

KANT, RADICAL AGNOSTICISM, AND METHODOLOGICAL ELIMINATIVISM ABOUT THINGS IN THEMSELVES

Robert Hanna
Independent Philosopher

1. Introduction

In a recent paper, ‘Things in Themselves: An Interim Report’¹, Robert Howell says that a philosophically satisfactory resolution of all-too-familiar problems about Kant’s views on the object of knowledge and the nature of things in themselves should meet three conditions:

- (1) it should be reasonably faithful to Kant’s views,
- (2) it must show that his views are internally consistent (and that his major arguments are valid),

and

- (3) it must not rest on premises that are themselves philosophically implausible.

Then Howell proceeds to examine several attempts by recent or not-so-recent analytically-oriented Kantians—Peter Strawson, Rae Langton, Henry Allison, and Desmond Hogan—to provide an account of Kant’s views that meets these conditions, and finds them all wanting in various respects.

Howell concludes that it would be philosophically good to find a satisfactory resolution of these problems, but also suspects that no such resolution will be found.

I fully agree with Howell’s three conditions of adequacy on a philosophically satisfactory resolution of the problems, and also fully agree that it would be philosophically good to find a resolution of them.

But I sharply disagree that no such resolution can be found. Indeed, I believe that I’ve found such a resolution.

In my books *Kant and the Foundations of Analytic Philosophy*,² *Kant, Science, and Human Nature*,³ and *Cognition, Content, and the A Priori*,⁴ I’ve argued for the following four Kantian doctrines:

- (1) what I call the *Two Concept or Two Property* theory of things in themselves (TC/TP),

¹ See Howell (2014).

² See Hanna (2001), esp. chs. 1–2.

³ See Hanna (2006), esp. chs. 1–4, 6, and 8.

⁴ See Hanna (2015), esp. chs. 3 and 7.

(2) what I call *Weak or Counterfactual Transcendental Idealism* (WCTI),

(3) what I call *Kantian Radical Agnosticism* about things in themselves (KRA),

and

(4) what I call *Kantian Methodological Eliminativism* about things in themselves (KME).

TC/TP says that Kant's distinction between empirically real, directly perceivable manifest objects (comprising both 'appearances' and 'objects of experience') and things in themselves (as positive noumenal objects of the understanding, i.e., *Verstandeswesen*, such that, if they were to exist, would be non-empirical, non-spatiotemporal, and constituted by intrinsic non-relational properties) is best captured by a corresponding concept-dualism and meta-physical property-dualism, with no further ontological commitments as to the instantiation or non-instantiation of positive noumenal properties.

WCTI says that Kant's transcendental idealism is best captured by the synthetic a priori counterfactual thesis to the effect that, necessarily, the manifest world exists only if, if cognisers like us (i.e., 'human' cognisers in the sense that they possess essentially the same sorts of innately specified cognitive and practical faculties or powers as we do) *were to exist*, then they *would be able to know* that world directly through sensible intuitions, objectively valid concepts, and judgements of experience, and through also analytic and synthetic a priori judgements or propositions, to some non-trivial extent.

KRA says that we can and do know a priori that we cannot know *either* the nature of things in themselves *or* whether things in themselves exist or do not exist, nor can we cognise (in the narrow, B-edition sense of *Erkenntnis*) any facts about things in themselves or their negations, nor can we prove (in the strict scientific sense) any facts about things in themselves or their negations.

And KME says that for the specific purposes of working out a philosophically adequate anthropocentric, 'human-faced' metaphysics of transcendental idealism, that is to say, WCTI, we can completely ignore things in themselves.

The arguments for WCTI are somewhat long and involved, since they also require interpretations and defences of Kant's *analytic/synthetic distinction*, his *empirical realism*, and *Kantian Non-Conceptualism*, as well as close, critical readings and interpretations of the Refutation of Idealism and the Transcendental Aesthetic.

So obviously I can't, and won't, repeat them here.

But in any case, they can be found in detail in chapters 3–5 of *Kant and the Foundations of Analytic Philosophy*; in chapters 1–4, 6, and 8 of *Kant, Science, and Human Nature*; and in chapters 2–4 and 7 of *Cognition, Content, and the A Priori*.

In the rest of this paper, then, I'll focus on unpacking and defending KRA and KME, and also indicate how they are internally related to TC/TP.

2. Kantian Radical Agnosticism about Things in Themselves

The fundamental twofold theme of the *Critique of Pure Reason* is

- (1) the critical rejection of classical metaphysics, especially including classical Rationalist metaphysics, but also including various versions of dualism and materialism or physicalism—according to Kant, metaphysical materialism must be sharply distinguished from classical Empiricism, in the sense that it is logically possible to accept materialism or physicalism and reject classical Empiricism, and conversely,

and

- (2) the revolutionary replacement of classical metaphysics by a new, inherently anthropocentric, and essentially mitigated kind of rationalist, real metaphysics: transcendental idealism.

Kant's revolutionary anthropocentric turn to the mitigated rationalism or real metaphysics of transcendental idealism in fact has *three* conjoined and equally important philosophical sources:

First, Hume's sceptical Empiricism about the content, truth, and justification of human cognition, especially as applied to the classical Rationalist metaphysical concepts of causation and causal necessity, remembered by Kant in 1771 or 1772.

Second, Kant's own revolutionary thesis in real metaphysics, transcendental idealism, about the necessary conformity of the ontic structures of apparent, phenomenal, or manifest natural spacetime and the causal-dynamic relations between manifest natural objects and states-of-affairs, to the innate mentalistic structures of human sensibility, understanding, and inferential reason, discovered and formulated by him between the *Directions in Space* essay in 1768 and 'the year of great light', 1769, and 1772, the year of Kant's famous letter to Herz.

Third, the self-annihilating character of all classical metaphysical reasoning, especially classical Rationalist metaphysical reasoning, demonstrated by the antinomy of pure reason, and thereby the possibility of the First *Critique*, discovered by Kant in 1766.

These three sources combined to produce in Kant—a middle-aged, mid-career, and middling-successful academic philosopher who was 42 in 1766 and 48 in 1772, a smallish (5 feet 2 inches tall) confirmed bachelor with mild *angina pectoris*, and a hypochondriac—an intellectual and spiritual rebirth comparable to those of Augustine, Luther, and Pascal.⁵

But unlike the earlier thinkers, Kant's personal rebirth was not brought by discovering or re-discovering God: it was brought about by discovering or re-discovering the anthropocentric, existential epiphany of *the rational human condition*.⁶

The A-Preface of the First *Critique* conveys this 'human, all too human' epiphany by means of a rhetorically quite effective and semantically pregnant *framing metaphor* based on the tragic plight of Hecuba.

⁵ See Kuehn (2001:148ff., 238).

⁶ See also Hanna (2015b).

Hecuba was a mythic queen of Troy, wife of King Priam, and the mother of Hector, Cassandra, and some forty-eight other children (some of them, presumably, the offspring of Priam's concubines), many of whom died violent deaths.

Hecuba was enslaved by the Achaeans after the Fall of Troy. Kant cites her tragic lament—"Greatest of all by race and birth, I am now cast out, powerless" (Aix)—as movingly portrayed in Ovid's *Metamorphoses*.⁷

There are other equally moving portrayals in Euripides's *Hecuba*⁸ and *The Trojan Women*⁹, and in Dante's 16th century *Inferno*, the first part of *The Divine Comedy*.

Significantly, Dante's description of Hecuba in the *Inferno* has her ultimately descending into madness, and barking like a dog.¹⁰

Kant's framing metaphor aligns Hecuba's plight with the tragic career of classical Rationalist metaphysics, the erstwhile "queen of all the sciences" (Aviii).

The truly memorable opening line of the *Critique of Pure Reason*—

Reason has the peculiar fate in one species of its cognition that it is burdened with questions that it cannot dismiss, since they are given to it as problems by the very nature of reason itself, but which also it cannot answer, since they transcend every capacity of human reason. (Avii)

tells us that there is an innate cognitive capacity, faculty, or power (*Vermögen*), namely reason or *Vernunft*, which is partially *constitutive* of human cognition, but at the same time this very faculty of reason is the inherent source of insoluble difficulties for human cognition.

More precisely, the faculty of reason in its human involvement naturally presents certain philosophical problems that lead inevitably to corresponding questions, which in turn simply cannot be answered, even in principle, since the answers would require a kind of knowledge that inherently transcends the scope of human reason itself.

So in other words, the very cognitive faculty that, according to classical rationalism, is supposed to be the strict determiner of the content, truth, and justification of necessary a priori judgements, by a wholly tragic reversal of fortune, turns out to be the primitive source of its own rational self-stultification, or cognitive suicide—in effect, a self-inflicted descent into barking madness, like poor Hecuba. How the philosophically mighty have fallen.

Kant then provides a psychologically-oriented *meta*-philosophical diagnosis of the logico-metaphysical and pragmatic self-stultification, or *cognitive suicide*, of the faculty of reason.

It is not reason's fault. It is built into the very nature of human reason that it cognises and follows principles that are adequately warranted by their application to human experience.

But the recursive application of these same principles—their consistent, constructive application to the results of previous applications of the very same principles—leads reason bey-

⁷ See Ovid (2004:508ff.)

⁸ See Euripides (1958).

⁹ See Euripides (1958b).

¹⁰ See Dante (1995, canto XXX, lines 13ff.)

and the original data of human experience, and into applications that extend unrestrictedly beyond experience, and ultimately into “obscurity and contradictions”, *Dunkelheit und Widersprüche* (Aviii).

What kind of obscurity and what kind of contradictions? Here there is a strong anticipation by Kant of the logico-metaphysical phenomenon of what contemporary philosophers of logic and mathematics call *impredicativity*, and even more precisely, of what can be called *vicious impredicativity*.

Impredicativity is the construction or definition of sets or totalities of objects in terms of, or by reference to, those very totalities themselves.

For example, the set or totality of all non-bachelors is itself a non-bachelor, and thus belongs to the membership of that very set or totality.

On the other hand, the set or totality of bachelors is not a member of the set or totality of bachelors.

So, some sets or totalities are members of themselves, and some sets or totalities are not members of themselves.

Let us call the self-membership of the set or totality of non-bachelors an instance of *benign impredicativity*, and let us call the rule which says that totalities may permissibly be constructed or defined by impredicative operations, *The Principle of Benign Impredicativity*.

Correspondingly, the construction or definition of sets or totalities according to The Principle of Benign Impredicativity seems perfectly logically and mathematically legitimate.

Vicious impredicativity, by sharp contrast, is the generation of logical paradoxes, hyper-contradictions, or instances of *dialetheia* (i.e., propositions such that, logically necessarily, they are true if and only if they are false, hence logically necessarily they are both true *and* false, aka ‘truth-value gluts’) by means of impredicative reasoning.

Here is a famous and historically highly relevant example of vicious impredicativity taken from the logical foundations of mathematics, which, for Frege in *Basic Laws of Arithmetic*, had logically catastrophic results, and almost literally Hecuba-like personal consequences.¹¹

As Frege wrote to Russell, who had discovered the contradiction in basic law V of Frege’s logical system, the unrestricted comprehension principle for naïve set theory, which says that any well-formed predicate intension uniquely determines a well-formed concept, whose extension is a well-formed set:

Your discovery of the contradiction has surprised me beyond words and, I should like to say, left me thunderstruck because it has rocked the ground on which I mean to build arithmetic [...]. It is all the more serious as the collapse of my law V seems to undermine not only the foundations of my arithmetic but the only possible foundations of arithmetic as such. (Monk 1996:153)

Russell’s discovery was not only logically catastrophic and personally tragic for Frege. It was also, from the retrospective standpoint of the First *Critique*, logico-metaphysically *ironic*. This is because Frege’s goal in *Basic Laws* was precisely to provide an explanatory and ontological reduction of arithmetic to logic and thereby, by showing that arithmetic truths are analytic truths, refute Kant’s thesis that arithmetic is synthetic a priori.

¹¹ See Frege (1964).

Nevertheless, as we have just seen, Frege’s attempt to resuscitate the project of classical Rationalist metaphysics against Kant in the guise of Logicism ended in *Dunkelheit und Widersprüche*.

Here is how it happened, bounded in a nutshell. In order to reduce arithmetic to logic, Frege presupposed and used naïve set theory as the *reducing* theory, and characterised numbers as sets of all sets whose membership can be put into one-to-one correspondence with each other (equinumerosity).

We have seen that by an application of The Principle of Benign Impredicativity to the objects of ordinary experience, there are some sets or totalities that *are* members of themselves, like the set or totality of all non-bachelors, and that there are other sets or totalities that are *not* members of themselves, like the set or totality of bachelors.

But now, by a recursion that clearly extends beyond the scope of human experience, what about the set or totality consisting of *all* sets or totalities that are not members of themselves?

Call this non-experiential set or totality *K*, and now let us try to apply The Principle of Benign Impredicativity to it.

If *K* is a member of itself, then it is not a member of itself. But if *K* is not a member of itself, then it is a member of itself.

So, by recursively applying beyond the limits of human experience what seemed to be the otherwise perfectly legitimate Principle of Benign Impredicativity, we discover to our rational dismay that *K* is a member of itself if and only if it is not a member of itself.

In other words, *Dunkelheit und Widersprüche*, and we directly bear witness to the fact of reason’s self-inflicted descent into barking madness.

This is a particularly crisp and vivid example of what Kant later in the First *Critique* calls the antinomy of pure reason, the discovery of which in 1766, as we saw above, was the third basic source of Kant’s revolutionary anthropocentric turn to the mitigated rationalism and real metaphysics of transcendental idealism.

In 1903 and 1908 Russell called instances of the antinomy of pure reason *The Contradictions*,¹² but *like* Frege and *unlike* Kant, Russell also refused to trace their generation to the innate constitution of human reason itself and its natural and ‘human, all too human’ psychological tendency to engage in what I shall call *transcendent impredicative reasoning*.

Transcendent impredicative reasoning is the iterative self-including construction of higher-order set or totalities whose first-order membership does *not* consist exclusively of elements that are actual or possible objects of human experience—leaving aside *the empty set*, or “empty intuition without an object” (B348), which is a special representation of pure intuition that is also minimally contained in every non-empty first-order set or totality whatsoever.¹³

¹² See Russell (1996 and 1971:59ff.).

¹³ See A290ff./B346ff.

Sets or totalities created by transcendent impredicative reasoning are what I shall call *ill-founded* or *noumenal* sets or totalities. This in turn leads to the corresponding notion of a *well-founded* or *phenomenal* set or totality:

A set or totality is well-founded or phenomenal if and only if either (i) every element of its membership (leaving aside the empty set) is an actual or possible object of human experience, or (ii) all its iterative self-including constructions necessarily presuppose that every element of its first-order membership (leaving aside the empty set) is an actual or possible object of human experience.

Kant ultimately argues in the Dialectic of Pure Reason that no well-founded or phenomenal set or totality is antinomous, whereas at least *some* of the ill-founded or noumenal sets or totalities, and possibly *all* of them, are antinomous.

Russell's refusal to trace the genesis of The Contradictions to the innate constitution of human reason and its natural psychological tendency to engage in transcendent impredicative reasoning, however, like Frege's similar refusal, stemmed directly from the explicitly anti-Kantian metaphysical and epistemic commitments of Russell's Logicism:

[T]he Kantian view [...] asserted that mathematical reasoning is not strictly formal, but always uses intuitions, *i.e.* the *a priori* knowledge of space and time. Thanks to the progress of Symbolic Logic, especially as treated by Professor Peano, this part of the Kantian philosophy is now capable of a final and irrevocable refutation. (Russell 1996:4)

Ever since I abandoned the philosophy of Kant [...] I have sought solutions of philosophical problems by means of analysis; and I remain firmly persuaded [...] that only by analysing is progress possible. (Russell 1959:14–5)

Just as Frege's Logicism had failed because of The Contradictions' emergence in basic law V, so too Russell's Logicism failed for a somewhat different—although not so *very* different—reason.

In 1931, Kurt Gödel proved two seminal Incompleteness theorems which show that all classical second-order logical systems, like Russell and Whitehead's system in *Principia Mathematica*, which contain enough of Peano's five axioms for arithmetic, must also contain logically unprovable sentences (in effect, self-referring versions of The Contradiction formally equivalent to the Liar Paradox) and are therefore not only

(1) incomplete (*i.e.*, not all of their tautologies are theorems),

but also

(2) consistent if and only if the ground of truth for such systems is outside the system itself.¹⁴

Gödel's seminal Incompleteness results are formally highly analogous to Kant's dialectical logical analysis of the Antinomy of Pure Reason, which shows that classical metaphysics, especially including classical Rationalist metaphysics, logically entails antinomies if and only if it is assumed that there is no fundamental ontological difference between appearances or phenomena, and things in themselves or noumena.

¹⁴ See Gödel (1967:596ff.)

Clearly, then, there is some deep and essential logico-metaphysical connection, discovered by Kant, and later re-discovered by Gödel, between collapsing the fundamental ontological difference between phenomena and noumena, transcendent impredicative reasoning and its vicious impredicativity, The Contradictions, Incompleteness, and anti-Logicism.

In any case, both Frege’s and Russell’s versions of Logicism ultimately failed because of their inability to control the logically explosive power of The Contradictions, and to constrain transcendent impredicative reasoning together with its vicious impredicativity.

To summarise the philosophical plot up to here:

Kant’s first two main points in the epiphanic A-Preface, to be reprised and worked out in full detail later in the chapter entitled The Antinomy of Pure Reason, are these:

- (1) classical metaphysics, especially including classical Rationalist metaphysics, is inherently *capable* of generating an antinomy of pure reason precisely because the innate constitution of the faculty of human reason provides for a natural psychological tendency to engage in transcendent impredicative reasoning.
- (2) the faculty of human reason on its own, without appealing to any *other* basic human cognitive faculty—e.g., the faculty of sensibility or *Sinnlichkeit*—is also inherently *incapable* of comprehending what has gone wrong in its reasoning processes when it generates an antinomy of pure reason.

Indeed, Kant even goes so far here as to *identify* metaphysics with reasoning that satisfies these two conditions: “The battlefield of these endless controversies is called *metaphysics*” (Aviii).

Now it is bad enough, and certainly philosophically tragic, that classical Rationalist metaphysics commits cognitive suicide by means of transcendent impredicative reasoning, its vicious impredicativity, and the antinomy of pure reason.

But its inevitable self-ignorance about how and why this is happening is *also* a direct violation of the original Socratic philosophical imperative to ‘know thyself’.

Hence at this point, one might well think: “Well, so much the worse for human reason, and its supposedly innate faculty for infallible clear and distinct rational intuition!”

And like Hume in the first *Enquiry*, one might well think that as a consequence we should simply *burn* all classical Rationalist metaphysics *at the stake*:

When we run over libraries, persuaded of these [Empiricist] principles, what havoc must we make? If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask, *Does it contain any abstract reasoning concerning quantity or number?* No. *Does it contain any experimental reasoning concerning matter of fact and existence?* No. Commit it then to the flames: for it can contain nothing but sophistry and illusion. (Hume 1977:165)

But although in 1771 or 1772 Kant was indeed philosophically awakened and enlightened by remembering Hume, nevertheless he was never *mesmerised* by Hume’s sceptical Empiricism.

So the very fact that human reason has this ‘peculiar fate’ (*besondere Schicksal*) driving it towards cognitive self-annihilation, according to Kant, also ultimately drives self-critical human reason through and beyond Hecuba’s tragedy towards a tragic rational *catharsis* in Aris-

tote's sense in the *Poetics*, that is to say, towards a purging of otherwise harmful emotions for the sake of self-knowledge.

And by means of this process of rational catharsis, as self-clarifying, self-critical, self-knowing metaphysicians and knowers, we rid ourselves of the very *need* to transcend human experience, and fully accept our 'human, all too human' finitude.

According to Kant we thereby become what I call *mitigated rationalists*, and *real metaphysicians*, not Humean Empiricist *mitigated sceptics* about human reason.

Back now to unfortunate Hecuba, fallen 'queen of all the sciences', namely classical metaphysics and especially classical Rationalist metaphysics.

Within this metaphorical frame, now rhetorically extended to an analogy with the political history of pre-Enlightenment despotism, Kant sketches an apocalyptic history of the rise and fall of 17th and 18th century European metaphysics.

The basic details are these:

The metaphysical theories of classical 17th or early 18th century Rationalists like Descartes, Spinoza, Leibniz, Wolff, and Baumgarten were dogmatic and needlessly strict ("despotic," and retaining "traces of ancient barbarism" [Aix]). As a consequence, classical Rationalism was wide open to highly effective critical attacks by a few Empiricist sceptics, especially Hume ("nomads" [Aix]).

But although the Rationalists did attempt to reply to these attacks, they did so only in a piecemeal way, and without any decisive reply to the sceptics ("continually attempting to rebuild, though never according to a plan unanimously accepted among themselves" [Aix]). The result was philosophical confusion and endless controversy.

For a brief period in the late 17th century, however, it seemed that Locke's 1689 *Essay Concerning Human Understanding*¹⁵ might resolve the controversies, by the application of its positive Empiricist epistemology and philosophical psychology to the analysis of the human faculty for understanding or *Verstand* ("a certain *physiology* of the human understanding" [Aix]).

Yet although Locke's *Essay* plausibly argued that necessarily, all human cognition has its causal origins in human experience ("the birth of the purported queen was traced to the rabble of common experience" [Aix]), his overall positive Empiricist line of argument ultimately failed because of some basic flaws ("this *genealogy* was attributed to her falsely" [Aix]), and classical Rationalist metaphysics was able simply to re-assert its basic principles and stubbornly hold the line ("metaphysics fell back into the same old worm-eaten *dogmatism*" [Ax]).

The result of this unresolved Rationalist-Empiricist dialectic was that by the present time—that is to say, the latter half of the 18th century—many or even most uncommitted philosophers were intellectually bored to death by the whole enterprise and simply refused to take sides, or to undertake any fruitful or new philosophical work ("what rules is tedium and complete *indifferentism*" [Ax]), so that the science of metaphysics was, in effect, going to hell in a handbasket ("the mother of chaos and night in the sciences" [Ax]).

¹⁵ See Locke (1975).

Kant's notion of 'complete indifferentism' is what I shall call *zombie-like agnosticism*, that is to say, the intellectually lifeless, mechanical and puppet-like, empty-headedly and heartlessly rote refusal to advance from cognitive neutrality to any sort of deeper, responsibly-held belief or authentic commitment. In other words, it is *cognitive inauthenticity* in the Existentialist sense.

Generally speaking, however, agnosticism is the negative doxic attitude of *comprehensive non-belief* or *suspension of judgement*.

More specifically characterised, a cognitive subject *X* is agnostic or indifferent that *P* if and only if *X* does not believe that *P* and *X* does not believe that not-*P*. As characterised in this way, agnosticism is then to be contrasted with the slightly more limited negative doxic attitude of *doubt*. More specifically, a cognitive subject *X* doubts that *P* if and only if either *X* does not believe that *P* or *X* believes that not-*P*.

Thus every form of agnosticism includes doubt, but not all forms of doubt are agnostic.

Both Locke and Hume had fruitfully explored the negative doxic attitudes of agnosticism and doubt, but Kant pushes those explorations significantly beyond those of the classical Empiricists.

I have already characterised 'complete indifferentism', aka *zombie-like agnosticism*, as the intellectually lifeless, mechanical and puppet-like, empty-headedly and heartlessly rote refusal to advance from cognitive neutrality to any sort of deeper, responsibly-held belief or authentic commitment.

Another distinct form of agnosticism, which for clarity's sake I shall dub *constructive agnosticism* or *indifferentism*, consists in a doxic attitude of reasonable cognitive neutrality or open-mindedness, and in a cautious refusal to take a stand until all the relevant evidence is in.

Constructive agnosticism or indifferentism follows naturally from the act or process of reasonable doubt that Hume characterises as *mitigated scepticism*.¹⁶ Mitigated scepticism leading to constructive agnosticism or indifferentism is fully appropriate in everyday reasoning, exact science, and legal contexts, for example, in an ordinary well-conducted court of law.

In such contexts, the presiding judge should deploy reasonable doubt and remain in a doxic attitude or state that is constructively agnostic or indifferent until it is appropriate to form a judgement, given all the relevant evidence.

This, in turn, is sharply opposed to what I shall call *destructive agnosticism* or *indifferentism*, which consists in a deeply regressive doxic attitude of cognitive paralysis or self-annihilating confusion in the face of 'obscurity and contradictions', essentially similar to what the Stoic radical sceptics called 'equipollence'.

The end result of destructive agnosticism or indifferentism is *doxic nihilism*, that is to say, not only the refusal to proceed in appropriate evidential circumstances from cognitive neutrality or open-mindedness to judgement, but also the serious attempt to devastate and undermine any attempt to break out of this state of cognitive neutrality.

¹⁶ See Hume (1977:102ff.)

Doxic nihilism is therefore one giant step beyond complete indifferentism or zombie-like agnosticism. It is when the latter turns into a fully devolutionary and regressive cognitive counterforce—zombies with razor-sharp teeth, hence flesh-eating, intellect-eating monsters.

Destructive agnosticism or indifferentism follows naturally from the corrosive and radical form of scepticism that Hume dubs *Pyrrhonian* scepticism, which seeks to undermine the grounds of all rational belief, and which in fact is *a direct attack on human rationality itself*.¹⁷

A good contemporary example of the doxic nihilism of destructive agnosticism or indifferentism is the current craze for *debunking strategies* and *error theories* in early 21st century mainstream Analytic philosophy.

As we have already seen, the underlying cause of destructive agnosticism or indifferentism, and equally of the pathological mechanism of Pyrrhonian radical equipollence scepticism in its application to classical metaphysics, is *transcendent impredicative reasoning, its vicious impredicativity, and the antinomies of pure reason*.

So Kant is saying that this is the intellectual killer-zombie abyss and apocalypse where metaphysics as a science *could* end up forever, if some radical metaphysical therapy is not carried out, *fast*.

According to Kant, moreover, this radical metaphysical therapy turns on a *third* form of indifferentism or agnosticism, that is to say, a radical form of non-belief or suspension of judgement that is in sharp contrast to *either* zombie-like agnosticism, constructive agnosticism, *or* destructive agnosticism.

This essentially more robust negative doxic attitude, or what I shall call *radical* agnosticism or *radical* indifferentism, is in fact a higher-order state of negative self-knowledge—that is to say, of *knowing that it is impossible for me to know some first-order object or state of affairs one way or the other*.

This can be more specifically spelled out. Call any such first-order object, *X*, and call any such first-order state of affairs, the state of affairs that *P*. Then radical agnosticism about *X* and the state of affairs that *P* is knowing that it is impossible for me to know whether it is either true of *X* that *P* or false of *X* that *P*.

Kant holds that the mature rational human faculty or power of judgement, or *Urteilkraft*, is capable of advancing to radical agnosticism about the ‘illusory knowledge’ (*Scheinwissen*) dogmatically postulated by classical metaphysics, and especially classical Rationalist metaphysics.

That is, our faculty of judgement can advance to the higher-order negative self-knowledge that it is impossible for us to know any objects or states of affairs *beyond the limits of human experience*, one way or the other.

For example, we can have the higher-order negative self-knowledge that it is impossible for us to know whether an omnipotent, omniscient, and all-good (aka ‘a 3-O’) God exists or does not exist, whether noumenal agent-causal libertarian freedom exists or does not exist, and whether immortality of the soul exists or does not exist.

¹⁷ See Hume (1977:102ff.)

Similarly, as I noted above—and not altogether coincidentally, in view of what we already know about the explicitly anti-Kantian project of Logicism—Gödel proved in 1931 that there are unprovable sentences in classical second-order logic plus enough of Peano’s five axioms for arithmetic.

Consider, now, any such unprovable sentence of Peano arithmetic.¹⁸ By means of Gödel’s Incompleteness proof, we have the higher-order logical self-knowledge that it is impossible for us to know, *by means of proof-theoretic means alone*, whether this sentence is true or false.

The inherently reasonable process of mitigated scepticism leads to constructive agnosticism or indifferentism, and the inherently *anti*-reasonable process of Pyrrhonian radical scepticism, that is to say, destructive agnosticism or indifferentism, leads to doxic nihilism.

For Kant, the *paradigmatically philosophical* rational process of self-criticism that leads to the higher-order negative self-knowledge of radical agnosticism or indifferentism is nothing more and nothing less than *the critique of pure reason*.

This rational process of self-criticism is *paradigmatically philosophical* precisely because it fully captures and also fully clarifies the epistemic force of *Socratic Ignorance*:

The truly Critical philosopher knows only that she *cannot* know, as to either their existence or their nature, one way or the other, all and only those objects and states of affairs that transcend the limits of human experience, that is to say, all and only *noumena* or *things in themselves*.

3. Kantian Methodological Eliminativism about Things in Themselves

Kantian Methodological Eliminativism about things in themselves (KME) says that although things in themselves are logically possible, nevertheless precisely because we cannot cognise them, then we also know a priori that we cannot know whether they really do exist or do not exist, and therefore for the constructive purposes of real metaphysics, we can completely ignore things in themselves.

In order to defend KME, in turn, I must look more closely at transcendental idealism’s seminal distinction between appearances/phenomena and things in themselves/noumena.

This is without a doubt the most controversial, difficult, and written-about issue about the *First Critique*, not only because resolving it is essential for understanding the nature of transcendental idealism and for critically evaluating its philosophical truth or falsity, but also because, sadly, Kant himself *never fully resolved it*, and more or less systematically oscillated between at least three different versions of the distinction:

- (1) *The Two Object or Two World Theory*, which is a specifically Kantian version of *substance dualism*,
- (2) *The Two Aspect or Two Standpoint Theory*, which is a specifically Kantian version of ontological monism, that is to say, a One World Theory, together with an *epistemic dualism*, and

¹⁸ See Peano (1967:83ff.).

- (3) *The Two Concept or Two Property Theory*, aka TC/TP, which is also a specifically Kantian version of ontological monism or One World theory, together with a *property-dualism-without-substance-dualism*.¹⁹

One thing we *do* know is that appearances or phenomena, are manifest natural objects such that

- (1) necessarily, they conform to the non-empirical structures of our innate cognitive faculties, hence necessarily, they are related to human minds, and in at least *that* sense they are mind-dependent,
- (2) necessarily, they either are or can be directly present to human intuition and sensibility, and also fall under human concepts and understanding, hence necessarily, they are sensory and have sensory properties,
- (3) necessarily, they are spatiotemporal,
- (4) necessarily, they are causally empowered under natural laws, and
- (5) necessarily, their existence and specific character is conditioned by causes that spatially exist earlier in time.

And, since things in themselves or noumena are, by hypothesis, *essentially distinct* from appearances or phenomena, it follows that, *if they really exist*, then they are non-manifest, non-natural objects such that

- (1*) necessarily, they do not conform to the non-empirical structures of our innate cognitive faculties, hence necessarily, they are not related to human minds, and in at least *that* sense they are mind-independent,
- (2*) necessarily, they either are not or cannot be directly present to human intuition and sensibility, and also fall under human concepts and understanding, hence necessarily, they are *non-sensory* and have *non-sensory* properties,
- (3*) necessarily, they are *non-spatiotemporal*,
- (4*) necessarily, they are *not* causally empowered under natural laws, and
- (5*) necessarily, their existence and specific character is *unconditioned* by any causes that spatially exist earlier in time.

Nevertheless, these five features occur explicitly under a conditional existential supposition, and we therefore need to know whether, according to Kant, things in themselves really do exist or not. Let us call this *The Real Existence Question*.

Here is what Kant tells us about *The Real Existence Question*: “Our rational cognition *a priori* [...] reaches appearances only, leaving the thing in itself [*Sache an sich selbst*] as something actual [*wirklich*] for itself, but uncognized by us” (Bxx). Unfortunately, this formulation is at least four ways ambiguous and does not tell us whether Kant is saying that

- (1) things in themselves *do* really exist but are *uncognised* by us,

¹⁹ See Hanna (2006:195–6, 425ff.)

- (2) things in themselves *do* really exist but are *uncognisable* by us,
- (3) things in themselves *possibly* really exist but *cannot be known* either really to exist or not exist precisely because they are *uncognised* by us, or
- (4) things in themselves *possibly* really exist but *cannot be known* either really to exist or not exist precisely because they are *uncognisable* by us.

Only formulation (4) adequately captures the cognitive attitude of Kantian Radical Agnosticism about things in themselves, aka KRA, that I spelled out and defended in section 2.

In my opinion, formulations (1) to (3) each have apparently insuperable philosophical problems associated with them. But for reasons of space limitation, I shall not go into these worries here. In any case, in the First *Critique*, Kant has not decisively plumped for any one of them.

Finally, the other thing that Kant tells us here is that

[T]he same objects can be considered from two different sides, *on the one* side as objects of the senses and the understanding for experience, and *on the other side* as objects that are merely thought at most for isolated reason striving beyond the bounds of experience. If we now find that there is agreement with the principle of pure reason when things are considered from this twofold standpoint, but that an unavoidable conflict of pure reason with itself arises with a single standpoint, then the experiment decides for the correctness of that distinction. (Bxviii–xix n.)

This is a classic statement of The Two Aspect or Two Standpoint Theory of the appearances/phenomena versus things in themselves/noumena distinction.

It is also a classic statement of the One World or ontologically monistic interpretation of the appearances/phenomena versus things in themselves/noumena distinction.

Again sadly, however, it is at least *three* ways ambiguous, since it does not tell us whether Kant is saying that:

- (1) there is one and only one class of apparent or phenomenal objects (i.e., one apparent or phenomenal world, hence *phenomenal monism*) such that each member of this class can be considered phenomenally or noumenally by us,
- (2) there is one and only one class of things in themselves or noumenal objects (i.e., one in-itself or noumenal world, hence *noumenal monism*) such that each member of this class can be considered phenomenally or noumenally by us, or
- (3) there is one and only one class of non-phenomenal, non-noumenal objects (one non-phenomenal, non-noumenal world, hence *neutral monism*) such that each member of this class can be considered phenomenally or noumenally by us.

Even granting the fourfold ambiguity of what Kant has said explicitly about *The Real Existence Question*, given what he *has* explicitly said about it, then it seems extremely unlikely that formulation (3) captures the correct interpretation of The Two Aspect or Two Standpoint Theory.

For in that connection, nothing whatsoever has been said by Kant about any *third* or neutral class of objects apart from the class of appearances or phenomena and the class of things in themselves or noumena.

Nevertheless, again granting that fourfold ambiguity, unfortunately, formulation (1) and formulation (2) are both consistent with what Kant has explicitly said.

Hence either of them, so far, could be a correct interpretation of The Two Aspect or Two Standpoint Theory.

It must also be frankly noted that there is also the very real critical question of whether the The Two Aspect or Two Standpoint Theory is philosophically acceptable under *any* reasonable interpretation of it. But again owing to space limitations, I shall not, in this connection, offer a definite answer to that question. Still it is crucial to note here that the One World or monistic interpretation of the appearances/phenomena versus things in themselves/noumena distinction is itself *logically independent* of The Two Aspect or Two Standpoint Theory.

For it is perfectly consistent to hold, as a version of the One World or monistic interpretation, that

- (1¹) there is one and only one class of apparent or phenomenal objects (i.e., one apparent or phenomenal world, hence phenomenal monism), and
- (1²) each member of the one and only class of apparent or phenomenal objects can be considered phenomenally by us, yet
- (1³) it makes no rational sense to consider apparent or phenomenal objects as in-themselves or noumenally, given the ontologically disjoint character of the classes of appearances/phenomena and things in themselves/noumena, hence no member of the one and only class of phenomenal objects can ever be intelligibly considered noumenally by us, but
- (1⁴) at the same time, we *can consistently think about*, although *never cognise and never know*, things in themselves or noumena.

In this way one can consistently deny The Two Aspect or Two Standpoint Theory and also accept a phenomenally monistic version of the One World interpretation of the appearances/phenomena versus things in themselves/noumena distinction.

It should also be noted that (1¹) through (1⁴) are all smoothly consistent with formulation (4) of what Kant has explicitly said about *The Real Existence Question*, which in turn adequately captures the cognitive attitude of radical agnosticism about things in themselves.

So even leaving aside the question of whether Kant *himself* holds (1¹) through (1⁴) as his all-things-considered version of the One World interpretation of the appearances/phenomena versus things in themselves/noumena distinction, together with formulation (4)'s version of an answer to *The Real Existence Question*, together with radical agnosticism about things in themselves, we could still plausibly assert that Kant rationally *ought to have held* all these views, and also that their conjunction constitutes a philosophically defensible and recognisably Kantian, or at least contemporary Kantian, philosophical theory that combines KRA, KME, and TC/TP.

BIBLIOGRAPHY

- Dante, A. (1955) ‘The Inferno’, in D. L. Slayers (trans.) *The Divine Comedy* in 3 vols. (Harmondsworth: Penguin).
- Euripides (1958) ‘Hecuba’, in W. Arrowsmith (trans.) *Euripides III* (Chicago: University of Chicago Press), pp. 9–68.
- (1958b) ‘The Trojan Woman’, in R. Lattimore (trans.) *Euripides III* (Chicago: University of Chicago Press), pp. 127–175.
- Frege, G. (1964) *The Basic Laws of Arithmetic*, trans. M. Furth (Berkeley/Los Angeles: University of California Press).
- Gödel, G. (1967) ‘On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems’, in J. Van Heijenoort (ed.) *From Frege to Gödel* (Cambridge, MA: Harvard University Press), pp. 596–617.
- Hanna, R. (2001) *Kant and the Foundations of Analytic Philosophy* (Oxford: Oxford University Press).
- (2006) *Kant, Science, and Human Nature* (Oxford: Oxford Univ. Press).
- (2015) *Cognition, Content, and the A Priori: A Study in the Philosophy of Mind and Knowledge* (Oxford: Oxford University Press).
- (2015b) ‘Preface and Introduction to The Rational Human Condition’, Summer 2015. https://www.academia.edu/12512529/Preface_and_General_Introduction_to_The_Rational_Human_Condition_Summer_2015_version_comments_welcomed_.
- Howell, R. (2014) ‘Things in Themselves: An Interim Report’. https://phil.hse.ru/data/2016/04/12/1126937318/howell_1.pdf.
- Hume, D. (1977) *Enquiry Concerning Human Understanding* (Indianapolis IN: Hackett).
- Kuehn, M. (2001) *Kant: A Biography* (Cambridge: Cambridge Univ. Press).
- Locke, J. (1975) *Essay Concerning Human Understanding* (Oxford: Oxford University Press).
- Monk, R. (1996) *Bertrand Russell: The Spirit of Solitude* (London: Jonathan Cape).
- Ovid (2004) *Metamorphoses*, trans. D. Raeburn (New York/London: Random House/Penguin).
- Peano (1967) ‘The Principles of Arithmetic, Presented by a New Method’, in J. van Heijenoort (ed.) *A Source Book in Mathematical Logic, 1879–1931* (Cambridge: Harvard University Press), pp. 83–97.
- Russell, B. (1959) *My Philosophical Development* (London: Allen & Unwin).

- (1971) ‘Mathematical Logic as Based on the Theory of Types’, in B. Russell, *Logic and Knowledge* (New York: G.P. Putnam’s Sons).
- (1996) *The Principles of Mathematics*, 2nd ed. (New York: W. W. Norton).