

# Every Man Has His Price: Kant's Argument for Universal Radical Evil

(Forthcoming in *Inquiry: An Interdisciplinary Journal of Philosophy*. Penultimate draft – please cite published version)

Jonas Jervell Indregard

National Research University Higher School of Economics, Moscow

*Abstract:* Kant famously claims that we have all freely chosen evil. This paper offers a novel account of the much-debated justification for this claim. I reconstruct Kant's argument from his affirmation that we all have a price – we can all succumb to temptation. I argue that this follows a priori from a theoretical principle of the *Critique of Pure Reason*, namely that all empirical powers have a finite, changeable degree, an intensive magnitude. Because of this, our reason can always be overpowered by sensible inclinations. Kant moreover holds that this necessary feature of our moral psychology should not have been the case: We ought to instead be like the divine human being, for whom the moral law yields a greater incentive than any possible temptation. On Kant's view, we are thus responsible for having a price, and the synthetic a priori fact that we do proves that we each made an initial choice of evil.

*Keywords:* Immanuel Kant, radical evil, intensive magnitude, freedom, empirical character, causality

## 1. Introduction

Kant's *Religion* famously claims that *all* natural human beings are initially evil. Kant seemingly thinks this follows from the truth of the saying “every man has his price, for which he sells himself” (RGV, 6:38-39). However, interpreters struggling with Kant's puzzling doctrine of universal evil have hitherto not found this saying helpful. Why must *every* man have his price? And how could we know? Even the “member of the English parliament” (RGV, 6:38) to whom Kant ascribes the saying presents it as a generalization concerning *most* men rather than a universal, a priori truth: “It is an old maxim, that every man has his price, if you can but come up to it: this, I hope, does not hold true of every man, but I am afraid it too generally holds true; and that of a great many it may hold true, is what I believe was never doubted of.”<sup>1</sup>

---

<sup>1</sup> Sir William Wyndham, speech to the House, March 13, 1734; quoted in Coxe (1798: 415). The Cambridge Translation of Kant's *Religion* instead refers the saying to Sir Robert Walpole, the English Prime Minister at the

This article argues that, initial appearances aside, the saying “every man has his price” – when properly understood – forms the basis of a novel and plausible reconstruction of the Kantian argument. This argument amounts to a synthetic a priori proof that all human beings within possible experience begin from freely chosen evil. Readings of Kant’s doctrine of radical evil traditionally divide into two groups: “Empirical” interpretations take Kant’s claim to be based on empirical generalization and deny that there is a formal proof;<sup>2</sup> “a priori” interpretations propose synthetic a priori proofs of universal radical evil on Kant’s behalf.<sup>3</sup> My interpretation belongs to the latter group: We can infer from the necessary conditions of possible experience that every man has his price.<sup>4</sup> Unlike extant proposals, I identify a specific synthetic a priori cognition from *KrV* on which the argument rests.

Kant’s endorsement of the saying “every man has his price” reflects neither a dubious empirical psychology, nor an equally dubious claim about our free will’s susceptibility to monetary temptation (which would then not be true freedom at all). Rather, it reflects the essential nature of the finite causal powers of the mind (*Gemüthskräfte*). These powers fall under a synthetic a priori principle Kant argues for in the *Critique of Pure Reason*: that every empirical reality has a finite degree, an intensive magnitude. Applied to the empirical powers of the mind, this principle brings out the underlying truth of “every man has his price,” namely what I call Possible Overpowering (PO): When it comes to our empirical ability to choose between different courses of action (our *Willkür* as phenomenon), the degree of influence exerted by reason can always be

---

time, who in fact made a slightly different claim several years later, that “[a]ll those men [referring to “certain patriots”] have their price.” Nonetheless, the saying has since been commonly (mis)attributed to him. Which of these Kant had in mind is unclear – in any case, neither version of the claim has universal scope.

<sup>2</sup> See Frierson (2003), Grenberg (2005), Loudon (2009), Quinn (1988); they disagree on the extent to which this empirical generalization is justified.

<sup>3</sup> See Allison (1990), Palmquist (2008), Morgan (2005), Papish (2018), Muchnik (2009) (although Muchnik labels his argument as “hybrid,” incorporating empirical elements), and Formosa (2007) (although Formosa thinks Kant’s argument fails).

<sup>4</sup> One could add a third group of “subjective” interpretations: whereas the two other groups adduce objective (a priori or empirical) grounds for universal radical evil, this third group holds that universal radical evil is affirmed on *subjective* – but nonetheless universal – grounds. They may for instance hold that it is a teleological claim about human biology, psychology, or culture; or a necessary regulative claim on moral grounds. See Anderson-Gold (1991), Wood (1991) and (1999), Grimm (2002), Kemp (2011), Serck-Hanssen (2012), Kohl (forthcoming).

overpowered by the degree of influence exerted by sensibility. Kant sees this as ineluctable for all natural human beings, who can therefore always be influenced by sensibility to immoral action: the “subjective ground of the possibility of the deviation of the maxims from the moral law (...) belongs to the human being universally” (RGV, 6:29).

PO is a synthetic a priori cognition concerning human psychology, independent of empirical premises or generalization. Kant’s affirmation of universal radical evil follows if we are *responsible* for PO. I will show that this is indeed Kant’s view. Our empirical powers, and PO as a feature of how these powers relate, belong to our empirical character, which on Kant’s view has our freely chosen intelligible character as its “transcendental cause” (KrV, A546/B574). We are responsible for the causality of our intelligible character and hence for PO. Our powers *should* instead relate in a way that we can think (but not experience), namely like in the *divine* human being (*göttlich gesinnter Mensch*) for whom PO does not hold. We can reject evil and strive towards this ideal by influencing our empirical powers in the right direction – towards virtue, the “moral strength of a human being’s will” (MS, 6:405). But as natural human beings, we remain infinitely removed from the ideal (see RGV, 6:66), since our moral strength – within experience – must remain finite in degree. We will always have a price.

Some caveats: First, Kant famously does not himself provide an a priori argument. As often quoted, with varying degrees of frustration, Kant holds that we can “spare ourselves the formal proof that there must be such a corrupt propensity rooted in the human being, in view of the multitude of woeful examples that the experience of human *deeds* parades before us” (RGV, 6:32-33).<sup>5</sup> My proposal is thus a reconstruction. Kant’s phrasing does suggest that there *is* a formal proof available (cf. Grimm 2002: 164-165),<sup>6</sup> and if my proposal is correct the lack of proof is explained by Kant’s expressed intention in the *Religion* not to require knowledge of the theoretical principles of *KrV* from its readers. I hope further to make plausible that the “every man has his price” saying, and other passages considered below, gesture towards the proposed argument. Second, my ambition is only to provide a plausible reconstruction of Kant’s reasoning

---

<sup>5</sup> Already a year after its publication, Eberhard (1794: 41-42) targeted this lack of proof in Kant’s *Religion*; see further Loudon (2009: 109-110).

<sup>6</sup> Kant later states: “Of the evil that lies in the human heart and of which nobody is free (...) everyone can be convinced *through his reason*” (RGV, 6:163, my italics).

– defending its independent philosophical plausibility is beyond my scope. Third, I assume from the outset that Kant’s talk of powers of the mind is legitimate, and specifically, that he justifiably appeals to *empirical* mental powers. Defending this surprisingly controversial assumption would take us too far afield (for discussion see Frierson 2014). I use it to develop a novel reconstruction of Kant’s argument for universal radical evil; indirectly, success at this task supports the assumption by showing its fruitfulness.

## 2. Human Predispositions and Power of Choice

This section and the next aims to gradually introduce the premises underlying the “every man has his price” saying, beginning from a weak, analytic assumption concerning the constitution of the human mind, and moving incrementally towards stronger, more controversial, and synthetic assumptions needed for universal radical evil to follow. This approach helps pinpoint the role of the synthetic a priori premise I attribute to Kant’s argument and distinguish it from analytic claims. Moreover, previous commentators have held weaker assumptions to be sufficient grounds for Kant’s claim of universal evil; explaining their insufficiency positions my reconstruction within the existing literature.

Kant famously holds that human beings have an essentially *composite* nature, comprising a sensuous and a rational side:

*Composite:* Human beings have two distinct faculties, namely sensibility and reason.

This distinguishes us from beings with intellectual intuition – from the intuitive intellect of God; and from animals lacking rationality, whose power of choice is necessitated by sensible inclinations alone (V-Met-L2/Pölitz, 28:588; V-Met/Arnoldt, 29:1015; V-Mo/Mron II, 29:611). On occasion, Henry Allison’s influential reading appears to infer universal radical evil directly from Composite: He interprets Kant’s contrast between human beings and beings with a “holy will” as a contrast between “a will (or agent) [that] is not sensuously affected” and “finite, sensuously affected rational beings such as ourselves” (Allison 1990: 156). This may be taken to mean that by having sensibility, human beings are *ipso facto* subject to temptation, which Allison

takes to imply radical evil.<sup>7</sup> However, Allison plausibly intends something more than Composite by the term “sensuously affected,” namely a sensible influence specifically on our choice, *Willkür*. There could be creatures whose sensibility did not exert any potentially immoral influence on their *choice*: “finite holy beings” (MS, 6:383) may constitute an example.<sup>8</sup> Such angelic beings would regard sensibility merely as a source of information, utilised by reason in determining how to act. However, sensibility would not be a separate source of desires or inclinations. A slightly stronger assumption might therefore better capture Allison’s intention:

*Potential Conflict*: Human beings have two potentially conflicting influences on their choice (*Willkür*), namely sensibility and reason.

We face not just distinct faculties but a potential *clash* between sensible inclinations and moral demands of reason. For Kant, Potential Conflict distinguishes human beings from “finite holy beings (who could never be tempted to violate duty)” (MS, 6:383). Human *Willkür* is “affected, but not [completely] determined” by sensible impulses (MS, 6:214; cf. KrV, A534/B562; V-Met/Dohna, 28:677). Our actions are “in great part occasioned but not completely determined by sensibility” (Refl 5611, 18:252); complementarily, we are not creatures that necessarily comply with oughts, as we would be “if reason completely determined the will” (KpV, 5:20).

What is the status of Potential Conflict? Textual evidence suggests that Kant sees it as an *analytic* claim about human beings. Consider the *Religion*’s account of the three “predispositions (...) that relate immediately to the faculty of desire and the exercise of the power of choice [*Gebrauch der Willkür*]” (RGV, 6:28): One is the “predisposition to animality” (RGV, 6:26), another the “predisposition to personality” (RGV, 6:27) – simplifying somewhat these indicate the influence of sensibility and of reason, respectively, on the power of choice. Both of them are for Kant “*original*, for they belong to the possibility of human nature” (RGV, 6:28). Predispositions are original if “they belong with necessity to the possibility of this being, (...) [whereas they are] *contingent* if the being in question is possible in itself also without them”

---

<sup>7</sup> Wood’s criticism of Allison for turning evil into a “trivial practical corollary of our finitude” (1999: 287) seems to interpret Allison in this way.

<sup>8</sup> Early lecture notes indicate that Kant has angels in mind, cf. V-PP/Herder, 27:13. (Occasionally, lecture notes report Kant as denying that finite holy beings are possible, cf. V-Phil-Th/Pölitiz, 28:1113. However, this is unlikely to be his considered view.)

(RGV, 6:28). If a human being without these predispositions is not “possible in itself” (not just within possible experience), they must belong analytically to the concept of a human being.<sup>9</sup>

Is Potential Conflict enough to qualify as *evil*? Here, I part way with several interpreters in holding that it is not.<sup>10</sup> I reason as follows: Potential Conflict does not guarantee both conflicting parties potential *victory*. One can conceive of a being for whom Potential Conflict holds, but where reason (or sensibility) always, and inevitably, has the upper hand in cases of conflict. In fact, Kant himself describes such a being: the *divine human being* characterized in the *Religion* (6:60ff.). This being is “afflicted by just the same needs and hence also the same sufferings, by just the same natural inclinations and hence also the same temptations to transgression, as we are” (RGV, 6:64, see also 6:61), and thus encumbered with Potential Conflict.<sup>11</sup> Yet, crucially, the divine human being is nonetheless “free from innate propensity to evil” (RGV, 6:80n.).<sup>12</sup>

In the divine human being, Potential Conflict coexists with a premise ala ‘Necessary Overpowering: Reason necessarily overpowers sensibility in cases where their influences on *Willkür* conflict.’<sup>13</sup> All temptations are thereby resisted. Radical evil entails that “pleasure can induce them to break the moral law, even though they recognize its authority” (MS, 6:379). Since Potential Conflict combined with Necessary Overpowering makes this impossible,

---

<sup>9</sup> See also: “The human being has many instincts belonging to animality, and (...) has to have them if he is to continue being human” (V-Phil-Th/Pölit, 28:1078).

<sup>10</sup> Against for instance Allison (1990); Grimm (2002), who seems to read Potential Conflict as a teleological, rather than analytic, claim about human nature that suffices for radical evil; and Caswell, who claims that “temptations to immorality” (Caswell 2006: 202) suffices for evil.

<sup>11</sup> Baxley (2010: 55-56) assumes that the divine human being is a finite holy will. However, in subjecting itself to Potential Conflict, it *divests* itself of an aspect of this holiness: “[W]e represent to ourselves this God-like human being, our prototype, in such a way that, though himself holy and hence not bound to submit to sufferings, he *nonetheless takes these upon himself in the fullest measure* for the sake of promoting the world's greatest good” (RGV, 6:61, my emphasis). In becoming human, the divine being brings suffering and temptation (incompatible with holiness as such) upon itself “for the sake of promoting the world's greatest good.”

<sup>12</sup> Some interpreters dispute this; for discussion see footnote 24 below.

<sup>13</sup> Flew (1962) independently emphasizes the distinction I find in Kant between having no temptations (finite holy wills) and having temptations but necessarily overcoming them (the divine human being).

Potential Conflict by itself is insufficient for radical evil.<sup>14</sup> If I am right that Kant sees Potential Conflict as analytic, this is no surprise: Kant is clear that universal radical evil is not an analytic claim (RGV, 6:29, 6:32). Further indication that Potential Conflict is insufficient for radical evil is found in Kant's insistence that the mere presence of sensible inclinations that conflict with demands of morality is not evil:

The ground of this evil cannot (...) be placed, as is commonly done, in the sensuous nature of the human being, and in the natural inclinations originating from it. For not only do these bear no direct relation to evil (they rather give the occasion for what the moral disposition can demonstrate in its power, for virtue): we also cannot presume ourselves responsible for their existence (we cannot because, as conatural to us, natural inclinations do not have us as their author), though we can well be responsible for the propensity to evil (RGV, 6:34f.).

This passage confirms both points already made: First, that Potential Conflict is not sufficient for evil because it could be merely the occasion for demonstrating the power of the moral disposition, i.e., it does not guarantee the potential *victory* of sensible inclinations over moral demands; second, that Potential Conflict, as "conatural to us," pertains analytically to human beings and thus cannot be something we are responsible for. At this point we must therefore ask: Is there a synthetic a priori judgment that we can add in order to reach evil, i.e., a claim that would establish the human condition within experience as one where sensibility might win out over reason in cases of conflict?

### 3. Degrees of Power

---

<sup>14</sup> Doesn't our *freedom* rule out Necessary Overpowering, and thereby suffice, together with Potential Conflict, for radical evil? The suggestion is that Potential Conflict, plus the premise that we are *free* to choose either side of the conflict, entails that we can always be tempted to actually do evil. Therefore, the victory of sensibility over reason in cases of conflict is a real and inextirpable possibility. Against this suggestion, Kantian universal radical evil is not merely the *possibility* of freely choosing evil, but the much stronger thesis that all natural human beings have *in fact* made an initial choice to subordinate morality to sensibility (see RGV, 6:43). The divine human being has not made this initial choice, despite being both free and subjected to Potential Conflict.

I propose a novel candidate for the synthetic a priori basis for universal radical evil, namely the principle of the Anticipations of Perception in *KrV*: “**In all appearances the real, which is an object of the sensation, has intensive magnitude**, i.e., a degree” (B207). “The real” here means the positive, non-trivial, and non-geometrical properties of appearances that Kant calls their *realities*.<sup>15</sup> Without going into the intricacies of the principle and Kant’s justification of it in *KrV*, the following point shows why the principle is relevant to radical evil: the two conflicting influences on our *Willkür* – sensibility and reason – are realities, positive properties of ourselves as inner objects of experience, and therefore have degrees. We thus get the following premise:

*Conflict of Finite Powers*: Natural human beings have two potentially conflicting powers with finite degrees influencing choice (*Willkür*), namely sensibility and reason.

This adds to Potential Conflict the idea that these conflicting influences on *Willkür* are, so to speak, on the same playing field: they both have a *finite, quantifiable* impact.<sup>16</sup> A synthetic a priori element thereby supplements the analytic claims. Our powers having a finite degree is not contained analytically in the concept of a human being, but, given the *KrV*’s principle of intensive magnitude, is nonetheless inevitably the case within experience.

Reading sensibility and reason as empirical realities is well supported by textual evidence. At one point Kant specifically ascribes to all the different faculties of the mind a “degree of reality”: “[O]ne (...) cannot deny to [the soul as object of inner sense], any more than to any other existence, an intensive magnitude, i.e., a degree of reality in regard to all its faculties, indeed to everything in general that constitutes its existence” (*KrV*, B414; see MAN, 4:542; V-Met/Mron, 29:905f., 29:912f.; V-Met-K2/Heinze, 28:761, 28:763f.; V-Met/Arnoldt, 29:1037; Refl 5650, 18:299; VNAEF, 8:413). Sensibility is said to have a “determinate degree of receptivity” (*KrV*, A172/B214), and “[i]mpulses of nature, accordingly, involve *obstacles* within the human being’s mind to his fulfillment of duty and (sometimes powerful [*mächtige*]) forces [*Kräfte*] opposing it” (MS, 6:379f.). Even reason, the faculty Kant interpreters perhaps most hesitate to ‘naturalize’ as

---

<sup>15</sup> For more on Kant’s conception of reality and its relation to causal power, see Warren (2001: ch. 1); Giovanelli (2011). Kant’s argument for the principle is defended in Jankowiak (2013).

<sup>16</sup> The divine human being’s reason does *not* have a finite degree in this sense (Kant describes it as a “maximum” (V-Mo/Mron II, 29:604f.)). This reflects the well-known conception of the divine as having certain positive properties *fully* or *infinitely*, whereas we have them to a finite, limited degree (*KpV*, 5:131n.). See further below.



an empirical causal power, is characterized by Kant as having empirical causality. We should consider empirical reason more closely, since reason is not *merely* an empirical power:

A human being really finds in himself a capacity by which he distinguishes himself from all other things, even from himself insofar as he is affected by objects, and that is *reason*. (...) Because of this a rational being must regard himself *as intelligence* (hence not from the side of his lower powers) as belonging not to the world of sense but to the world of understanding (GMS, 4:452; see KrV, A546/B574f.).

Given its importance for his moral philosophy, Kant mostly emphasizes this intelligible aspect of our reason. However, there are clear indications that reason is *also* an empirical power: “We have in us a faculty that not only stands in connection with its subjectively determining grounds, which are the natural causes of its actions – *and thus far is the faculty of a being which itself belongs to appearances* – but that is also related to objective grounds that are mere ideas, insofar as these ideas can determine this faculty, a connection that is expressed by *ought*. This faculty is called *reason*” (Prol, 4:344-5, first emphasis mine). In *KrV*, Kant states that “even though it is reason, it must nevertheless exhibit an empirical character” (A549/B577) and that we “cognize practical freedom through experience, as one of the natural causes, namely a causality of reason in the determination of the will” (A803/B831; see *MS*, 6:418). Elsewhere he reportedly says that “even [one’s] reason, as subjected to the laws of nature, can be considered devoid of all freedom” (V-MS/Vigil, 27:502). Systematically, empirical psychology’s investigation of actual human representation, thought, and action cannot plausibly leave out the effects of reason – and indeed, Kant’s empirical discussions of mental faculties always include reason.<sup>17</sup>

Returning to Conflict of Finite Powers, it helps explain what so many commentators have found intolerable about Kant’s *Religion*: The idea that Kant *has* a formal proof of universal radical evil available, yet declines to offer it. Why not present the proof and (with the 20/20 characteristic of hindsight) spare us centuries of inconclusive speculation and labor? However, Kant explicitly intends the *Religion* not to invoke the complicated theoretical argumentation of the *KrV*: “[O]nly

---

<sup>17</sup> The “higher powers,” including reason, are a staple of discussion in Kant’s empirical psychology and anthropology, see e.g. Anth, 7:196f.; V-Met/Herder, 28:863-865; V-Met/Mron, 29:888-890; V-Anth/Pillau, 25:773f.; V-Anth/Mensch, 25:1032f.; V-Anth/Mron, 25:1296; V-Anth/Busolt, 25:1476; see further Frierson (2005: 14, 23f.; 2014: 7-8), Kain (2003: 235). For more on reason as a natural cause, see Rauscher (2006), (2015: 114f.).

common morality is needed to understand the essentials of this text, without venturing into the critique of practical reason, still less into that of theoretical reason” (RGV, 6:14). Hence it is unsuited for a formal proof based on one of *KrV*’s First Principles of Pure Understanding, explicating its consequences within Kant’s overall theory of human powers and faculty of desire. One may still disapprove of Kant’s pedagogical choice here, of course. This reading at least has him proceeding consistently with his expressed intentions, rather than haphazardly neglecting to give (or gesturing towards without actually having) what seems like a crucial proof. As far as I can tell, this explanation is not available to other extant a priori interpretations of the universality of evil, since they do not appeal to specific critical principles ‘of theoretical reason’.

Conflict of Finite Powers is not enough for inferring universal radical evil. Radical evil entails that sensibility might win out over reason in cases where their influence on our *Willkür* conflict. Even given Conflict of Finite Powers, one power might inevitably overpower the other – as, say, the finite number 10 is always greater than the finite number 5. A further premise must be added:

*Changeability*: The strength of one’s powers is subject to possible change over time.

On Kant’s view, our empirical powers may change over time (see further Indregard 2018). Kant considers Changeability a corollary to the principle of the Anticipations of Perception: In explicating the principle he states that “every reality in appearance, however small it may be, has a degree, i.e., an intensive magnitude, *which can still always be diminished*” (KrV, A169/B211, my emphasis), “every reality has its degree that can decrease to nothing” (KrV, A172-173/B214), and again, “everything real has for the same quality its degree (...) [which] can become infinitely smaller until it is transformed into emptiness and disappears” (KrV, A174/B216).<sup>18</sup> Kant uses this point specifically with respect to empirical *mental* powers in the Paralogisms, in arguing that the soul cannot be shown to be a substance:

---

<sup>18</sup> Some read these passages as making a merely phenomenological claim about the possibility of *imagining* change in degree (see e.g. Buroker 2006: 152f.), without implying the *real* possibility of change in degree of reality. However, admitting “imagined possibility” – as something more than logical, yet less than real possibility – is controversial (Stang (2016) counts only real and logical possibility among Kant’s main concepts of possibility, see e.g. the figure on (2016: 271)). Moreover, passages I cite in what follows, for instance concerning the duty to strengthen and develop one’s powers, clearly require real and not merely imagined possibility of change in degree.

[O]ne (...) cannot deny to [the soul], any more than to any other existence, an intensive magnitude, i.e., a degree of reality in regard to all its faculties, which might diminish through all the infinitely many smaller degrees (...) [T]he supposed substance (the thing whose persistence has not otherwise been established already) could be transformed into nothing, although not by disintegration, but by a gradual remission (*remissio*) of all its powers (KrV, B414-415; cf. MAN, 4:542; V-Met/Mron, 29:905f., 29:912f.; V-Met-K2/Heinze, 28:761, 28:763f.; V-Met/Arnoldt, 29:1037; Refl 5650, 18:299).

Kant thus takes the Anticipations of Perception principle to imply that the degree of strength of one's powers can change over time. We rediscover this throughout Kant's philosophy: Kant repeatedly encourages changing, for moral purposes, the strength of one's powers, referring to "the duty of man to develop his powers *quoad maxime*" (V-MS/Vigil, 27:625f.; see also Päd, 9:472ff.) and gradually acquiring virtue as "moral strength" (MS, 6:405).<sup>19</sup> This may involve combatting the "weakness in the use of one's understanding" (MS, 6:408; cf. Anth, 7:253) or "to *cultivate* [moral feeling] and to strengthen it" (MS, 6:399-400).<sup>20</sup> Conversely, Kant allows for the weakening of one's reason: "[T]he basis of great crimes is merely the force of inclinations that *weaken* reason" (MS, 6:384, emphasis in original). Kant's account of aesthetic experience also refers to changes in strength of mental powers: perceiving beauty brings "a consciousness of the harmony of our powers of representation (...) in which we feel our entire cognitive faculty (understanding and imagination) strengthened" (KpV, 5:160; see KU, 5:313, 5:366).

Conflict of Finite Powers, together with Changeability, plausibly suffices to get us the claim underlying Kant's thesis of radical evil, namely that it is always possible for our sensibility to win out over the demands of morality in cases of conflict. Before concluding this section, I will briefly discuss another premise that also seems crucial for Kant's moral psychology, although its precise relation to the proof of radical evil is not quite clear:

---

<sup>19</sup> The *Anthropology* states that "the wisdom of nature has planted in us the predisposition to compassion in order to handle the reins *provisionally*, until reason has achieved the necessary strength" (Anth, 7:253).

<sup>20</sup> I here simplify Kant's view by conceiving sensibility and reason merely as two potentially opposed powers influencing our *Willkür*. As the idea of cultivating moral *feeling* indicates, the actual relation between them is considerably more complex, and Kant does not advocate a view of virtue on which it simply means suppressing one's sensibility and strengthening one's reason. For our purposes these complications can be set aside.

*Modulation*: The strength of sensibility's influence on our *Willkür* changes with the nature of the external influences on sensibility, e.g., with the strength of the affection of sensibility.

The input sensibility receives makes a difference to the strength of its influence on our *Willkür*. We must countenance not only the strength of sensibility itself, as a power, but also the strength of its various inputs. The influence of sensibility on our choice differs depending on whether it is affected by something moderately tempting (say, a piece of fruit) or something we find practically irresistible (say, a delicious piece of cake). The idea behind *Modulation* is something like this: we can conceive of the strength of the influence on choice as the strength of the input multiplied by the strength of the power of sensibility. Of course, this is a simplification – the details will vary depending on empirical factors and do not matter for my purposes.

An interesting result would arise if we could add to *Modulation* that the strength of the input (the impression made on sensibility) is *positively, non-asymptotically correlated* with the strength of sensibility's influence on *Willkür*. We could then infer that no matter how weak one's sensibility, and no matter how strong one's reason, there is always a *possible* input to one's sensibility of such strength that it 'modulates' sensibility to overpower reason's influence on *Willkür*. If the cake smells sufficiently delicious, or the pain is sufficiently excruciating, even the most stalwart and stoic of us would succumb to temptation. This begins to sound a lot like Kant's suggestion: "Every man has his price, for which he sells himself" (RGV, 6:38-39).

Strengthened in this way, *Modulation* and *Conflict of Finite Powers* suffice, even without *Changeability*, to ground Kant's argument for radical evil – entailing that every man indeed has his price. However, there is a problem: The strengthened version of *Modulation* is not plausibly a priori. It does not appear derivable from the *Anticipations of Perception*, nor from additional a priori (synthetic or analytic) claims. The correlation being non-asymptotic seems particularly difficult to justify a priori (or even empirically, for that matter). It is conceivable that Kant (if strengthened *Modulation* plays a role in the "formal proof" he had in mind) simply did not

consider the possibility of an asymptotic correlation.<sup>21</sup> However, it seems preferable not to give strengthened Modulation an essential role in Kant's proof.

With Changeability, however, even non-strengthened Modulation can ground the radical evil thesis. Together, they entail that there is a *possible* degree of strength of input on one's sensibility, and a *possible* degree of strength/weakness of one's empirical powers of reason and sensibility, so that sensibility exerts a stronger influence than reason on one's *Willkür*. For given Changeability and Modulation, no starting point can rule out that the strength of one's powers and sensible input might change so as to actualize this possibility in the future. A basis for universal radical evil, and the underlying truth of the "every man has his price" saying, follows:

*Possible Overpowering* (PO): The degree of influence exerted by reason on our empirical *Willkür* can be overpowered by the degree of influence exerted by sensibility.

Given Changeability, PO is not falsified even if one's sensibility cannot *currently* overpower one's reason (perhaps because one's reason is strong and sensibility's influence is only asymptotically strengthened by the strength of its inputs), since this may change over time.

PO indicates that *frailty* – the possibility of being overcome by temptation, even if one has revolutionized one's disposition and attained a good will – is the most fundamental human weakness, and the inescapable signal of our initial evil disposition.<sup>22</sup> This safeguards Kant's

---

<sup>21</sup> This worry about asymptotes arguably underlies a natural skepticism towards Kant's "every man has his price" claim: Sure we can be tempted by pains or gains, and the greater the pain or gain the greater the temptation. However, there are diminishing returns and eventually further increase is irrelevant – if there are things one will not do for a billion dollars, an additional billion will presumably not matter. The Changeability premise allows us to accept this point and nonetheless affirm universal radical evil.

<sup>22</sup> Frailty is the lowest of the three *grades* of evil recounted by Kant (RGV, 6:29f.). On my account, it inextricably pertains to all natural human beings, whereas impurity can be, and depravity necessarily is, avoided in the aftermath of a "revolution (...) in the mode of thought" (RGV, 6:47). I read frailty as encompassing not just *actual* instances of weakness of will, but also the empirical character in which the *potential* for weakness of will is present. I thus disagree with Hare's claim that "frailty is where the inclination is not only present (because we are finite) but subjectively *stronger* than the good will" (2013: 308); frailty as the "general weakness of the human heart in complying with the adopted maxims" (RGV, 6:29) refers to a general condition rather than specific instances, plausibly the condition opposed to virtue as moral *strength*. Since perfect virtue is unattainable within experience, frailty is inescapable. See further McCarty (2009: ch. 7).

thesis from an implausible quasi-empirical consequence often attributed to it: We need not say that all human beings have freely chosen to perform immoral *deeds*. One could find saint-like human beings, whose perceivable deeds provide no grounds for suspecting depravity, or even impurity, of heart. Their actions all seem to stem from mere respect for the moral law – of course, this cannot be *known* due to opacity, but a situation is possible where no evidence suggests otherwise. Kant is committed to a merely counterfactual claim: You are a priori theoretically justified not in holding that all human beings *have sold* themselves (given in to temptation), but that even the best, most saint-like human beings have a *price* for which they *would* sell themselves, i.e., *possible* sensuous inclinations of such strength as to overpower their rational wills.<sup>23</sup> As Kant suggests: “[N]owhere is a virtue which no level of temptation *can* overthrow” (RGV, 6:38f., my emphasis; cf. V-Mo/Collins, 27:434). Regardless of whether the possibility is ever actualized, the mere possibility already signals an initial evil disposition, a radical evil that can never (within experience) be fully extirpated, even by attaining a good will. This renders transcendent to theoretical cognition our concept of a *divine* (rather than natural) human being: the “elevation of such a Holy One above every frailty of human nature” (RGV, 6:64).<sup>24</sup> As Kant reportedly says in one of his lectures: “The ideal is the *prototypon* of morality.

---

<sup>23</sup> As implied by the question: “How many people who have lived long and guiltless lives may be merely *fortunate* in having escaped so many temptations?” (MS, 6:392f.; cf. V-Mo/Collins, 27:434).

<sup>24</sup> The section in question invites a potential objection to my reading: Kant seems to say that we should *not* elevate the divine human being above frailty, as this would “stand in the way of the practical adoption of the idea of such a being for our imitation” (RGV, 6:64). Thus, frailty (Palmquist 2012: 427; Pasternack 2015: 497), and even radical evil (Palmquist 2012: 428-9), must be ascribed also to the divine human being. My reply: This section of the *Religion* concerns the “objective reality” (RGV, 6:62) of the idea of a divine human being, an objective reality Kant first claims is evident “from the *practical* point of view” (RGV, 6:62, my italics). He then discusses problems associated with ascribing objective reality to it from a *theoretical* point of view. *Conceived as theoretical cognition*, a denial of frailty is illegitimate and practically unfortunate. The paragraph ends with an important footnote distinguishing the “schematism of analogy” from the “schematism of object-determination,” which “from the moral point of view (in religion) (...) has most injurious consequences” (RGV, 6:65n.). Kant’s point is not to theoretically *ascribe* frailty to the divine human being; the error lies in conceiving the divine human being through a “schematism of object-determination” at all, whether we thereby affirm (from principles of experience) or deny (from speculative reason) frailty. We can thus take at face value Kant’s characterizations of the divine human being as not subject to weakness (V-Mo/Mron II, 29:606), and as “a person free from innate propensity to evil” (RGV, 6:80n., cf. 6:80) –

A natural man can never be the ideal, for he is still always subject to weakness” (V-Mo/Mron II, 29:605).

#### 4. Culpability and Empirical Character

Readers familiar with the major interpretative strands on Kant’s practical philosophy have probably noticed that I advocate a so-called ‘battle of forces’ model of empirical choice.<sup>25</sup> Actions are empirically determined by the relative strength of the two powers influencing choice – reason and sensibility. Many Kant interpreters reject such an account as incompatible with Kant’s theory of freedom and rationality.<sup>26</sup> Let me therefore first point out that my concern has been the *empirical* power of choice, our “choice [*Willkür*] as a *phenomenon*” (MS, 6:226; cf. KrV, A549/B577, A552/B580; V-Met/Arnoldt, 29:1015). Although it is denied by some commentators,<sup>27</sup> there is strong textual evidence that Kant is committed to empirical determinism not just in the physical but also in the mental realm.<sup>28</sup> Such determinism coheres well with a ‘battle of forces’ account on the *empirical* level, especially if one takes on board the recent reevaluation of Kant’s theory of causality convincingly proposed by Watkins (2005), where causal powers take center stage. And as I will go on to argue, a ‘battle of forces’ account of empirical choice is in fact uniquely *well-suited* to accommodate Kant’s theory of freedom.

The idea is as follows: Kant’s account of noumenal freedom fits his account of an empirical battle of forces, if free agency is conceived as a capacity to *influence* the battling forces – and thereby also the outcome. Along lines recently proposed within contemporary philosophy of agency by Brian Ellis, Kantian freedom is a “meta-causal power,” the “power we have to modify

---

from a *practical* point of view. Similarly, while the Paralogisms argue against ascribing immortality to the soul on theoretical grounds, we must nonetheless ascribe immortality to our soul on practical grounds.

<sup>25</sup> Proponents of similar views include Frierson (2014); McCarty (2009); Zinkin (2006).

<sup>26</sup> See Allison (1990: 164); Ameriks (2006: 8f.); Baron (2006: 25); Moran (2012: 115-116); Reath (2006: 12f.); Wood (1999: 33).

<sup>27</sup> See Gouaux (1972); Hanna (2009); Nayak and Sotnak (1995); Westphal (2004).

<sup>28</sup> See e.g. KrV, Bxxvii, A347/B405, A798/B826; KpV, 5:96f.; Anth, 7:141, 7:231; V-Met-L2/Pöhlitz, 28:582; Refl 5662, 18:320f.

our (...) causal power profiles” (Ellis 2013: 186).<sup>29</sup> As indicated in our discussion of Changeability above, Kant conceives of our freedom as having such meta-causality: We exercise empirical causality through our empirical mental powers, but this “empirical causality itself, without the least interruption of its connection with natural causes, could nevertheless be an effect of a causality that is not empirical, but rather intelligible” (KrV, A544/B572). Empirical actions are, in one sense, inexorably produced by the mental powers working within us. But since we have the free meta-causal power to modify our empirical causality, this does not turn morality into “an empirically explainable natural phenomenon” (Reath 2006: 13). Rather, our free choice of morality – or self-love – has a real and empirically unexplainable influence *on* the phenomena, and specifically on the mental powers of the empirical subject as a natural phenomenon.<sup>30</sup> One therefore has a “*duty to govern oneself* (...) [which] involves cultivation of the mental powers to those ends with which they are collectively compatible, and constitutes, therefore, the essential in the soul’s capacity or readiness to enlarge the *facultates animi* for all moral ends, and to direct them thereunto” (V-MS/Vigil, 27:627). Our free meta-causal power can, given Changeability, carry out this duty to cultivate our powers and acquire moral strength, i.e., virtue. By influencing and directing one’s mental powers, one freely strives towards virtue rather than being at the mercy of external circumstance:

The power that the soul has over all its faculties (...) to subordinate them to its free choice, without being necessitated to do so, is a monarchy. If a man does not busy himself with this monarchy, he is a plaything of other forces [*Kräfte*] and impressions, against his choice, and is dependent upon chance and the arbitrary course of circumstances (V-Mo/Collins, 27:362).

This reading gives a reasonable delimitation of one’s *scope of responsibility* (cf. Vilhauer 2010): Each person is directly responsible for his or her mental powers, how they are cultivated and

---

<sup>29</sup> Although Ellis does not ascribe to Kant’s transcendently idealist framework, where the meta-causal power of freedom belongs to us considered as things in ourselves rather than as appearances.

<sup>30</sup> Crucially, causal powers – like causal laws but unlike empirical events and states of affairs – do not themselves require sufficient empirical causes (see Frierson 2014: 50); hence, a free capacity to influence our causal powers does not contradict Kant’s empirical determinism. Causal laws and causal powers of nature are on a different ontological level from states and events: the causal *grounds* that, given prior states of affairs, determine the subsequent states of affairs need not themselves be empirically determined. See further Indregard (2018).



used. However, on one point my reading does considerably extend our moral responsibility compared to natural expectations: Radical evil results only if our free agency is *responsible for PO*. Kant thinks that free agency could only ground PO by initially choosing an *evil* disposition. Since we know synthetically a priori that we are all subject to PO, we also know that we all initially chose to subordinate morality.<sup>31</sup> But *why* does Kant think that we, as particular persons, are responsible for PO? If it is a synthetic a priori claim, why is it not an unfortunate, inevitable fact about human existence? I will show that PO forms part of our empirical character, and for Kant, we are responsible through free choice for our empirical character.

An initial indication that we are responsible for PO is that it *ought not* to be valid. When discussing weakness of will, Kant claims that the morally good, incorporated into our maxim, is an “irresistible [*unüberwindlich*] incentive objectively” (RGV, 6:29). “Objectively” here has not a theoretical sense (the moral law is not in fact irresistible for us, as our weakness clearly shows), but a practical, normative sense: it *ought to be* irresistible.<sup>32</sup> PO is incompatible with such irresistibility, so we can deduce that PO ought not to hold. Within experience, however, we cannot rid ourselves of PO, as we can only ever attain a *finite* degree of moral strength, i.e., virtue.<sup>33</sup> We are therefore in a sense infinitely removed from how we ought to be:

The distance between the goodness which we ought to effect in ourselves and the evil from which we start is, however, infinite (...). Nevertheless, the human being’s moral constitution ought to agree with this holiness. The latter must therefore be assumed in his disposition (RGV, 6:66; cf. KpV, 5:122).

---

<sup>31</sup> Note that Kant is not attempting to explain *why* we all chose evil. This is inscrutable (see Grimm 2002: 171). But such explanation is not (*pace* Grimm) supposed to be a part of Kant’s argument: Kant’s argument for universal radical evil is completed by pointing to (necessary) features of the world of experience that allow us to infer *that* we all chose evil.

<sup>32</sup> Compare Kant's references to duties as "objectively necessary" at GMS, 4:413; MS, 6:218; V-Mo/Collins, 27:246f., 27:256; V-MS/Vigil, 27:481, 27:485f.; V-Met/Mron, 29:900; V-Met/Arnoldt, 29:1016.

<sup>33</sup> A lecture equates increasing virtue with increasing freedom (*practical* freedom to act independently of sensible inclinations): “The more a man considers a moral act to be irresistible, and the more he is compelled to it by duty, the freer he is. For in that case he is employing the power he has, to rule over his strong inclinations” (V-Mo/Mron II, 29:617).

Our initial evil choice dooms our experiential existence to a causal power profile in line with PO: “The deficiency (...) is in principle inseparable from the existence of a temporal being, [namely] never being able to become fully what he has in mind” (RGV, 6:67n.). We can subsequently enact a moral evolution in our mode of thought and establish a good *disposition*, enabling *progress* towards a virtuous empirical character, through the “gradual influence that it has on the mind” (RGV, 6:83). It is in fact always possible to perform, from morality, the action that we morally ought to perform – given Changeability our reason’s strength can always be gradually changed to overpower a specific inclination of sensibility (cf. V-Met/Mron, 29:897). But within experience the evil *propensity* in our character cannot be fully extirpated: No gradual change in the strength of our mental powers, no attained finite degree of virtue, dislodges PO – the price may gradually increase, but there will always be a price. Hence the duty to moral perfection is a “*narrow* and perfect one in terms of its quality; but it is *wide* and imperfect in terms of its degree, because of the *frailty* (*fragilitas*) of human nature” (MS, 6:446; cf. RGV, 6:74-75n.).<sup>34</sup>

PO is not just a coincidental, unfortunate fact about our worldly existence, like disease or natural disaster. As an essential yet synthetic aspect of our empirical character we are responsible for it. According to Kant, human beings have an empirical character:

The human being is one of the appearances in the world of sense, and to that extent also one of the natural causes whose causality must stand under empirical laws. As such he must accordingly also have an empirical character, just like all other natural things. We notice it through powers and faculties which it expresses in its effects (KrV, A546/B574).

Crucially, just prior to this Kant states that our “intelligible character (...) is the transcendental cause of the [empirical character]” (KrV, A546/B574); elsewhere, he similarly claims that the “empirical character is (...) determined in the intelligible character” (KrV, A551/B579). Our intelligible character determines the “law of [the empirical subject’s] causality (...) through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws” (KrV, A539/B567; see Prol, 4:346; GMS, 4:453). The intelligible character determines not just contingent actions but the *laws* of our

---

<sup>34</sup> Despite Kant’s ‘ought implies can’ principle, then, it is not possible (at least within experience) to fully rid oneself of radical evil, *even though one ought to*. Kant deals with various difficulties arising from this in the section “Difficulties That Stand in the Way of the Reality of This Idea, and Their Solution” (RGV, 6:66f.); see also 6:116f.

empirical agency, and the principle of intensive magnitude and PO as a synthetic first principle or meta-law of empirical agency thereby falls under the remit of responsibility.

This account of our moral responsibility for PO relies heavily on Kant's transcendental idealism. Specifically, it appeals to the more strongly metaphysical aspect of that doctrine, according to which the things in themselves *ground* the world of appearances. Kant understands the spontaneity of the subject as in some sense the author of the laws of nature in general: "Categories are concepts that prescribe laws *a priori* to appearances, thus to nature as the sum total of all appearances" (KrV, B163; cf. Prol, 4:319-320). What PO suggests is that theoretical reason's authorship of the general laws of nature is *already* tinged by the moral – or more precisely, *immoral* – disposition instigated by one's free power of choice.<sup>35</sup>

This reveals a potential cost of my interpretation: Is it implausible to see (one of) the most fundamental principles of theoretical reason as grounded on a contingent choice of evil? Three points can be made in response: First, any synthetic a priori interpretation of Kant's argument will say that the necessary conditions of possible experience – i.e., synthetic a priori propositions – imply universal radical evil. The implication is thus shared by all interpretations of this kind; my reconstruction merely reveals and specifies it. Second, we cannot theoretically explain why we have the forms of intuition and categories that we have (see KrV, B145-146). The idea that our free choice of intelligible character influences our theoretical principles therefore cannot be ruled out. Third, a strong case can be made that Kant himself draws the implication in question, by considering again his account of the divine human being:

The proposed reconstruction entails that a human being unsullied by radical evil – and therefore not possessing a merely finite degree of rational power – cannot be cognized within the world of experience. This fits Kant's account of the divine human being. A moral disposition so strong that it *cannot* be overcome by sensible inclinations, i.e., so strong that it violates PO, cannot be cognized theoretically: "How it is possible that the mere idea of conformity to law in general be an even more powerful incentive of that power [of choice] than any conceivable as deriving from [individual] advantages, can neither be understood by reason nor verified by examples from

---

<sup>35</sup> Although a full development of this idea lies beyond the scope of this article, one may see this as supporting Kant's controversial claims concerning the 'primacy of practical reason' (KpV, 5:119f.) by revealing practical choice as reaching into the heart of Kant's theoretical framework.

experience” (RGV, 6:62).<sup>36</sup> Theoretical real possibility requires conformity “with the formal conditions of experience (in accordance with intuitions and concepts)” (KrV, A217/B265). An inability to prove such real possibility (either a priori or from experience) thus signals that our idea of a human to whom PO does not apply fails to accord with the formal conditions of experience.<sup>37</sup> Theoretical reason or experience instead only presents this idea analogically (cf. RGV, 6:65n.), as a human being that *is not* overcome by the strongest temptations imaginable:

We cannot think the ideal of a humanity pleasing to God (hence of such moral perfection as is possible to a being pertaining to this world and dependent on needs and inclinations) except in the idea of a human being willing not only to execute in person all human duties, (...) but also, though tempted by the greatest temptation, to take upon himself all sufferings, up to the most ignominious death, for the good of the world and even for his enemies. – For human beings cannot form for themselves any concept of the degree and the strength of a force like that of a moral disposition except by representing it surrounded by obstacles and yet – in the midst of the greatest possible temptation – victorious (RGV, 6:61, cf. 6:80-82; V-Mo/Collins, 27:441f.; V-Mo/Mron II, 29:604f.).

The objective reality, i.e., the real possibility, of such an idea of the divine human being, can be established only “[f]rom the practical point of view” (RGV, 6:62). This is what we would expect if, as my reconstruction entails, the way we ought to be and therefore in some sense *could have been* is nonetheless incompatible with the a priori principles of possible experience.<sup>38</sup>

---

<sup>36</sup> See also the idea that “according to the cognition we have of the human being through experience, he *cannot* be judged otherwise [than as evil]” (RGV, 6:32, my emphasis).

<sup>37</sup> One could object that my account entails something stronger: That the concept of the divine human being lacks not just (theoretical) objective validity, but even logical *consistency*. If the divine human being takes part in the world of experience yet fails to accord with the conditions of possible experience, we must ascribe contradictory predicates to it. My response is to reiterate the importance of restricting the concept to *practical* rather than theoretical use: Theoretically asserting that the divine human being, qua divine, *takes part in the world of experience* – as the objection requires – is already an (illegitimate) determination of the concept according to the “schematism of object-determination” (RGV, 6:65n.).

<sup>38</sup> We can thereby also offer a speculative, yet perhaps still illuminating, answer to the often-expressed worry about the compatibility of universal evil with freedom: How can it be that we are free to choose between good and evil and yet somehow we *all* choose evil (see e.g. Grimm 2002: 163-164)? The proposed reconstruction suggests that while we can conceive of humans initially choosing good, these *cannot appear within experience*. Wherever these truly

## 5. Conclusion

I have proposed a novel account of the missing “formal proof” of universal radical evil in Kant. The synthetic a priori element of this proof is the principle of intensive magnitude of *KrV*, which, when applied to our mental powers and their influence on our empirical power of choice (*Willkür*), leads to the conclusion that every man has his price: We are all subject to Possible Overpowering – the possibility of sensible inclinations overpowering moral demands so as to make us choose an immoral action.

We are moreover culpable for making the initial choice of disposition that grounds Possible Overpowering as a synthetic a priori principle of the world of experience. From the practical point of view we can cognize a different possible human being, the *divine* human being, as an object of faith [*Glauben*] and as the prototype of humanity (RGV, 6:62). From the theoretical point of view, however, the “incomprehensibility of the union of this holiness with human sensible nature in the moral disposition” (RGV, 6:82) means that no theoretical cognition of our empirical character can be satisfactory or furnish sufficient proof of a human being without an evil propensity.

Recent work on Kant has highlighted the underappreciated importance of causal powers (Watkins 2005), including for his empirical psychology (Frierson 2014). This article has argued that pursuing the causal power framework further helps shed new light on the persistent problem of universal radical evil. A key innovation, enabling this further development, is interpreting our freedom as able to influence the causal powers of the empirical mind. The proposed argument for universal radical evil reveals perhaps the most fundamental, radical consequence of how, for

---

good people are, they are not – cannot be – here. (Of course, one should not ascribe “any worth or validity to such figments of the brain” (KrV, B417n.), beyond dispelling the semblance of a contradiction between freedom and the universality of evil.) Similarly, one may ask: If *your* form of cognition is sufficient for me to appear as evil, then how does *my* intelligible character contribute? How am I responsible, if your evil choice already ensures that all those who appear within experience must be radically evil? The answer is that your form of cognition *isn't* sufficient for me to appear as evil – since it isn't sufficient for me to *appear at all*. Only if *I* have chosen evil do I appear within possible experience, together with others who have made the same choice (*how* these choices coordinate to form the world of experience is beyond our knowledge).

Kant, our mental powers are grounded in freedom: A first principle of theoretical understanding, namely the principle of intensive magnitude, implies – in its application to our mental powers – radical evil, and must therefore be understood as a result of our initial evil choice of intelligible character.

## **Bibliography:**

### Kant's Works

For the *Critique of Pure Reason* I refer to the standard A/B edition pagination. Otherwise all references to Kant give volume and page from the *Akademie Ausgabe* (AA):

Kant, Immanuel. (1900–) *Kants gesammelte Schriften*. Edited by the Akademie der Wissenschaften. Berlin: Walter de Gruyter.

Translations are from:

Kant, Immanuel. (1992–) *The Cambridge Edition of the Works of Immanuel Kant*. Edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.

### Abbreviations:

Anth	Anthropologie in pragmatischer Hinsicht (AA 07)
GMS	Grundlegung zur Metaphysik der Sitten (AA 04)
KpV	Kritik der praktischen Vernunft (AA 05)
KrV	Kritik der reinen Vernunft (AA 03-04)
KU	Kritik der Urteilskraft (AA 05)
MAN	Metaphysische Anfangsgründe der Naturwissenschaften (AA 04)
MS	Die Metaphysik der Sitten (AA 06)
Päd	Pädagogik (AA 09)
ProI	Prolegomena zu einer jeden künftigen Metaphysik (AA 04)

Refl	Reflexion (AA 14-19)
RGV	Die Religion innerhalb der Grenzen der bloßen Vernunft (AA 06)
VNAEF	Verkündigung des nahen Abschlusses eines Tractats zum ewigen Frieden in der Philosophie (AA 08)
V-Anth/Busolt	Anthropologie 1788/1789 Busolt (AA 25)
V-Anth/Mensch	Anthropologie 1781/1782 Menschenkunde, Petersburg (AA 25)
V-Anth/Mron	Anthropologie 1784/1785 Mrongovius (AA 25)
V-Anth/Pillau	Anthropologie 1777/1778 Pillau (AA 25)
V-Mo/Collins	Moralphilosophie Collins (AA 27)
V-Mo/Mron II	Moral Mrongovius II (AA 29)
V-Met/Arnoldt	Metaphysik Arnoldt (K3) (AA 29)
V-Met/Dohna	Metaphysik Dohna (AA 28)
V-Met/Herder	Metaphysik Herder (AA 28)
V-Met-K2/Heinze	Metaphysik K2 (Heinze, Schlapp) (AA 28)
V-Met-L2/Pölit	Metaphysik L2 (Pölit, Original) (AA 28)
V-Met/Mron	Metaphysik Mrongovius (AA 29)
V-MS/Vigil	Die Metaphysik der Sitten Vigilantius (AA 27)
V-Phil-Th/Pölit	Philosophische Religionslehre nach Pölit (AA 28)
V-PP/Herder	Praktische Philosophie Herder (AA 27)

Other Works:

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

Ameriks, Karl. (2006) "Kant and Motivational Externalism." In H. Klemme, M. Kühn & D. Schönecker (eds.), *Moralische Motivation. Kant und die Alternativen*. Hamburg: Felix Meiner, 3-22.

- Anderson-Gold, Sharon. (1991) "God and Community: An Inquiry into the Religious Implications of the Highest Good." In P. Rossi & M. Wreen (eds.), *Kant's Philosophy of Religion Reconsidered*. Bloomington: Indiana University Press, 113-131.
- Baron, Marcia. (2006) "Overdetermined Actions and Imperfect Duties." In H. Klemme, M. Kühn & D. Schönecker (eds.), *Moralische Motivation. Kant und die Alternativen*. Hamburg: Felix Meiner, 23-37.
- Baxley, Anne Margaret. (2010) *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press.
- Buroker, Jill Vance. (2006) *Kant's Critique of Pure Reason: An Introduction*. Cambridge: Cambridge University Press.
- Caswell, Matthew. (2006) "Kant's conception of the Highest Good, the *Gesinnung*, and the Theory of Radical Evil." *Kant-studien*, 97, 184-209.
- Coxe, William. (1798) *Memoirs of the Life and Administration of Sir Robert Walpole, Earl of Oxford*. Vol. 1. London.
- Eberhard, Johann Augustus. (1794) "Ueber das Kantische radikale Böse in der menschlichen Natur." *Philosophisches Archiv*, 2.2, 34-47.
- Ellis, Brian. (2013) "The Power of Agency." In R. Groff and J. Greco (eds.), *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge, 186-206.
- Flew, Antony. (1962) "Are Ninian Smart's Temptations Irresistible?" *Philosophy*, 37.139, 57-60.
- Formosa, Paul. (2007) "Kant on the Radical Evil of Human Nature." *Philosophical Forum*, 38.3, 221-245.
- Frierson, Patrick. (2003) *Freedom and Anthropology in Kant's Moral Philosophy*. Cambridge: Cambridge University Press.
- (2005) "Kant's Empirical Account of Human Action." *Philosophers' Imprint*, 5.7, 1-34.



- (2014) *Kant's Empirical Psychology*. Cambridge: Cambridge University Press.
- Giovanelli, Marco. (2011) *Reality and Negation – Kant's Principle of Anticipations of Perception*. Dordrecht: Springer.
- Gouaux, Charles. (1972) "Kant's View on the Nature of Empirical Psychology." *Journal of the History of the Behavioral Sciences*, 8.2, 237-242.
- Grenberg, Jeanine. (2005) *Kant and the Ethics of Humility*. Cambridge: Cambridge University Press.
- Grimm, Stephen R. (2002) "Kant's Argument for Radical Evil." *European Journal of Philosophy*, 10.2, 160-177.
- Hanna, Robert. (2009) "Freedom, Teleology, and Rational Causation." *Kant Yearbook*, 1, 99-142.
- Hare, John. (2013) "The Place of Kant's Theism in His Moral Philosophy." In S. Baiasu and M. Timmons (eds.), *Kant on Practical Justification*. Oxford: Oxford University Press, 300-314.
- Indregard, Jonas Jervell. (2018) "A Gradual Reformation: Empirical Character and Causal Powers in Kant." *Canadian Journal of Philosophy*, 48.5, 662-683.
- Jankowiak, Tim. (2013) "Kant's Argument for the Principle of Intensive Magnitudes." *Kantian Review*, 18.3, 387-412.
- Kain, Patrick. (2003) "Prudential Reason in Kant's Anthropology." In B. Jacobs and P. Kain (eds.), *Essays on Kant's Anthropology*. Cambridge: Cambridge University Press, 230-265.
- Kemp, Ryan. (2011) "The Contingency of Evil: Rethinking the Problem of Universal Evil in Kant's *Religion*." In O. Thorndike (ed.), *Rethinking Kant: Volume 3*. Newcastle: Cambridge Scholars Publishing, 100-123.
- Kohl, Markus. (forthcoming) "Radical Evil as Regulative Idea." *Journal for the History of Philosophy*.

- Louden, Robert B. (2009) "Evil Everywhere: The Ordinariness of Kantian Radical Evil." In S. Anderson-Gold and P. Muchnik (eds.), *Kant's Anatomy of Evil*. Cambridge: Cambridge University Press, 93-115.
- McCarty, Richard. (2009) *Kant's Theory of Action*. Oxford: Oxford University Press.
- Moran, Kate. (2012) *Community and Progress in Kant's Moral Philosophy*. Washington: Catholic University of America Press.
- Morgan, Seiriol. (2005) "The Missing Formal Proof of Humanity's Radical Evil in Kant's Religion." *Philosophical Review*, 114.1, 63-114.
- Muchnik, Pablo. (2009) "An Alternative Proof of the Universal Propensity to Evil." In S. Anderson-Gold and P. Muchnik (eds.), *Kant's Anatomy of Evil*. Cambridge: Cambridge University Press, 116-143.
- Nayak, Abhaya C. & Sotnak, Eric. (1995) "Kant on the Impossibility of the 'Soft Sciences'." *Philosophy and Phenomenological Research*, 55.1, 133-151.
- Palmquist, Stephen R. (2008) "Kant's Quasi-Transcendental Argument for a Necessary and Universal Evil Propensity in Human Nature." *Southern Journal of Philosophy*, 46.2, 261-297.
- (2012) "Could Kant's Jesus Be God?" *International Philosophical Quarterly*, 52.4, 421-437.
- Papish, Laura. (2018) *Kant on Evil, Self-Deception, and Moral Reform*. Oxford: Oxford University Press.
- Pasternack, Lawrence. (2015) "Kant's 'Appraisal' of Christianity: Biblical Interpretation and the Pure Rational System of Religion." *Journal of the History of Philosophy*, 53.3, 485-506.
- Quinn, Philip. (1988) "In Adam's Fall, We Sinned All." *Philosophical Topics*, 16.2, 89-118.
- Rauscher, Frederick. (2006) "Reason as a Natural Cause." In H. Klemme, M. Kühn & D. Schönecker (eds.), *Moralische Motivation. Kant und die Alternativen*. Hamburg: Felix Meiner, 97-110.

————— (2015) *Naturalism and Realism in Kant's Ethics*. Cambridge: Cambridge University Press.

Reath, Andrews. (2006) *Agency and Autonomy in Kant's Ethical Theory*. Oxford: Oxford University Press.

Serck-Hanssen, Camilla. (2012) "Das radikale Böse, das Ich und die Illusion der praktischen Vernunft." In B. Dörflinger and G. Kruck (eds.), *Worauf Vernunft hinaussieht. Kants regulative Ideen im Kontext von Teleologie und praktischer Philosophie*. Hildesheim: Olms, 125-137.

Stang, Nicholas. (2016) *Kant's Modal Metaphysics*. Oxford: Oxford University Press.

Vilhauer, Benjamin. (2010) "The Scope of Responsibility in Kant's Theory of Free Will." *British Journal for the History of Philosophy*, 18.1, 45-71.

Warren, Daniel. (2001) *Reality and Impenetrability in Kant's Philosophy of Nature*. New York: Routledge.

Watkins, Eric. (2005) *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press.

Westphal, Kenneth. (2004) *Kant's Transcendental Proof of Realism*. Cambridge: Cambridge University Press.

Wood, Allen. (1991) "Unsociable Sociability: The Anthropological Basis of Kantian Ethics." *Philosophical Topics*, 19, 325-351.

————— (1999) *Kant's Ethical Thought*. New York: Cambridge University Press.

Zinkin, Melissa. (2006) "Respect for the Law and the Use of Dynamical Terms in Kant's Theory of Moral Motivation." *Archiv für Geschichte der Philosophie*, 88.1, 31-53.