upon ourselves in the course of our development.[33] For Kant, contra Cato, any account of human development that passes over the fact of our corruption cannot really be an account of *the human being*.

C36P30 Radical evil must then be understood in terms of the practical commitments of a human being who has come into the use of her reason. (Pre-rational human beings, we may suppose, act perfectly appropriately, guided by the predisposition to animality.) The two basic incentives on human action, self-love and morality, are attributed to the original predisposition (Rel. 6: 36); imputable agents act on practical principles that incorporate these incentives. And, as noted above (Section 1), there must be some overarching practical commitment that characterizes the causality of an imputable agent's will. Kant then explains radical evil as the freely undertaken commitment to treat self-love as the condition that must be satisfied in any action complying with the moral law (Rel. 6: 36). Although he claims that we cannot explain *why* we make this choice (Rel. 6: 32, 43), its inevitability can nevertheless be appreciated from a developmental perspective. For while the incentives of self-love and morality are laid in us in the way we are created, we only form principles to act on these incentives once we come into the use of our reason. At this point, any human being will have plenty of practice acting—pre-rationally—from the incentive of self-love. But no human being will have had practice acting on the incentive of morality, since by Kant's lights this is not possible at all before the resources of self-conscious rationality are at hand. The human being who has just come into the use of her reason is not developmentally prepared to choose in any other way than to bring on the propensity to evil, at least as Kant understands this.[34]

C36P31 In this way, radical evil corrupts the agent's entire evaluative outlook: it 'corrupts the ground of all maxims' (Rel. 6: 37).[35] Thus a person cannot overcome this corruption by tinkering piecemeal with elements of this evaluative outlook; she can only entirely overthrow it, in an equally incomprehensible act of free choice (Rel. 6: 44–45), in a move Kant deems a 'revolution', and links to the sort of spiritual rebirth invoked in biblical

sources (Rel. 6: 47).[36] But this point should be linked to the Stoic teleological background of Kant's argument, as well. Since radical evil is the corruption of the original predisposition to good, it follows that everything we do while thus corrupted is inappropriate (*zweckwidrig*): completion-hindering rather than completion-promoting. Only when we overthrow radical evil—only when we 'restore' the original predisposition 'to its power' (Rel. 6: 44)—will we stand to act appropriately, in completion-promoting ways. Only then, as Kant puts it, can a human being be 'a subject receptive to the good' (Rel. 6: 48). This is a point about engaged *attraction*, which Kant encourages his audience to read in a Christian light as a 'change of heart' (Rel. 6: 47), but which can equally, again, be read in the light of Stoic *oikeiōsis*, as a transformation in what one regards as most dear, most properly 'one's own'.

C36P32 My aim in this section was to set out a particular version of a teleological account of radical evil. Generally, a teleological approach allows us to consider radical evil as inevitable in the course of the development of 'the human being'—but not strictly necessary. We now need to look further into that account of human development, which, as we are about to see, does not concern 'the human being' simply as an individual, but also as a *species*. Hence the psychological account of radical evil in *Religion* 1 presents only one side of the story about the inevitability of radical evil in 'the human being'.

C36S4

4 RADICAL EVIL AND HUMAN SOCIALITY IN *RELIGION* 3

C36P33 *Religion* 1 offers a psychological account of radical evil as the corruption of the original predisposition to good, and concludes with remarks on the possibility of overthrowing this corruption, or 'restoring' the original predisposition 'to its power'. In this account, radical evil is a corruption that *we each* bring upon ourselves. Indeed, since both this corruption and its overthrowing are possible only in acts of free choice, the *Religion* 1 account appears to be entirely concerned with the doings of individuals.

C36P34 But this impression is challenged at the start of *Religion* 3, where we learn that the 'morally well-disposed human being'—the post-revolution hero of *Religion* 1—labours under constant assault from 'the evil principle' (6: 93). This 'perilous state' appears to be nothing other than ordinary human society, wherein we corrupt *one another*, without trying or intending, but simply by existing alongside one another in a social condition:

C36P35 > it suffices that they are there, that they surround him, and that they are human beings; they will mutually corrupt each other's moral disposition and make one another evil. (Rel. 6: 94)[37]

C36P36 Restoring the original predisposition 'to its power' is not the end of the story: it seems an individual's predisposition may be so restored, but our common situation in the world

remains inappropriate, *zweckwidrig*—*completion*-hindering rather than completion-promoting. Moreover, 'the human being' is in this condition 'through his own fault' (Rel. 6: 93): the situation itself is culpably bad (*böse*), which essentially distinguishes it from some unfortunate circumstance (something *übel*) that we could reasonably look to escape, as one might migrate from inhospitable terrain. *Religion* 3 thus begins with a renewed expression of pessimism: evil is everywhere, and the hero of *Religion* 1—indeed even a legion of such heroes—may be powerless against it.

C36P37 How should we make sense of this abrupt turn in Kant's account? The answer lies in further examination of the premise of Kant's argument in the *Religion*. The original predisposition to good consists of the creaturely endowments by which we are induced (but not directly compelled) to act in the appropriate ways; *and we situate ourselves in the world* through the exercise of those predispositions. Now, the Stoic natural teleology that informs this assumption is distinctively *comprehensive*: not only must a creature be suitably endowed, psychologically and somatically, to act in the ways that will enable it to manifest fully as a creature of a certain kind; a creature must also be suitably situated, in its environment and in relation to other creatures, for this end.[38] In the case of non-rational animals, this distinction is notional at best: thus Cicero observes that bees do not swarm for the purpose of building honeycombs, but swarm because they are 'gregarious by nature', which finds determinate expression in (among other things) the building of honeycombs (Officiis 1.157).[39] He also twice points to the cooperative arrangement between the *pinna nobilis*, a kind of Mediterranean bivalve, and the little shrimp that lives inside of it, and nips at it to alert it to threats, as well as to the availability of food, tiny fish that they apparently eat in common (Finibus 3.63, Deorum 2.123–24).[40] They are directly compelled by their predispositions to take this arrangement, and find it appropriate.

C36P38 These passages consider how social animals—creatures with an inborn affinity[41] to congregate—are unified as this affinity expresses itself in cooperative action, whereby these creatures are situated in the world in some determinate way. We are like such creatures, Cicero says—but we are bound by *deliberating* and acting together (Officiis 1.157). With characteristic sanguinity, Cicero takes this to mean that our ties are stronger than those binding other social creatures,[42] since—as part of an account of human *oikeiōsis* that abstracts from our corruption—it can be supposed that our ties are forged in cooperative action that is self-consciously aimed at benefiting human beings (Finibus 3.64).

C36P39 But by Kant's lights no account of human development can bracket our corruption, as if it were merely some distracting complication. For this corruption is inevitable in the course of our development, even on the supposition that we have been providentially endowed with everything we need to complete our essentially rational nature. As we have seen, Kant argues in *Religion* 1 that we inevitably corrupt this endowment, the original predisposition to good, in the course of our development as individuals. Now Kant adds another detail from the general Stoic theory of *oikeiōsis*: this providential endowment determines how a creature is situated in the world. This yields a new perspective on radical evil in *Religion* 3, one that considers how we, in the necessary exercise and

development of our predispositions, collectively create social conditions that countermand the telic drive in question.

C36P40 To understand Kant's move in *Religion* 3, it helps to look back to his 1784 'Idea for a Universal History with a Cosmopolitan Aim'. That essay also provides an account of human development that proceeds from the broadly Stoic natural teleology we have been considering. He thus begins with a conception of animals generally as creatures providentially endowed with predispositions to complete their natures; and then, in extending this conception of the animal to the particular case of the human being, marks the essential difference that in us this completion is possible '*only in the species, but not in the individual*' (UH 8: 18). He takes this to be the case on grounds familiar from Cicero: namely, that the particular determination of how we should live and be situated in the world is not fixed by our predispositions. If predispositions lead a creature to act appropriately, then the human being must be induced to act on his own initiative. Hence, Kant reasons, Nature has been miserly in the dispensation of our creaturely endowments of the animal sort, so that we are driven to exercise ingenuity to solve the problem of how we should live—and thereby develop 'those predispositions whose goal is the use of [our] reason' (UH 8: 19–20). We develop these predispositions by coming up with solutions, and learning from others, diachronically over generations.

C36P41 While this development will require cooperative action of some sort, the development of any individual's predisposition nevertheless requires her to take an appropriate interest in the powers of her own agency. There are problems to be solved in the face of the fact that nature gave us 'neither the horns of the steer, nor the claws of a lion' (UH 8: 19) and so forth; and to be the source of a solution to those problems has a value more fundamental than that of the utility of any solution itself—at least from a teleological perspective, one concerned with appropriate, or completion-promoting, action in an essentially rational animal (UH 8: 20). This developmental imperative gives each 'a great propensity [Hang] to *individualise* (isolate) himself' (UH 8: 21). As a result, the bonds of human fellowship are not formed in the unadulterated expression of our deliberating and acting together, as Cicero suggested. We are, rather, bound together in 'unsociable sociability': we are driven to congregate, as this is the condition in which we feel ourselves human, and can develop our natural predispositions; but we are sociable in an unsociable way, since the developmentally appropriate interest any human being takes in the powers of her own agency finds expression in a drive to distinguish these powers over against those of others (UH 8: 20–21).

C36P42 Let us bring this to bear on the *Religion* 3 preamble. To recap: the post-revolution hero of *Religion* 1, who on the terms of that account should now be oriented 'to good', sees that our social situation is itself inappropriate for this goal. Yet this situation is one that the human being has brought upon himself. How? His situation is further analysed as follows:

C36P43 > His needs are but limited, and his state of mind in providing for them moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and *will despise him for it* [*darüber*

> *verachten möchten*]. Envy, addition to power, avarice, and the malignant inclinations associated with these, assail his nature, which on its own is undemanding, *as soon as he is among human beings*. (Rel. 6: 93–94; first emphasis mine)

C36P44 The passage turns on the idea of his being *despised* for being poor: this is to regard his poverty as a fault of his agency, for which he may be blamed. As we have just seen, the human being has a developmental imperative to take an interest in the powers of his own agency. Here Kant points to envy, addiction to power, and avarice as some of the twisted ways we respond to this imperative: first, we allow wealth and power to stand proxy for genuine resources of agency; and we then treat their possession as genuinely *good*, and their absence as genuinely *bad*.

C36P45 One point of the passage, then, is to indicate how socially manifested evaluative authority induces evaluative confusion in individuals. But this can only be one half of a mutually sustaining relation. From the other side, the evaluative confusion of individuals corrupts social relations—for example, if, in the very attitude of 'considering himself poor', the man subordinates himself to the masters of these proxy goods.[43] Thus, as Sharon Anderson-Gold deftly puts it: 'The propensity to evil is not something that is simply 'within me' or 'within you', but something that operates within our very mode of association' (2001: 46). This result follows naturally when we recognize that Kant understands radical evil to be the corruption of the original predisposition to good, and conceives of that predisposition as in broadly Stoic terms as the psychological and somatic affinities that orient us towards the completion of our nature, and includes our being drawn into community with other human beings.

C36S5

5 CONCLUSION: INEVITABLE, BUT NOT NECESSARY

C36P46 I have been arguing that Kant's claims about the inevitability of radical evil should be understood in teleological terms—specifically, in terms of an account of human development that draws on the Stoic doctrine of *oikeiōsis*. This teleology is providential: it starts with the thought that we have been endowed with everything we need to complete our essentially rational nature. But if we inevitably corrupt ourselves at some point in the course of our development from this 'natural principle', how is there not a developmental imperative to do so? And if this is the case, how does Kant meaningfully avoid thinking of radical evil as something *mandated* by our nature, and thus as *necessary*?

C36P47 One might attempt to respond to these worries by treating radical evil as a kind of by-product of acting on these developmental imperatives. But the developmental imperatives in question are issued only once the human being has come into the use of her reason, just as she corrupts the original predisposition to good. As we are compelled

by corrupted developmental imperatives, so we act and situate ourselves in the world inappropriately.

C36P48

But radical evil is unnecessary. We understand this point in the context of the full story of human development. Radical evil is unnecessary because it can be otherwise: we can overcome it. But only the *species* can overcome radical evil, because we must associate on new terms, or 'congregate under' the principle of virtue (Rel. 6: 94). Kant calls this association the 'ethical community', a corporate body of all humanity united towards the telos of virtue, as a good common to all (Rel. 6: 94–100). Although the ethical community is a complex topic unto itself, for our purposes we can note two things about it. The duty to enter into the ethical community is not one that individuals have to other individuals, but one that the species has to itself (Rel. 6: 97). Second, it can only be fulfilled through profound evaluative transformation: for the individual who enters into the ethical community acts as a member of a corporate body, and no longer values virtue as if it might be her own possession, the good she might wreak in her own soul. This tells us why the psychological and social accounts of radical evil, and its overcoming, must be of a piece: for ultimately, it is not *my* virtue, or *your* virtue that matters in the ethical community. If we find it hard to fully comprehend such a thought, let alone manifest it in our social condition, it can only be due to the entrenched propensity of each to value her own agency in fundamentally perverted ways.[44]

Notes

1. Ruppel (2019) indicates that while the modern German terms do eventually assume, in the main, the significances that Kant proposes here, the development is gradual and largely postdates Kant.
2. On this second kind of confusion, see Louden (2010).
3. My account focuses on the *Religion*, and does not attempt to offer a comprehensive assessment of its compatibility with the second *Critique* or other works.
4. Otherwise only some part of the person would be the source of the action—a good part if the action is good, a bad part if the action is bad—which Kant rejects in his defence of 'rigorism' (Rel. 6: 23–25).
5. In an oft-mentioned 1793 letter to Herder, Goethe regards Kantian radical evil as a foul stain upon his critical philosophy, which he supposes Kant undertook to expand its audience to religious Christians; for discussion see Fackenheim (1954).
6. See e.g. the set-up in Michaelson (1990: 8), or Wolterstorff (1991) on related questions about atonement and grace.
7. I could hardly hope to improve upon Pasternack (2020); see Mariña (1997) for an opposing perspective.
8. Kant himself attempts to block this assumption in the second-edition Preface (Rel. 6: 13–14).
9. E.g. from the approximate midpoint of CPrR, at 5:72.
10. Hence the title of Anderson-Gold (2001), *Unnecessary Evil*, a work with which I stand in broad agreement.
11. Hereafter, Parts of the *Religion* will be indicated by succeeding Arabic numeral: e.g. *Religion* 1.

12. The implications of this point for Kant's view of grace fall under the question of his 'Augustinianism', which I am not taking up; see n.7.
13. E.g. Michaelson (1990: 42).
14. I have sanitized Kant's actual example, which is 'all savages have a propensity for intoxicants' (Rel. 6: 28n.).
15. The ensuing inclination she might form, however, is imputable. Kant explains inclination as 'habitual desire' (Rel. 6: 28): an inclination is a desire that has become habitual through repeated gratification, drawing some motive force from the habit itself. This is why inclinations can seem to assail us, driving us against our will or contrary to our better judgment. But since inclinations are formed through repeated acts of gratifying a certain desire in a certain way, and our actions are expressions of (typically tacit) commitment to maxims, we are rationally complicit in them: see also Wood (2018: 100).
16. See the 'deduction' provided in Allison (1990: 152–57); the 'transcendental justification' reconstructed by Morgan (2005), or the 'a priori proof' that 'should resemble transcendental argumentation' in taking as given only 'those features of human agency that are necessarily operative in any context of deliberation and choice' in Papish (2018: 134). Sometimes the reconstructed proof is deemed 'quasi-transcendental' because it must draw on 'anthropological assumptions about the workings of the human mind and the dynamics of social relations' (Muchnik 2009: 73); somewhat differently, Palmquist (2008) reconstructs a 'quasi-transcendental' proof that explains radical evil as a condition of the possibility of religious experience.
17. E.g. Grenberg (2010: 175) on Wood. While Wood takes radical evil to have a social origin, he also thinks that its imputability can *only* be understood in individual-psychological terms (Wood 1999: 288–90)—though he provides no textual evidence for the latter claim. My account in Section 4 suggests the latter claim may be incorrect, but limitations on space prohibit further discussion of this aspect of his view.
18. The term, usually left transliterated, is drawn from the standard word for expressing relations of belonging within a family; approximations include 'orientation', 'affinity', 'endearment', and 'appropriation'—see discussion in Inwood (1985) and Long (1996).
19. See e.g. Cicero (Deorum 2.120–21) for this distinction, which Kant evidently takes for granted in his explication of the concept of life (CPrR 5: 9n.).
20. My sketch of *oikeiōsis* draws mostly on Seneca (Epistulae 121 and 124), and Cicero (Finibus 3.16–24, 3.62–63), which remain among our best sources of the doctrine; on Kant's access to and understanding of Stoic texts relevant to the *Religion*, see Merritt (2020).
21. Seneca (Epistulae 121.15–17) conceives of the primary attachment to one's own constitution (and in that sense to oneself) as a constant that underlies developmental changes in the constitution itself; see Inwood's commentary in Seneca (2007: 339–41)
22. The acquisition of the resources of reason must not be confused with the development of those resources to completion, or virtue.
23. On these points of Stoic psychology, see Inwood (1985) and Brennan (2003); for discussion of its role in the emergence of a conception of free will in later Stoicism, see Frede (2011).
24. Stoics suppose that we acquire reason once we acquire 'the stock of common notions naturally shared by all human beings', around age fourteen (Frede 1996: 11).
25. On this point see Cicero Finibus 3.21–22, and Frede (1999) for discussion.
26. See Long (1968), Cooper (1999).
27. Cato only briefly mentions the *pathē* as manifestations of this corruption at Finibus (3.35), outside either part of the account of *oikeiōsis* (3.16–24 and 3.62–63).

28. I've presented Kant's slight misquotation of the Latin, which he leaves untranslated, and my own translation based on that in Seneca (2010).
29. Consider Kant's similar usage of *Rechtschaffenheit* ('uprightness') e.g. at CPrR 5: 77.
30. On the *Bestimmung* debate see Kuehn (2009), and Brandt (2003; 2007) regarding neo-Stoic influences on it.
31. On the humanity/animality distinction, see also MM 6: 392.
32. The predisposition to personality, however, admits no vicious graft (Rel. 6: 27, 45).
33. See also his remark on usage at (Rel. 6: 46.33–34).
34. A differently elaborated version of these ideas can be found in Merritt (2020).
35. Commentators devote much discussion to Kant's claim there are three 'different grades' to the propensity to evil: frailty, impurity, and depravity (Rel. 6: 29–30); Muchnik (2009: 15–161) takes each grade to be a corrupted form of a corresponding element of the original predisposition to good. But this won't work for two basic reasons: radical evil is the corruption of an entire evaluative outlook, not some part of it; and depravity cannot be the corruption of the predisposition to personality, since Kant takes this element to be incorruptible as such (Rel. 6: 27, 45). As Kant later indicates that depravity captures the core idea of radical evil (Rel. 6: 37), I forego further comment here.
36. Kant mentions John (3:5) and alludes to Colossians (3:9–10).
37. Punctuation altered for clarity. Wood (1999: 283–300) and Wood (2010) take the *Religion* 3 preamble to draw directly from Rousseau; but the language of this passage closely echoes Seneca, who captures the workings of the social forces of corruption through a metaphor of proximate, mechanical force (thus not requiring individuals' intentions). See *De Vita Beata* 1.4 (Seneca 2014: 240–41), and also the metaphors of contagion at Epistulae (7.2–7.5) and *De Tranquillitate Animi* 7.4 (Seneca 2014: 193).
38. We should therefore not confuse it with Aristotelian natural teleology, which is not normally understood to be comprehensive in this way; for background, see Henry (2015).
39. Cicero is widely understood to be expressing the Stoic position of Panaetius in *Officiis* 1–2 (see the commentary of Griffin and Adkins, in Cicero [1991, xviii–xix]).
40. From the speeches of Cato and Balbus, respectively, as representatives of Stoicism.
41. *commendatio* (Finibus 3.63).
42. This is explicit at Officiis (1.57), which helps explain Cato's remark to this effect at Finibus (3.63).
43. See (e.g.) Epictetus *Discourses* (1.4.19) for a fitting example of this point.
44. Thanks to Anil Gomes, Andrew Stephenson, and Reed Winegar for written comments and Markos Valaris for discussion. This research was supported by the Australian Research Council (FT180100494).

C36S6

REFERENCES

C36P49 Alexander of Aphrodisias (1991), *Ethical Problems*, trans. R.W. Sharples (Bloomsbury). **Cited by section and line number**.

C36P50 Allison, Henry (1990), *Kant's Theory of Freedom* (Cambridge University Press).

C36P51 Anderson-Gold, Sharon (2001), *Unnecessary Evil: History and Moral Progress in the Philosophy of Immanuel Kant* (State University of New York Press).

C36P52 Brandt, Reinhard (2003), 'The Vocation of the Human Being', in B. Jacobs and P. Kain (eds.), *Essays on Kant's Anthropology* (Cambridge University Press), 85–104.

C36P53 Brandt, Reinhard (2007), *Die Bestimmung des Menschen bei Kant* (Felix Meiner).

C36P54 Brennan, Tad (2003), 'Stoic Moral Psychology', in B. Inwood (ed.), *Cambridge Companion to the Stoics* (Cambridge University Press), 257–294.

C36P55 Cicero (1928), *De Officiis*, trans. W. Miller, Loeb Classical Library (William Heinemann Ltd). **Cited by abbreviated title, book, and section number**.

C36P56 Cicero (1967), *De Natura Deorum*, *Academica*, trans. H. Rackham, Loeb Classical Library (Harvard University Press). **Cited by abbreviated title, book, and section number**.

C36P57 Cicero (1991), *On Duties*, trans. M. T. Griffin and E. M. Atkins (Cambridge University Press). **Cited according to Cicero (1928).**

C36P58 Cicero (1997), *The Nature of the Gods*, trans. P. G. Walsh (Oxford University Press, Clarendon Press). **Cited according to Cicero (1967)**.

C36P59 Cicero (1998), *De Finibus Bonorum et Malorum*, Oxford Classical Texts (Oxford University Press, Clarendon Press). **Cited by abbreviated title, book, and section number**.

C36P60 Cicero (2001), *On Moral Ends*, ed. J. Annas, trans. R. Woolf (Cambridge University Press). **Cited according to Cicero (1998).**

C36P61 Cooper, John (1999), 'Posidonius on Emotions', in *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory* (Princeton University Press), 449–484.

C36P62 Diogenes Laertius (1925), *Lives of the Eminent Philosophers*, trans. R. D. Hicks, Loeb Classical Library (Harvard University Press). **Cited by book and section number**.

C36P63 Epictetus (1925), *Discourses*, Books 1–2, trans. W. A. Oldfather, Loeb Classical Library (Harvard University Press). **Cited by book, discourse, and section number**.

C36P64 Fackenheim, Emil (1954), 'Kant and Radical Evil', *University of Toronto Quarterly* 23/4: 339–353.

C36P65 Frede, Michael (1996), 'Introduction', in M. Frede and G. Striker (eds.), *Rationality in Greek Thought* (Oxford University Press, Clarendon Press), 1–28.

C36P66 Frede, Michael (1999), 'On the Stoic Conception of the Good', in K. Ierodiakonou (ed.), *Topics in Stoic Philosophy* (Oxford University Press, Clarendon Press), 71–94.

C36P67 Frede, Michael (2011), *A Free Will: Origins of the Notion in Ancient Thought* (University of California Press).

C36P68 Grenberg, Jeanine (2010), 'Social Dimensions of Kant's Conception of Radical Evil', in S. Anderson-Gold and P. Muchnik (eds.), *Kant's Anatomy of Evil* (Cambridge University Press), 173–194.

C36P69 Guyer, Paul (2009), 'The Crooked Timber of Mankind', in A. O. Rorty and J. Schmidt (eds.), *Kant's* Idea for a Universal History with a Cosmopolitan Aim: *A Critical Guide* (Cambridge University Press), 129–149.

C36P70 Henry, Devin (2015), 'Aristotle on the Cosmological Significance of Animal Generation', in D. Ebrey (ed.) *Theory and Practice in Aristotle's Natural Science* (Cambridge University Press), 100–118.

C36P71 Inwood, Brad (1985), *Ethics and Human Action in Early Stoicism* (Oxford University Press, Clarendon Press).

C36P72 Kuehn, Manfred (2009), 'Reason as a Species Characteristic', in A. O. Rorty and J. Schmidt (eds.), *Kant's* Idea for a Universal History with a Cosmopolitan Aim: *A Critical Guide* (Cambridge University Press), 68–93.

C36P73 Long, A. A. (1968), 'The Stoic Concept of Evil', *Philosophical Quarterly* 18/73: 329–343.

C36P74 Long, A. A. (1996), 'Hierocles on *Oikeiōsis* and Self-Perception', in *Stoic Studies* (University of California Press), 250–263.

C36P75 Louden, Robert (2010), 'Evil Everywhere: the Ordinariness of Kantian Radical Evil', in S. Anderson-Gold and P. Muchnik (eds.), *Kant's Anatomy of Evil* (Cambridge University Press), 93–115.

C36P76 Mariña, Jacqueline (1997), 'Kant on Grace: A Reply to His Critics', *Religious Studies* 33/4: 379–400.

C36P77 Merritt, Melissa (2020), 'Nature, Corruption, and Freedom: Stoic Ethics in Kant's *Religion*', *European Journal of Philosophy*, DOI: 10.1111/ejop.12553.

C36P78 Michaelson, Gordon (1990*), Fallen Freedom: Kant on Radical Evil and Moral Regeneration* (Cambridge University Press).

C36P79 Morgan, Seiriol. 2005. 'The Missing Formal Proof of Humanity's Radical Evil in Kant's *Religion*', *Philosophical Review* 114/1: 63–114.

C36P80 Muchnik, Pablo (2009), *Kant's Theory of Evil: An Essay on the Dangers of Self-Love and the Apriority of History* (Rowman & Littlefield, Lexington Books).

C36P81 Palmquist, Stephen (2008), 'Kant's Quasi-Transcendental Argument for a Necessary and Universal Evil Propensity in Human Nature', *The Southern Journal of Philosophy* 46/2: 261–297.

C36P82 Papish, Laura (2018), *Kant on Evil, Self-Deception, and Moral Reform* (Oxford University Press).

C36P83 Pasternack, Lawrence (2020), 'On the Alleged Augustinianism in Kant's *Religion*', *Kantian Review* 25/1: 103–124.

C36P84 Rossi, Philip (2005), *The Social Authority of Reason: Kant's Critique, Radical Evil, and the Destiny of Humankind* (State University of New York Press).

C36P85 Ruppel, Antonia (2019), 'Kakology: A Study of Some Evil Words', in A. Chignell (ed.), *Evil: A History* (Oxford University Press), 43–59.

C36P86 Seneca (1917, 1920, and 1925), *Epistulae Morales*, 3 vols., trans. R. M. Gummere, Loeb Classical Library (Harvard University Press). **Cited by abbreviated title, letter, and section number.**

C36P87 Seneca (1928 and 1932), *Moral Essays*, 2 vols., trans. J. W. Basore, Loeb Classical Library (Harvard University Press). **Individual works are cited by abbreviated Latin title, part, and section number.**

C36P88 Seneca (2007), *Selected Philosophical Letters*, trans. B. Inwood, with commentary (Oxford University Press, Clarendon Press). **Cited according to Seneca (1917, 1920, and 1925).**

C36P89 Seneca (2010), *Anger, Mercy, Revenge*, trans. R. Kaster and M. C. Nussbaum (University of Chicago Press). **Cited according to Seneca (1928 and 1932).**

C36P90 Seneca (2014). *Hardship and Happiness*, trans. E. Fantham, H. Hine, J. Ker, G. Williams (University of Chicago Press). **Individual works from this volume are cited according to Seneca (1928 and 1932).**

C36P91 Wolterstorff, Nicholas (1991), 'Conundrums in Kant's Rational Religion', in P. Rossi and M. Wreen (eds.), *Kant's Philosophy of Religion Reconsidered* (Indiana University Press), 40–53.

C36P92 Wood, Allen (1999), *Kant's Ethical Thought* (Cambridge University Press).

C36P93 Wood, Allen (2010), 'Kant and the Intelligibility of Evil', in S. Anderson-Gold and P. Muchnik (eds.), *Kant's Anatomy of Evil* (Cambridge University Press), 114–172.

C36P94 Wood, Allen (2014), 'The Evil in Human Nature', in G. Michaelson (ed.), *Kant's* Religion within the Boundaries of Mere Reason: *A Critical Guide* (Cambridge University Press), 31–57.

C36P95 Wood, Allen (2018), 'Feeling and Desire in the Rational Animal', in K. Sorensen and D. Williamson (eds.), *Kant and the Faculty of Feeling* (Cambridge University Press), 88–106.

C36P96 Wood, Allen (2020), *Kant and Religion* (Cambridge University Press).