

ON THE FOURFOLD ROOT OF KANT'S PRINCIPLE OF SUFFICIENT REASON

Robert Hanna

Independent philosopher, USA

That *every proposition must have a reason* is the logical (formal) principle of cognition, which is subordinated to, and not set beside, the principle of contradiction. That *every thing must have its ground* is the transcendental (material) principle, which no one has ever proven or will prove by means of the principle of contradiction (and in general from mere concepts without relation to sensory intuition). It is clear enough, and has been stated countless times in the *Critique*, that a transcendental principle must determine something *a priori* in regard to objects and their possibility; consequently, it does not, like the logical principles (which abstract completely from everything concerning the possibility of the object), merely concern itself with the formal conditions of judgment. (Üe, 8193)

1. Introduction

The title of my essay is of course a shameless rip-off of the title of Schopenhauer's first book, *On the Fourfold Root of the Principle of Sufficient Reason* (1788). But now, having shamelessly used this allusion to Schopenhauer's book as philosophical click-bait, actually from here on in I'm going to be talking exclusively about *Kant's* version of the principle of sufficient reason (aka *principium rationis sufficientis*, aka *Satz des zureichenden Grundes*).

As the text I've quoted as this essay's epigraph—lifted from Kant's late-breaking polemical essay from 1790, “On a Discovery Whereby Any New Critique of Pure Reason is to be Made Superfluous by an Older One”—clearly shows, for him, the principle of sufficient reason, aka the PSR, is both

(i) *logico-semantic*, and thus a principle of the determination of *propositions*, especially including *true* propositions, and in turn, especially including *necessary truths* and *necessary rules or laws*, especially including valid or sound *arguments* or *proofs*, and also

(ii) *real-metaphysical*, and thus a principle of the determination of *beings* in a very broad sense that includes individual things, states of individual things, properties, relations, facts, events, processes, rational agents or persons, choices, and acts.

One implication of this is that Kant thinks that there are different kinds of *sufficiency* expressed by the PSR: e.g., the kind of sufficiency whereby propositions, especially true propositions, necessarily generate or produce other propositions, especially other true propositions; and, distinct from that, the kind of sufficiency whereby beings necessarily generate or produce other beings. So Kant is what I'll call a *sufficiency pluralist* with respect to the PSR.

And another implication of this is that Kant thinks that there are different kinds of *reasons* expressed by the PSR: for example, reasons that are propositions; reasons that are facts, especially *normative* facts; and reasons that are beings, especially including *material things other than human*

persons, but also and most especially, *human persons*. Indeed, as we'll see below, for Kant, a human person, via her free will and practical agency, involving both *desires* (aka "internal reasons") and *ends* (aka "external reasons"), can herself *be* a sufficient reason—an *ultimate author, creator, or source*—of her own choices and acts. So Kant is also what I'll call a *reasons-pluralist* with respect to the PSR.

Correspondingly, my basic aim in this essay is to argue that for Kant, the PSR has *four* importantly distinct, co-primitive, and mutually irreducible—hence basic—*modes*:

(i) *analytic or logical sufficiency and reasonhood*, applying to analytical or logical truths or laws, logically valid arguments/proofs, and logical consequence or entailment,

(ii) *mathematical sufficiency and reasonhood*, applying to mathematical truths, axioms, or laws, mathematically valid arguments/proofs, and mathematical consequence or entailment,

(iii) *causal-physical-mechanical sufficiency and reasonhood*, applying to deterministic, mechanistic natural laws and states of individual things, properties, relations, facts, events, and processes, and finally

(iv) *causal-deontological-spontaneous sufficiency and reasonhood*, applying to non-deterministic, non-mechanical, non-instrumental moral and practical laws, free choices and free acts, and to the autonomous willing of human persons and the practical agency of rational human animals more generally.

So Kant is what I'll call a *fourfold-root pluralist* with respect to the PSR, in addition to his being *both* a sufficiency-pluralist and *also* a reasons-pluralist with respect to the PSR. In other words, in Kant's hands, the PSR is an emphatically *robust* and thoroughly *non-reductive* principle; and therefore my interpretation of Kant's doctrine of the PSR is in sharp and in fact diametric contrast to Béatrice Longuenesse's well-known interpretation, according to which Kant is a *deflationist* or *reductivist* (or in her terminology, a "deconstructionist") about the PSR.¹

Finally, I'll also argue that for Kant, all four basic modes (roots) of sufficiency and reasonhood under the PSR, whether analytic-logical, mathematical, or real-metaphysical (*including* causal-physical-mechanical, as well as causal-deontological-spontaneous), are *inherently and irreducibly categorically normative* and also that this categorical normativity derives from the categorical normativity of *pure general logic*. Now by "normativity", I mean *that which directly concerns ideals, standards, and values*, especially *human* ideals, standards, and values; and by "categorical normativity", I mean normativity that's at once *non-instrumental* and is also grounded by the moral law in Kant's sense, aka the *Categorical Imperative*. Thus the logical-normativity-argument, in turn, will yield a specifically and uniquely *Kantian* solution to the well-known *circularity* or *reflexivity* problem with respect to the PSR, which consists in asking

(i) whether the PSR can be *applied to itself*, or not, and

(ii) if it *can* be applied to itself, *how* this self-application can obtain without *vicious circularity* or *paradoxical reflexivity*.

¹ Longuenesse, (2001).

2. Some Preliminaries:

A Priori-A Posteriori, Analytic-Synthetic, and Synthetic A Priori

Before I get to the philosophical main event, however, we'll need to make some preliminary remarks about Kant's famous or notorious a priori – a posteriori and analytic-synthetic distinctions, and also about his conception of the synthetic a priori.

The notion of “cognitive content” for Kant has two sharply distinct senses:

(i) intension or *Inhalt*, which is objective and representational (semantic content), and

(ii) sensory matter or *Materie*, which is subjective and non-representational, reflecting only the immediate conscious response of the mind to the external impressions or inputs that trigger the operations of the faculty of sensibility (phenomenal qualitative content) (A19-20/B34, A320/B376).

To be sure, for Kant just as for the Empiricists, all cognition “begins with” (*mit ... anfang*) the raw data of sensory impressions. But in a crucial departure from Empiricism and towards what might be called a *mitigated rationalism*, Kant also holds that not all cognition “arises from” (*entspringt ... aus*) sensory impressions: so for him, a significant and unique contribution to both the form and the objective representational content of cognition arises from the innate spontaneous cognitive capacities (*CPR* B1). This notion of cognition's “arising from” either sensory impressions or innate spontaneous cognitive capacities can best be construed as a strict determination relation (similar to what is nowadays called “strong supervenience”) such that *X* strictly determines *Y* if and only if the *X*-features of something are sufficient for its *Y*-features, and there cannot be a change in anything's *Y*-features without a corresponding change in its *X*-features. This allows us to say that a cognition is a posteriori, empirical, or *dependent on sensory impressions and/or contingent natural objects or facts* just in case it is strictly determined in its form or in its semantic content by sensory impressions and/or contingent natural objects or facts; but a cognition is a priori, non-empirical, or *absolutely independent of all sensory impressions and/or contingent natural objects or facts* just in case it is not strictly determined in its form or in its semantic content by sensory impressions and/or contingent natural objects or facts and is instead strictly determined in its form or in its semantic content by our innate spontaneous cognitive faculties (*CPR* B2-3). It should be noted that the apriority of a cognition in this sense is perfectly consistent with all sorts of associated sensory impressions and also with the actual presence of sensory matter in that cognition, caused by contingent natural objects or facts, so long as neither the form nor the semantic content is strictly determined by those sensory impressions and/or contingent natural objects or facts. “Pure” a priori cognitions are those that in addition to being a priori or absolutely independent of all sensory impressions and/or contingent natural objects or facts, also contain no sensory matter whatsoever (B3). So in other words, some but not all a priori cognitions are pure.

Applying these notions to judgments or propositions, it follows that a judgment or proposition is a posteriori if and only if either its logical form or its propositional content is strictly determined by sensory impressions and/or contingent natural objects or facts; and a judgment or proposition is a priori if and only if neither its logical form nor its propositional content is strictly determined by sensory impressions and/or contingent natural objects or facts, and both are instead strictly determined by our innate spontaneous cognitive faculties, whether or not that cognition also contains sensory matter. Kant also holds that a judgment or proposition is a priori if and only

if it is necessarily true (Axv, B3-4, A76/B101). This strong connection between necessity and apriority expresses

(i) Kant's view that the contingency of a judgment is bound up with the modal dependence of its semantic content on sensory impressions and/or contingent natural objects or facts, i.e., its aposteriority (B3),

(ii) his corresponding view that necessity is equivalent with strict universality or *strengte Allgemeinheit*, which he defines in turn as a proposition's lack of any possible counterexamples or falsity-makers (B4), and

(iii) his further view that necessity entails truth (A75-76/B100-101).

Moreover Kant explicitly holds that not only do a priori judgments really exist in various sciences, including physics and legitimate or real (i.e., transcendental idealist) metaphysics, but also that there really are some *pure* a priori judgments, e.g., in mathematics (B4-5, B14-18).

Now what about the analytic-synthetic distinction, aka the A-S distinction?

Kant's theory of the A-S distinction gradually emerged from the long and winding road of his philosophical struggles with classical Rationalism in general, and with the Leibnizian-Wolffian version of the principle of sufficient reason in particular, between 1755 and 1781, the year of the publication of the first or A edition of the *Critique of Pure Reason*. After 1781, Kant did indeed make some substantive additions and modifications to his theory of the A-S distinction, but his objections to classical Rationalism and the Leibniz-Wolff doctrine remained essentially unchanged after that.

In his "New Elucidation of the First Principles of Metaphysical Cognition" in 1755, "The Only Possible Argument in Support of a Demonstration of the Existence of God" in 1763, the "Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality", aka the "Prize Essay", in 1764, "On the Form and Principles of the Sensible and Intelligible Worlds", aka the "Inaugural Dissertation", in 1770, in the two editions of the *Critique of Pure Reason* in 1781 and 1787, especially in the "Antinomies of Pure Reason" and "Ideal of Pure Reason" sections of the Dialectic of Pure Reason, and finally in "On a Discovery" in 1790, Kant repeatedly engaged critically with the Leibniz-Wolff doctrine of the PSR, and more generally, consistently doubted and rejected the thesis that there is a single, univocal, universal criterion or principle of truth and existence.

Throughout all these texts—which span Kant's pre-Critical (1755-1770), proto-Critical (1768-1781), Critical (1781-1787), and post-Critical (1788-1804) periods²—Kant's fundamental reason for doubting and rejecting that there is or ever can be such a single, univocal, universal criterion or principle of truth and existence is his thesis that the first principles and other truths of logic, mathematics, real metaphysics, and natural science are *all irreducibly distinct from one*

² My division of Kant's writings into *four* distinct periods (pre-Critical, proto-Critical, Critical, and post-Critical) periods is controversial. This is true not only with respect to the identification of a specifically *proto*-Critical period that begins with the crucial essay 'Directions in Space' in 1768, but also and especially with respect to narrowing the Critical period to the six years between the A and B editions of the first Critique, and thereby identifying a distinctively *post*-Critical period from 1787 to Kant's death in 1804, during which he makes a series of substantive additions and modifications to his Critical system. But in support of this fourfold division, see, e.g., Hanna (2016; 2018a); Hall (2014); and Thorndyke (2018).

another and nonequivalent with one another, logico-semantically, epistemically, and metaphysically.

Correspondingly, in the *Critique of Pure Reason*, Kant explicitly argues for *three* irreducibly distinct and nonequivalent *basic types* of truths:

- (i) analytic or logical a priori truths,
- (ii) synthetic a priori truths, and
- (iii) synthetic a posteriori truths—for example, empirical truths of natural science.

And in the first *Critique* and elsewhere, Kant also argues that under or within the basic type of synthetic a priori truths—i.e., basic type (ii)—there are *four* irreducibly distinct and nonequivalent basic *sub-types* of synthetic a priori truths:

- (ii.1) mathematical truths,
- (ii.2) real metaphysical truths,
- (ii.3) true nomological judgments about *the natural physical laws* of the manifestly real world, and finally
- (ii.4) true nomological judgments about *the natural practical laws* of the “human, all-too-human” world of human persons or rational human animals, namely,
 - either (ii.4.1) *the moral law(s)* of ethics, aka “the doctrine of virtue,”
 - or (ii.4.2) *the civil-juridical laws* of politics, aka “the doctrine of right.”

This Kantian scheme of types of truth is already fairly complicated; but it’s made even *more* complicated by the fact that Kant *also* formulates *two further pairs of distinctions* that cut across these three basic types as well as across the four basic sub-types under or within the class of synthetic a priori truths.

First, there’s a distinction between

- (i) what Kant calls *the nominal definition* of truth, that is, the trivially correct analytic definition of the concept of truth, namely that all truth whatsoever is an “agreement” (*Übereinstimmung*) or correspondence between human cognition (and, more precisely, between a judgment or proposition [A293/B350]) and its specific object [CPR A58/B82]), and
- (ii) what I’ll call the *real definition* of truth, that is, *the criterion of truth*.

Kant also explicitly holds that there is *no single general criterion of truth for all judgments whatsoever* (A58-59/B83), whether it be the Leibniz-Wolff principle of sufficient reason (or principle of contradiction, or principle(s) of identity) or Descartes’s principle to the effect that, necessarily, a judgment or proposition is true if and only if it is clearly and distinctly intuited.

But this doesn't mean that Kant is in any way a *skeptic* about truth-criteria; on the contrary, it means instead that he's a *constrained pluralist* about truth-criteria.³ More precisely, since, as we've seen, for Kant there are *three* basic types of truth and *four* sub-types of truth under one of the basic types, then it follows that for Kant there *seven distinct criteria of truth* altogether.

That was the first cross-cutting distinction.

But **second**, there's also another cross-cutting distinction between

(i) *formal truth*, that is, *analytic or logical truth in pure general logic*, and

(ii) *material truth*, that is, *synthetic truth*, whether in mathematics, real metaphysics, the theoretical-nomological foundations of natural science, the empirical part of natural science, the practical-nomological foundations of ethics and politics, or empirical anthropology.

So again, and to summarize: for Kant, all truth whatsoever is agreement or correspondence between judgments or propositions and their specific objects; there is no single general criterion of truth for all judgments or propositions whatsoever; but there are three criteria of truth for the three basic types of truth (namely, analytic or logical a priori truth, synthetic a priori truth, and synthetic a posteriori truth); and there are four further criteria for the basic sub-types of truth under or within synthetic a priori truth (namely, mathematical truth, real metaphysical truth, *scientific* nomological truths about natural *physical* laws, and *anthropological* or *humane* nomological truths about the natural practical laws of ethics or politics).

Now back to the classical Rationalist Leibniz-Wolff doctrine that Kant is explicitly arguing *against*, namely the doctrine that there's a single, univocal, universal criterion or principle of truth and existence.

Actually, in the *Monadology*, sections 31-35, Leibniz says that

[o]ur reasonings are based on *two great principles, that of contradiction*, in virtue of which we judge that which involves a contradiction to be false, and that which is opposed or contradictory to the false to be true.... And *that of sufficient reason*, by virtue of which we consider that we can find no true or existent fact, no true assertion, without there being a sufficient reason why it is thus and not otherwise.... There are also axioms and postulates, in brief, *primitive principles*, which cannot be proved and which need no proof. And these are *identical propositions*, whose opposite contains an explicit contradiction.⁴

Leibniz's formulations here are *prima facie* puzzling, since at one and the same time they apparently *distinguish between* the principle of contradiction, the PSR, and the principle(s) of identity, and yet also *collapse them* into a single, univocal, universal criterion of truth and existence, variously expressed as (what we would now call)

(i) *the principle of non-contradiction*, namely, Nec (P) (P & ~P),

(ii) *the PSR*, namely, "We can find no true or existent fact, no true assertion, without there being a sufficient reason why it is thus and not otherwise", and

³ See Hanna (2006, ch. 5; 2000b).

⁴ Leibniz (1989:217).

- (iii) one or more versions of *the Leibnizian principle(s) of identity* including,
 - (iiia) a *principle of necessary equivalence or equipollence for judgments or propositions*, namely, $\text{Nec } (P \leftrightarrow P)$,
 - (iiib) *principle of identity for individual substances*, namely, $\text{Nec } (a = a)$,
 - (iiic) *the Identity of Indiscernibles*, namely, $\text{Nec } (x)(y) [(Fx \leftrightarrow Fy) \rightarrow x=y]$, and
 - (iiid) *the Indiscernibility of Identicals*, namely $\text{Nec } (x)(y) [x=y \rightarrow (Fx \leftrightarrow Fy)]$.

Now Wolff was an *ultra-Leibnizian*, who philosophically flattened out many of the theoretical bumps and knots—that is, complexities and *prima facie* inconsistencies—in Leibniz’s doctrines.⁵ So in a spirit of retrospective rational interpretive charity, I think we can say that Leibniz and Wolff *both* held that all of these various formulations, as we would nowadays put it, are *necessarily equivalent in the modally strong sense of necessary intensional equivalence, as well as necessary extensional equivalence*, even if they’re also *non-synonymous*. Then correspondingly, I think that we can retrospectively rationally charitably interpret Kant’s objection to classical Rationalism in general and to the Leibniz-Wolff doctrine in particular as a thoroughgoing rejection of the claim that there is a single, univocal, universal criterion or principle of truth and existence *in the sense that this criterion or principle expresses the necessary intensional and extensional equivalence of the PSR, the principle of contradiction, and the principles of identity*.

In any case, and now time-jumping forward into the 20th and 21st centuries, Kant’s philosophically seminal distinction between analytic and synthetic judgments or propositions is the historical origin of, and therefore significantly related to, but—crucially—*not* strictly equivalent, either in intension or extension, with the nowadays more familiar analytic-synthetic distinction, deriving from the Logical Empiricist tradition, according to which

- (i) analyticity is truth by virtue of linguistic meaning alone, exclusive of empirical facts,
- (ii) syntheticity is truth by virtue of empirical facts, and
- (iii) the necessary statement vs. contingent statement distinction is formally and materially equivalent to the analytic-synthetic distinction.

By 1950 this more familiar distinction was accepted as gospel truth by virtually all Analytic philosophers: but in the two decades after the publication of W.V.O. Quine’s iconoclastic “Two Dogmas of Empiricism” in 1951,⁶ it was gradually replaced by the new-and-improved post-Quinean gospel truth that *there is no such thing as* an acceptable analytic-synthetic distinction. This plain historico-philosophical fact is closely related to the highly regrettable further fact that Kant’s A-S distinction is nowadays almost universally interpreted

⁵ See Anderson (2015, chs. 3-4). For the record, however, apart from the interesting material on Leibniz and Wolff, otherwise I pretty much *completely* disagree with Anderson’s interpretation of Kant’s A-S distinction; see Hanna (2017).

⁶ Quine (1961).

(i) in terms of the more familiar and now largely discredited Logical Empiricist version of the A-S distinction, and also

(ii) as reducible to an *epistemic distinction* between uninformatively or trivially true a priori judgments and informative judgments.

Alas. Ironically Frege, the founding grandfather of analytic philosophy, was much nearer the mark in the *Foundations of Arithmetic* when he correctly construed Kant's theory of analyticity *semantically*, as a theory about necessary internal relations between concepts; although at the same time he not quite so correctly said that Kantian analyticity boils down to “simply taking out of the box again what we have just put into it.”⁷ Backing away now from Frege, the crucial fact is that Kant's A-S distinction concerns two irreducibly different kinds of semantic content within objectively valid propositions,⁸ and this distinction is neither primarily epistemic in character (although it does have some important epistemic implications⁹) nor does it specifically concern the logical form of judgments (Prol, 4:266).

Frege famously regarded Kant's notion of analyticity as *trivial*. But in fact and on the contrary, Kant's notion of analyticity is *substantive*, by virtue of five important ideas:

(i) Kant's *pro-Leibnizian* idea that all concepts have intensional microstructures, or what he calls “logical essences” or “conceptual essences” (Log, 9:61),

(ii) his *anti-Leibnizian* idea that logically possible worlds are nothing but maximal logically consistent sets of concepts, not things-in-themselves (A571-573/B599-601),

(iii) his *referentialist* idea that all grammatically well-formed, sortally correct, and logically consistent concepts have non-empty cross-possible-worlds extensions (aka “comprehensions”) (A239/B298-299) (Log, 9:95-96),

(iv) his *semantic restrictionist* idea that all and only objectively valid propositions have truth-values, and finally, and most importantly,

(v) his *logical analyticity* idea that the notion of analyticity covers not only the so-called “containment” of predicate-concepts in subject-concepts in categorical propositions, and not only the intensional identity of subject-concepts and predicate-concepts, *but also* all the logical truths of truth-functional logic and monadic predicate logic.

Then for Kant a judgment is analytic if and only if its propositional content is necessarily true either by virtue of necessary internal relations between its objectively valid conceptual microstructures and/or its conceptual comprehensions, or by virtue of its truth-functional logical connectives, or by virtue of its monadic predicate logical connectives.¹⁰

⁷ Frege (1953:101).

⁸ Hanna (2001, chs. 3-4).

⁹ See, e.g., Hanna (2006:362-379; 1998).

¹⁰ See Hanna (2001, ch. 3).

So let this be repeated with strong emphasis: Kant does *not* define analyticity in terms of either the containment or the identity of concepts, which are at best sufficient conditions for analyticity and *not* also necessary conditions for analyticity. Sharply on the contrary, Kant explicitly states a universal, necessary, and sufficient semantic criterion for the truth of analytic judgments, namely that a judgment is analytically true if and only if its denial entails a contradiction, in a broad sense of “entailment” that includes intensional entailment and not merely classical deductive entailment (A151/B190-191).

This criterion also directly connects the notion of an *analytic truth* with the notion of a *logical truth* in a correspondingly broad sense that fully includes the tautologies and valid sentences of truth-functional logic and monadic predicate logic, but is neither restricted to nor reducible to (with the addition of “logical definitions,”¹¹ whatever *they* actually turn out to be on Frege’s account, which is not at all clear¹²) the truth-functional tautologies and valid sentences of classical polyadic predicate logic.

OK: so much for analyticity, for the time being; now what about syntheticity?

Since for Kant the A-S distinction is intended to be *exhaustive* in the sense that every judgment whatsoever is either analytic or synthetic but not both, his two-part doctrine of analyticity in turn provides him with a two-part negative doctrine of syntheticity:

A judgment is synthetic if and only if its truth is *not* strictly determined by relations between its conceptual microstructures or conceptual comprehensions alone, or by truth-functional logic or monadic predicate logic alone (which for Kant is expressively captured by the table of judgments and the table of pure concepts of the understanding); and a judgment is synthetically true if and only if it is true and its denial does *not* entail a contradiction.

But this negative characterization of course does not tell us what the truth of synthetic judgments positively consists in. In order to do this, Kant directly connects the semantics of syntheticity with the semantics of *intuitions* or *Anschauungen* just as he directly connects the semantics of analyticity with the semantics of *concepts* or *Begriffe* (including both empirical concepts and the pure concepts of the understanding). Then positively put, a judgment is synthetic if and only if its meaning and its truth are strictly determined by its constituent intuitions, whether empirical intuitions or pure intuitions (A8, A154-155/B193-194, A721/B749) (Ü, 8: 245) (Briefe, 11:38 [1790]). This is *not* to say either that synthetic judgments do *not* contain any concepts (in fact they always *do* contain empirical or pure concepts), or even that the conceptual components of a synthetic judgment are *irrelevant* to its meaning or truth (in fact empirical or pure concepts *always are* semantically relevant), but only to say that in a synthetic judgment it is the intuitional components that *strictly determine* its meaning and truth, not its empirical-conceptual or pure-conceptual components. In short, a synthetic judgment is *an intuition-based objectively valid or empirically meaningful proposition*.

Now what about the synthetic a priori?

Every reader of the *Critique of Pure Reason* knows that in the Introduction, Kant glosses his Critical philosophical project in that book as a complete and systematic answer to the question,

¹¹ See Frege (1953).

¹² See Benacerraf (1981).

“**how are synthetic *a priori* judgments possible?**” (B19, boldfacing in the original). And by the end of the “Transcendental Aesthetic” section, he provides a partial answer to his own question:

Here we now have one of the required pieces for the solution of the general problem of transcendental philosophy - **how are synthetic *a priori* propositions possible?** - namely pure *a priori* intuitions, space and time, in which, if we want to go beyond the given concept in an *a priori* judgment, we encounter that which is to be discovered *a priori* and synthetically connected with it, not in the concept but in the intuition that corresponds to it; but on this ground such a judgment never extends beyond the objects of the senses and can hold only for objects of possible experience. (B73, boldfacing in the original)

Relatedly, in a letter to Karl Reinhold on 12 May 1789, Kant says:

All synthetic judgments of theoretical cognition are possible only by relating a given concept to an intuition... If the synthetic judgment is experiential, then the underlying intuition must be empirical; if the judgment is synthetic a priori, the intuition must be pure. (Briefe, 11:38 [1790])

And in “On a Discovery” in 1790, he says that

the *Critique* made knowable *for the first time* the distinction between judgments which rest entirely on the principle of identity or contradiction, and those which require another principle, by naming them analytic as opposed to synthetic judgments. For by the term synthesis it is clearly indicated that something outside of the given concept must be added as a substrate, which makes it possible to go beyond the concept with my predicates; so that the investigation is directed to the possibility of a synthesis of representations with regard to cognition in general, which soon had to lead to a recognition of *intuition* as its indispensable condition, but of *pure intuition* for a [synthetic] priori cognition. (Ü, 8:245)

In this way, for Kant the meaning and truth of synthetic a priori judgments or propositions are grounded on *pure intuitions*, i.e., our a priori formal representations of space and time. Now since according to Kant our a priori formal representations of space and time are both necessary conditions of the possibility of human experience and *also* necessary conditions of the objective validity or empirical meaningfulness of judgments, which in turn confers truth-valuedness upon propositions, it then follows that a synthetic a priori judgment is a necessary objectively valid proposition that is true in all and only the humanly experienceable possible worlds and truth-valueless otherwise. Less abstractly put, a synthetic a priori judgment is *a necessary truth with a human face*.

In the first *Critique*, Kant explicitly holds that synthetic a priori judgments really exist in various sciences, including, mathematics, physics, and legitimate, i.e., transcendental idealist, metaphysics (A4–5, B4–5, B14–18). But Kant’s theory of the synthetic a priori also has a strongly critical, negative component. For although all of the basic judgments/propositions of traditional, and especially classical Rationalist, metaphysics are, *aspirationally*, synthetic a priori judgments (CPR B18), they are ultimately so *only* aspirationally. Hence his famous “critique of pure reason,” as applied to traditional and especially classical Rationalist metaphysics in the “Transcendental Dialectic” part of the first *Critique*, is nothing more and nothing less than an investigation of how objectively valid or empirically meaningful, true synthetic a priori judgments in classical and especially Rationalist metaphysics are *humanly impossible*.

Moreover—and this fact is generally unrecognized, even by Kant-scholars and other Kantians—in Kant’s post-Critical period, the synthetic a priori also extends beyond theoretical philosophy into pure practical philosophy, that is, into *the metaphysics of morals*. For Kant generally holds that “an axiom is an immediate intuitional judgment a priori” (Refl, 16:673, 3135), and also explicitly says in the *Metaphysics of Morals* that every immediately certain synthetic a priori proposition about “right” (*Recht*) is a pure practical “axiom of right” (MS, 6:250) or “axiom of outer freedom” (MS, 6:267-268). Correspondingly, there must also be pure practical *axioms of virtue* or *axioms of inner freedom or autonomy*, that is, moral principles flowing directly from the Categorical Imperative, although Kant never says this explicitly. In the metaphysics of morals, pure practical axioms are sensibly grounded in egoistic human empirical desires (for outer freedom) and in the non-egoistic a priori feeling of respect (for inner freedom). So the metaphysics of morals has a “human face” too.

3. The Philosophical Main Event: Four Basic Modes of Sufficiency and Reasonhood

Now for the philosophical main event.

As I pointed out in section 2, throughout his pre-Critical, proto-Critical, Critical, and Post-Critical periods, Kant repeatedly engaged critically with the Leibniz-Wolff doctrine of the PSR, in his “New Elucidation of the First Principles of Metaphysical Cognition” in 1755, “The Only Possible Argument in Support of a Demonstration of the Existence of God” in 1763, the “Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality”, aka the “Prize Essay”, in 1764, “On the Form and Principles of the Sensible and Intelligible Worlds”, aka the “Inaugural Dissertation”, in 1770, in the two editions of the *Critique of Pure Reason* in 1781 and 1787, especially in the “Antinomies of Pure Reason” and “Ideal of Pure Reason” sections of the Dialectic of Pure Reason, and finally in “On a Discovery” in 1790.

But at the same time, Kant was developing his own positive, original theory of the PSR, according to which, as I claimed in section 1, he is both a *fourfold-root pluralist* with respect to the PSR, in addition to his being *both* a sufficiency-pluralist and *also* a reasons-pluralist with respect to the PSR; and this in turn means that, in Kant’s hands, the PSR is an emphatically robust and thoroughly non-reductive principle; and therefore my interpretation of Kant’s doctrine of the PSR is in sharp and in fact diametric contrast to Béatrice Longuenesse’s well-known interpretation, according to which Kant is a deflationist or reductivist (or in her terminology, a “deconstructionist”) about the PSR.¹³

Correspondingly, I also claimed in section 1 that for Kant, the PSR comprehends four basic modes or “roots” of sufficiency and reasonhood:

- (i) *analytic or logical sufficiency and reasonhood,*
- (ii) *mathematical sufficiency and reasonhood,*
- (iii) *causal-physical-mechanical sufficiency and reasonhood,* and finally
- (iv) *causal-deontological-spontaneous sufficiency and reasonhood.*

¹³ See note 1 above.

I'll now spell these out more carefully, by way of defending my overall interpretive thesis.

3.1 Analytic or Logical Sufficiency and Reasonhood

What, according to Kant, is analytic or logical sufficiency and reasonhood? In order to be able to answer that question, we must first know what, according to Kant, *logic* and *truth* are.

As regards logic, in the *Critique of Pure Reason* he says that

a **general** but **pure** logic ... has to do with strictly a priori principles, and is a **canon** of the understanding and reason, but only in regard to what is formal in their use, be the content what it may. (A53/B77)

And in the *Jäsche Logic*, he says that

[the] science of the necessary laws of the understanding and of reason in general, or what is the same, of the mere form of thought as such, we call logic (Log, 9:13),

and also that logic is the “science that deals with all thought in general, without regard to objects as the matter of thought” (Log, 9:13). Now echoing Pontius Pilate, but more specifically according to Kant, what is truth? In the first *Critique*, Kant says that

[t]he nominal definition of truth, namely that it is the agreement (*Übereinstimmung*) of cognition with its object, is here granted and presupposed; but one demands to know what is the general and certain criterion of the truth of any cognition [that is, the *real definition* of truth—RH] (A58/B82),

and also that

truth and illusion are not in the object, insofar as it is intuited, but in the judgment about it, insofar as it is thought.... [T]ruth, as much as error, and thus also illusion as leading to the latter, are to be found only in judgments, i.e., only in the relation of the object to our understanding. (A293/B350)

And in the *Jäsche Logic*, he says that

truth must consist in the agreement of a cognition with just that determinate object to which it is related. (Log, 9:51)

So according to Kant, logic is the science of the necessary formal rules, aka laws, of thought and reason, and truth is the correspondence between a judgment or proposition and the specific object to which it is related.

Now we're in a position to be able to say what, according to Kant, analytic or logical sufficiency is. In the *Critique of Pure Reason*, he says that what he calls “the principle of contradiction” specifies both a strictly universal, *necessary*, and *negative* condition *on all possible judgments or propositions* and also a strictly universal, *necessary and sufficient* criterion of truth for analytic or logical truths, but *not* a sufficient criterion of truth for *non-analytic* truths:

Whatever the content of our cognition may be, and however it may be related to the object, the general though to be sure only negative condition of all our judgments whatsoever is that they do not contradict themselves; otherwise these judgments in themselves (even without regard to the object) are nothing. (A150/B189)

Now the proposition that no predicate pertains to a thing that contradicts it is called the principle of contradiction, and is a general though merely negative uncriterion of all truth, but on that account it also belongs merely to logic, since it holds of cognitions merely as cognitions in general, without regard to their content, and says that contradiction entirely annihilates and cancels them. (A151/B190)

[I]f the judgment is analytic, whether it be negative or affirmative, its truth must always be able to be cognized sufficiently in accordance with the principle of contradiction.... Hence we must allow the principle of contradiction to count as the universal and completely sufficient principle of all analytic cognition; but its authority and usefulness does not extend beyond this, as sufficient criterion of truth. For that no cognition can be opposed to it without annihilating itself certainly makes this principle into a *conditio sine qua non*, but not into a determining ground of the truth of our cognition. (A151/B191, boldfacing in the original)

And in the *Jäsche Logic*, Kant also says that the principle of contradiction and the principle of sufficient reason form a complementary pair of principles governing logical truth:

The formal criteria of truth in logic are:

1. *the principle of contradiction*,
2. *the principle of sufficient reason*.

Through the former, the *logical possibility* of a cognition is determined, through the latter its *logical actuality*. To the logical actuality of a cognition it pertains, namely:

First: that it be logically possible, i.e., *not contradict itself*. This characteristic of *internal* logical truth is only *negative*, however; for a cognition that contradicts itself is of course false, but if it does not contradict itself it is not always true.

Second: that it be logically grounded, i.e., that it (a) have grounds and (b) not have false consequences. This second criterion of external logical truth or of *accessibility to reason* (*Rationabilität*), which concerns the logical connection of a cognition with grounds and consequences, is positive. And here the following rules are valid:

1. From the *truth of the consequence*, we may infer the truth of the cognition as ground, but only *negatively*: if one false consequence flows from a cognition, then the cognition itself is false. For if the ground were true, then the consequence would also have to be true, because the consequence is determined by the ground.

But one cannot infer conversely that if no false consequence flows from a cognition, then it is true; for one can infer true consequences from a false ground.

2. *If all the consequences of a cognition are true, then the cognition is true too*. For if there were something false in the cognition, then there would have to be a false consequence too. (Log, 9: 51-52)

In other words, Kantian analytic or logical sufficiency and reasonhood comprehends all logical truths or laws, logically valid arguments/proofs, and logical consequence or entailment.

3.2 Mathematical Sufficiency and Reasonhood

What, according to Kant, is mathematical sufficiency and reasonhood? In the first *Critique*, he says:

Mathematical judgments are all synthetic.... [A] synthetic proposition can of course be comprehended in accordance with the principle of contradiction, but only insofar as another synthetic proposition is presupposed from which it can be deduced, never in itself.... [P]roperly mathematical propositions are always a priori judgments and are never empirical, because they

carry necessity with them, which cannot be derived from experience.... The arithmetical proposition is ... always synthetic.... Just as little is any proposition of pure geometry analytic. (A10/B14-17)

Geometry is a science that determines the properties of space synthetically and yet *a priori*. What then must the representation of space be for such a cognition of it to be possible? It must originally be intuition; for from a mere concept no propositions can be drawn that go beyond the concept, which, however, happens in geometry ... But this intuition must be encountered in us *a priori*, i.e., prior to all perception of an object, thus it must be pure, not empirical intuition. (A25/B40-41)

The pure image of all magnitudes (*quantorum*) for outer sense is space; for all objects of the senses in general, it is time. The pure **schema of magnitude (quantitatis)**, however, as a concept of the understanding, is **number**, which is a representation that summarizes the successive addition of one (homogenous) unit to another. Thus number is nothing other than the unity of the synthesis of the manifold of a homogenous intuition in general, because I generate time itself in the apprehension of the intuition. (A142-143/B182, boldfacing in the original)

[T]here are pure principles a priori that I may nevertheless not properly ascribe to the pure understanding, since they are not derived from pure concepts but from pure intuitions (although by means of the understanding) ... [and] [m]athematics has principles of this sort. (A160/B199)

All appearances are, as regards their intuition, **extensive magnitudes**. (A161, boldfacing in the original)

Now the consciousness of the homogeneous manifold in intuition in general, insofar as through it the representation of an object first becomes possible, is the concept of a magnitude (*Quanti*).... I call an extensive magnitude that in which the representations of the parts makes possible the representation of the whole (and necessarily precedes the latter).... Since the mere intuition in all appearances is either space or time, every appearance as intuition is an extensive magnitude, as it can only be cognized through successive synthesis (from part to part) in apprehension. All appearances are accordingly intuited as aggregates (multitudes of antecedently given parts), which is not the case with every kind of magnitude, but rather only with those that are represented and apprehended by us **extensive**. (A162/B203-A163/B204, boldfacing in the original)

In other words, mathematical sufficiency is just analytic or logical sufficiency, *together with* mediation and schematization by the pure representations of space and time, specifically governed by the principles of pure understanding that Kant calls the *Axioms of Intuition*, comprehending mathematical truths, axioms, or laws, mathematically valid arguments/proofs, and mathematical consequence or entailment.

3.3 Causal-Physical-Mechanical Sufficiency and Reasonhood

What, according to Kant, is causal-physical-mechanical sufficiency and reasonhood? In the first *Critique*, he says:

The schema of the cause and of the causality of a thing in general is the real upon which, whenever it is posited, something else always follows. It therefore consists in the succession of the manifold insofar as it subject to a rule. (A144/B183)

That something happens, therefore, is a perception that belongs to a possible experience, which becomes actual if I regard the position of the appearance as determined in time, thus if I regard it

as an object that can always be found in the connection of perceptions in accordance with a rule. This rule for determining something with respect to its temporal sequence, however, is that in what precedes, the condition is to be encountered under which the occurrence always (i. e., necessarily) follows. Thus the principle of sufficient reason is the ground of possible experience, namely the objective cognition of appearances with regard to their relation in the successive series of time. (A200–201/B246)

[The objects] of outer sense, thus the sum total of these, [is] corporeal nature.... [and] the mere concept of matter [is the concept of] ... impenetrable lifeless extension. (A846-847/B874-876)

In the *Metaphysics Foundations of Natural Science*, somewhat longwindedly, he says:

If the word nature is taken simply in its *formal* meaning, where it means the first inner principle of all that belongs to the existence of a thing, then there can be as many different natural sciences as there are specifically different things, each of which must contain its own peculiar inner principle of the determinations belonging to its existence. But nature is also taken otherwise in its *material* meaning, not as a constitution, but as the sum total of all things, insofar as they can be *objects of our senses*, and thus also of experience. Nature, in this meaning, is therefore understood as the whole of all appearances, that is, the sensible world, excluding all nonsensible objects. Now nature, taken in this meaning of the word, has two principal parts, in accordance with the principal division of our senses, where the one contains the objects of the *outer* senses, the other the object of *inner* sense. In this meaning, therefore, a twofold doctrine of nature is possible, the *doctrine of body* and the *doctrine of the soul*, where the first considers *extended* nature, the second *thinking* nature. (MAN, 4:467)

Properly so-called natural science presupposes, in the first place, metaphysics of nature. For laws, that is, principles of the necessity of that which belongs to the *existence* of a thing, are concerned with a concept that cannot be constructed, since existence cannot be presented a priori in any intuition. Thus proper natural science presupposes metaphysics of nature. Now this latter must always contain solely principles that are not empirical (for precisely this reason it bears the name of a metaphysics), but it can still either: *first*, treat the laws that make possible the concept of a nature in general, even without relation to any determinate object of experience, and thus undetermined with respect to the nature of this or that thing in the sensible world, in which case it is the *transcendental* part of the metaphysics of nature; or *second*, concern itself with a particular nature of this or that kind of thing, for which an empirical concept is given, but still in such a manner that, outside of what lies in this concept, no other empirical principle is used for its cognition (for example, it takes the empirical concept of matter or of a thinking being as its basis, and it seeks that sphere of cognition of which reason is capable a priori concerning these objects), and here such a science must still always be called a metaphysics of nature, namely, of corporeal or of thinking nature. However, [in this second case] it is then not a general, but a *special* metaphysical natural science (physics or psychology), in which the above transcendental principles are applied to the two species of objects of our senses. I assert, however, that in any special doctrine of nature there can be only as much *proper* science as there is *mathematics* therein. For, according to the preceding, proper science, and above all proper natural science, requires a pure part lying at the basis of the empirical part, and resting on a priori cognition of natural things. Now to cognize something a priori means to cognize it from its mere possibility. But the possibility of determinate natural things cannot be cognized from their mere concepts; for from these the possibility of the thought (that it does not contradict itself) can certainly be cognized, but not the possibility of the object, as a natural thing that can be given outside the thought (as existing). Hence, in order to cognize the possibility of determinate natural things, and thus to cognize them a priori, it is still

required that the *intuition* corresponding to the concept be given a priori, that is, that the concept be constructed. Now rational cognition through construction of concepts is mathematical. Hence, although a pure philosophy of nature in general, that is, that which investigates only what constitutes the concept of a nature in general, may indeed be possible even without mathematics, a pure doctrine of nature concerning *determinate* natural things (doctrine of body or doctrine of soul) is only possible by means of mathematics. And, since in any doctrine of nature there is only as much proper science as there is a priori knowledge therein, a doctrine of nature will contain only as much proper science as there is mathematics capable of application there.

So long, therefore, as there is still for chemical actions of matters on one another no concept to be discovered that can be constructed, that is, no law of the approach or withdrawal of the parts of matter can be specified according to which, perhaps in proportion to their density or the like, their motions and all the consequences thereof can be made intuitive and presented a priori in space (a demand that will only with great difficulty ever be fulfilled), then chemistry can be nothing more than a systematic art or experimental doctrine, but never a proper science, because its principles are merely empirical, and allow of no a priori presentation in intuition. Consequently, they do not in the least make the principles of chemical appearances conceivable with respect to their possibility, for they are not receptive to the application of mathematics. Yet the empirical doctrine of the soul must remain even further from the rank of a properly so-called natural science than chemistry. In the first place, because mathematics is not applicable to the phenomena of inner sense and their laws, the only option one would have would be to take the *law of continuity* in the flux of inner changes into account—which, however, would be an extension of cognition standing to that which mathematics provides for the doctrine of body approximately as the doctrine of the properties of the straight line stands to the whole of geometry. For the pure inner intuition in which the appearances of the soul are supposed to be constructed is *time*, which has only one dimension. [In the second place,] however, the empirical doctrine of the soul can also never approach chemistry even as a systematic art of analysis or experimental doctrine, for in it the manifold of inner observation can be separated only by mere division in thought, and cannot then be held separate and recombined at will (but still less does another thinking subject suffer himself to be experimented upon to suit our purpose), and even observation by itself already changes and displaces the state of the observed object. Therefore, the empirical doctrine of the soul can never become anything more than an historical doctrine of nature, and, as such, a natural doctrine of inner sense which is as systematic as possible, that is, a natural description of the soul, but never a science of the soul, nor even, indeed, an experimental psychological doctrine. This is also the reason for our having used, in accordance with common custom, the general title of natural science for this work, which actually contains the principles of the doctrine of body, for only to it does this title belong in the proper sense, and so no ambiguity is thereby produced. But in order to make possible the application of mathematics to the doctrine of body, which only through this can become natural science, principles for the *construction* of the concepts that belong to the possibility of matter in general must be introduced first. Therefore, a complete analysis of the concept of a matter in general will have to be taken as the basis, and this is a task for pure philosophy – which, for this purpose, makes use of no particular experiences, but only that which it finds in the isolated (although intrinsically empirical) concept itself, in relation to the pure intuitions in space and time, and in accordance with laws that already essentially attach to the concept of nature in general, and is therefore a genuine *metaphysics of corporeal nature*. Hence all natural philosophers who have wished to proceed mathematically in their occupation have always, and must have always, made use of metaphysical principles (albeit unconsciously), even if they themselves solemnly guarded against all claims of metaphysics upon their science. Undoubtedly they have understood by the latter the folly of contriving possibilities at will and playing with concepts, which can perhaps not be presented in intuition at all, and have no other certification of their objective reality than that they merely do not contradict themselves. All true metaphysics is drawn from the essence of the

faculty of thinking itself, and is in no way fictitiously invented on account of not being borrowed from experience. Rather, it contains the pure actions of thought, and thus a priori concepts and principles, which first bring the manifold of *empirical representations* into the law-governed connection through which it can become *empirical cognition*, that is, experience. Thus these mathematical physicists could in no way avoid metaphysical principles, and, among them, also not those that make the concept of their proper object, namely, matter, a priori suitable for application to outer experience, such as the concept of motion, the filling of space, inertia, and so on. But they rightly held that to let merely empirical principles govern these concepts would in no way be appropriate to the apodictic certainty they wished their laws of nature to possess, so they preferred to postulate such [principles], without investigating them with regard to their a priori sources. (MAN, 4:469-472)

And in the *Critique of Practical Reason*, more compactly, he says:

[A]ll necessity of events in time according to natural law can be called the “mechanism of nature,” even though it is not to be supposed that things which are subject to it must really be material machines. Here reference is made only to the necessity of the connection of events in a temporal series as they develop according to natural law, whether the subject in which this development occurs be called *automaton materiale* when the machinery is impelled by matter, or, with Leibniz, *automaton spirituale* when it is impelled by representations. And if the freedom of our will were nothing else than the latter, i.e., psychological and comparative and not at the same time transcendental or absolute, it would in essence be no better than the freedom of a turnspit, which when once wound up also carries its motions from itself. (KpV, 5:97, translation slightly modified)

So in other words, for Kant causal-physical-mechanical sufficiency and reasonhood is just mathematical sufficiency and reasonhood, *together with* the existential presupposition of inert matter, and a simple version of computability for basic physical quantities that’s based on Primitive Recursive Arithmetic, specifically governed by the Analogies of Experience and the Postulates of Empirical Thought, comprehending deterministic, mechanistic natural laws and states of individual things, properties, relations, facts, events, and processes.

III.4 Causal-Deontological-Spontaneous Sufficiency and Reasonhood

Now something philosophically amazing is going to happen.

Many or even most readers of Kant—including many or even most contemporary Kant-scholars, following Longuenesse’s lead—have assumed that Kant’s use of the PSR in relation to natural causation is *deflated* and *restricted* to causal-physical-mechanical sufficiency and reasonhood.¹⁴ But this is manifestly *not* the case, because Kant also explicitly extends the PSR to *organismic, non-mechanical, teleological* natural causation, *not only* in an explanatory or regulative sense, *but also* in a real-metaphysical or constitutive sense that applies directly to the manifestly real world of appearances.

As he famously or notoriously points out in the *Critique of the Power of Judgment*,

it is quite certain that we can never adequately come to know the organized beings and their internal possibility in accordance with merely mechanical principles of nature, let alone explain them; and indeed this is so certain that we can boldly say that it would be absurd for humans even to make such an attempt or to hope that there may yet arise a Newton who could make comprehensible the

¹⁴ See, e.g., Longuenesse (2001).

generation of a blade of grass according to natural laws that no intention has ordered; rather we must absolutely deny this insight to human beings. (KU, 5:400)

This Kantian doctrine directly implies, as Patricia Kauark-Leite very aptly puts it, that

the PSR, as a transcendental principle in its theoretical use, as applied to the realm of possible experience, cannot be *explanatorily* assimilated to the principle of causality in the “Second Analogy of Experience,” even for Newtonian physics. The ideas of reason, although they are not directly and constitutively applied to any object of experience, have a regulative use that is entirely legitimate and absolutely indispensable for completing the task of the understanding. While the understanding aims at unifying the multiple of experience by means of its concepts, reason aims at the unification of all empirical laws by means of its ideas, seeking the maximum possible expansion of our representation of the world of actual and possible experience. In this way, the restriction of the PSR to the deterministic, mechanical causal principle for the determination of an objective chronological order of the objects of perception, as established by the “Second Analogy of Experience,” is valid at the level of the understanding. However, reason demands more. It demands that the sufficiency of the scientific explanation must be sought in the idea of the complete unity of the concepts of the understanding, even if only in a hypothetical or counterfactual way. But these uses of the PSR entail *necessary* hypotheticals or counterfactuals that make manifest the explanatory inadequacy of deterministic, mechanical causal explanation.¹⁵

Yet Kant’s use of the PSR beyond deterministic, mechanical causation goes even beyond a merely *explanatory, regulative* application to non-deterministic, non-mechanical, organismic, teleological phenomena, to a *real-metaphysical, constitutive* application to those very phenomena. In the *Prolegomena to Any Future Metaphysics* and also in the third *Critique*, he explicitly says that *all* mindedness, including of course *human* mindedness, is *a form of organismic life*:

[L]ife is the subjective condition of all our possible experience. (Prol, 4: 335)

Life without the feeling of the corporeal organ is merely consciousness of one’s existence, but not a feeling of well- or ill-being, i.e., the promotion or inhibition of the powers of life; because the mind for itself is entirely life (the principle of life itself), and hindrances and promotions must be sought outside it, though in the human being himself, hence in combination with his body. (KU, 5:278)

Furthermore, in the first *Critique*, Kant says:

By freedom in the cosmological sense ... I understand the faculty of beginning a state **from itself** (*von selbst*), the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature. Freedom in this signification is a pure transcendental idea. (A533/B561, boldfacing in the original)

In the *Critique of Practical Reason*, he says (the second text here is one we’ve already seen):

It is this **transcendental** idea of freedom on which the practical concept of freedom is grounded.... **Freedom in the practical sense** is the independence of the power of choice (*Willkür*) from **necessitation** by impulses of sensibility. For a power of choice is **sensible** insofar as it is

¹⁵ Kauark-Leite (2018:1267).

pathologically affected (through moving-causes of sensibility); it is called an animal power of choice (*arbitrium brutum*) if it can be **pathologically necessitated**. The human power of choice is indeed an *arbitrium sensitivum*, yet not *brutum*, but *liberum*, because sensibility does not render its action necessary, but in the human being there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses. (A534/B562, boldfacing in the original)

[A]ll necessity of events in time according to natural law can be called the “mechanism of nature,” even though it is not to be supposed that things which are subject to it must really be material machines. Here reference is made only to the necessity of the connection of events in a temporal series as they develop according to natural law, whether the subject in which this development occurs be called *automaton materiale* when the machinery is impelled by matter, or, with Leibniz, *automaton spirituale* when it is impelled by representations. And if the freedom of our will were nothing else than the latter, i.e., psychological and comparative and not at the same time transcendental or absolute, it would in essence be no better than the freedom of a turnspit, which when once wound up also carries its motions from itself. (KpV, 5:97, translation slightly modified)

And perhaps most amazingly of all, in the first *Critique*, the third *Critique*, and the *Opus postumum*, he says:

Practical freedom can be proved through experience. For it is not merely that which stimulates the senses, i.e., immediate affects them, that determines human choice, but we always have a capacity to over-come impressions on our sensory faculty of desire by representations of that which is useful or injurious even in a more remote way; but these considerations about that which in regard to our whole condition is desirable, i.e., good and useful, depend on reason. Hence this also yields laws that are imperatives, i.e., objective laws of freedom, and that say what ought to happen, even though it never does happen We thus cognize practical freedom through experience, as one of the natural causes, namely a causality of reason in the determination of the will. (A802–803/B830–831, underlining added)

Now although there is an incalculable gulf fixed between the domain of the concept of nature, as the sensible, and the domain of the concept of freedom, as the supersensible ... yet the latter **should** have an influence on the former, namely the concept of freedom should make the end that is imposed by its laws real in the in the sensible world; and nature must consequently also be able to be conceived in such a way that the lawfulness of its form is at least in agreement with the possibility of the ends that are to be realized in it in accordance with the laws of freedom. (KU, 5: 176, boldfacing in the original, underlining added)

THE HUMAN BEING [IS] A BEING IN THE WORLD, SELF-LIMITED THROUGH NATURE AND DUTY. (Op. Post., 21:34, underlining added)

In other words, causal-deontological-spontaneous sufficiency and reasonhood is just mathematical sufficiency and reasonhood, together with the *non-mechanistic* existential presupposition of organismic living matter, together with the specifically human embodiment of the innately specified *spontaneous* and thereby *nondeterministic*, *non-mechanistic* innate cognitive capacities or powers, aka faculties, of

- consciousness
- sensibility: intuition and imagination
- understanding: concepts and thinking

- judgment
- affect (feelings, desires, emotions)
- will: either *Willkuer*, “the power of choice” (first-order effective affect) or *Wille* “pure will” (higher-order effective affect, inherently sensitive to hypothetical or categorical reasons)
- self-consciousness or apperception, and
- rationality (logical inference or practical decision-making, inherently governed by principles),

all of them governed by hypothetical imperatives and the Categorical Imperative, comprehending nondeterministic, non-mechanistic, non-instrumental moral and practical laws, free choices and free acts, and the autonomous willing of human persons and the practical agency of rational human animals more generally, *inside* the manifestly real world of appearances. As I mentioned in section 1, this means that for Kant, a human person, via her free will and practical agency, involving both desires (aka “internal reasons”) and ends (aka “external reasons”), can herself *be* a sufficient reason—an ultimate author, creator, or source—of her own choices and acts. This not merely amazing, but even *amazingly* amazing, Kantian doctrine is what I’ve also called *Kant’s Embodied Agency Theory of Freedom*.¹⁶

4. The Categorical Normativity of All Four Basic Modes of Kant’s Principle of Sufficient Reason, and a Kantian Solution to the Circularity or Reflexivity Problem

As I mentioned in section 1, by “categorical normativity,” I mean *irreducible, unconditional, non-instrumental normativity that is ultimately grounded on (the four or five intensionally and extensionally necessarily equivalent formulations of) the Categorical Imperative*. Perhaps surprisingly, the categorical normativity of all four basic modes of Kant’s principle of sufficient reason is guaranteed by his doctrine of *the categorical normativity of pure general logic*. How does this doctrine unfold?

In his discussion of the nature of logic in the *Critique of Pure Reason*, Kant focuses on what he calls “pure general analytic logic,” and takes it to be not only the paradigm of a completed, closed rational science, essentially in place since the Aristotelian, later Greek (especially including Stoic), and Scholastic logical traditions (Bviii-ix), but also a universal presupposition of, propaedeutic for, and canon (as opposed to an organon) of every other science (including mathematics, natural science, and metaphysics), and indeed for all cognition, thought (*Denken*), and scientific knowing (*Wissen*) whatsoever.

This, in turn, is precisely because pure general analytic logic provides a set of necessary rules or laws (e.g., the principle of contradiction) that count as minimally necessary, unconditionally or categorically (as opposed to merely hypothetically or instrumentally) normative principles for all cognition, thought, science (*Wissenschaft*), scientific knowing, and truth (*Wahrheit*) (A50-64/B74-88). Logic in this Kantian sense is:

- (i) *pure*, because it is a priori or strictly underdetermined by sensory, contingent facts (and indeed without any associated sensory content), hence also necessary,
- (ii) *general*, or “formal,” because it is absolutely universal and strictly underdetermined by all objectively valid representational contents, and thereby abstracts away from all specific

¹⁶ See Hanna (2006, ch. 8).

or particular differences between represented objects (as opposed to *particular* [*besondere*] logic, which is specifically sensitive to different kinds of objects), and “has to do with nothing but the mere form of thinking” (A54/B78),

(iii) *analytic*, because it deals with formal rules for contingent or necessary truth and valid consequence in reasoning, as opposed to “dialectic,” which systematically traces patterns of contingent or necessary falsity, fallacy, and illusion in reasoning, and

(iv) *logic*, because it is the science of the rules of the innate capacities or powers of understanding and reason (*Vernunft*) in general, including rules for the generation and use of concepts, judgments/propositions, and inferences.

Given these basic features of pure general analytic logic for Kant, four other very important derived features of it are

(i) that it is *anti-psychologistic*, in the sense that it is irreducible to and autonomous from empirical psychology,

(ii) that it is a *pure morality of thinking and reasoning*, in that it is related to applied logic (that is, systematically tracing patterns of cognition, thinking, and reasoning under concrete, real-world conditions) just as pure morality based on the Categorical Imperative is related to the applied doctrine of virtue (that is, systematically tracing morally significant patterns of emotion, willing, and acting under concrete, real-world conditions) (A54/B78),

(iii) that the *analyticity* of logical truths (that is, necessary truth grounded on intrinsic containment-relations or identity-relations between concepts, and criterially determined by the principle of contradiction) contrasts categorically with the syntheticity of non-logical truths (i.e., contingent or necessary truth grounded on non-intrinsic relations between concepts together with pure or empirical intuition) (A150-158/B189-197), and

(iv) that the basic logical forms of judgments/propositions are also the logico-metaphysically foundational *Pure Concepts of the Understanding or Categories*, insofar as they are inherently related to the intuition of actual or possible objects, whether these are objects in general, or objects specifically given in alien or human sensible intuition, via what Kant calls *transcendental logic* (A55-57/B79-82 and A66-83/B91-116).

All of these basic and derived features of pure general analytic logic (i.e., purity/apriority, necessity, generality/formality, analyticity, anti-psychologism, categorical normativity, and logico-metaphysical presuppositional foundationalism via transcendental logic) are also present in the definition of the concept of logic presented in the *Jäsche Logic*:

Logic is a science of reason, not [only] as to mere form but also as to matter (Materie); a science a priori of the necessary laws of thought, not in regard to particular objects, however, but to all objects in general; — hence a science of the correct use of the understanding and of reason in general, not subjectively, however, i.e., not according to empirical (psychological) principles for how the understanding does think, but objectively, i.e., according to principles a priori for how it ought to think. (Log, 9:16, italics in the original)

At the same time, there are three further derived features of logic as presented in the *Jäsche Logic* that go substantially beyond what is presented in the first *Critique*. The first feature is the deep analogy between pure general analytic logic and *universal grammar* as it is understood in the Port Royalist, “Cartesian linguistics” tradition (Log, 9:11-13). The second feature is the idea that logic is not only a *propaedeutic* for all the other sciences but also a *presuppositional foundation* (*Grundlage*) for all the other sciences (Log, 9:13). And the third feature is that “logic is to teach us ... that [correct use of the understanding] in which it agrees with itself” and thereby logic is a “self-cognition of the understanding and of reason, not as to their faculties in regard to objects, however, but merely as to form,” such that “[i]n logic the question is only, *How will the understanding cognize itself?*” (Log, 9:14).

This third feature says that pure general analytic logic is how human reason studies its own nature from a formal point of view. Otherwise put, pure general analytic logic for Kant is the science of *the formal essence of human rationality*. As a consequence, for Kant the normativity of logic, as categorical, and correspondingly logic’s role as a pure morality of thinking and reasoning, and as the science of the formal essence of human rationality, is not an accidental feature of pure general analytic logic, but instead *essential* to it. Hence radically unlike the mainstream trend in the history of logic, which tightly ties pure or formal logic to mathematics and to the exact sciences more generally, hence is fully oriented to the primacy of *theoretical* reason over practical reason, Kant’s pure general analytic logic by sharp contrast is fully oriented to the primacy of *practical* reason over theoretical reason.

In addition to the logical source of the irreducible categorical normativity of Kant’s PSR, there are also two further sources of a specifically *human* version of its categorical normativity. The first source flows from Kant’s conception of mathematics as a *moral science* that depends on the specifically human forms of intuition, the representations of space and time, thereby necessarily restricting it to the finite, egocentrically-centered, embodied human standpoint.¹⁷ And the second source flows from the essential roles of biological life, consciousness, and the spontaneous power of choice under practical laws, in causal-deontological-spontaneous sufficiency.

In these regards, the categorical normativity of Kant’s PSR is not only *systematically* but also (unlike most cases of overdetermination in metaphysics) *pleasingly* overdetermined.

And here’s one follow-up point on that, by way of a parallel with recent and contemporary Analytic philosophy. For Kant, like Donald Davidson’s well-known view *that reasons are also causes*,¹⁸ the third and fourth basic modes of sufficient reason are *also* causes. But unlike Davidson’s view, the rational character of sufficient reasons for Kant is *not* supervenient on the underlying deterministic, mechanistic physical causes; on the contrary, the the rational character of the reasons is *the immanent, intrinsic relational structure of the causes, whether deterministic-mechanistic or nondeterministic and non-mechanistic*. Therefore, *all* causation for Kant is *rational causation, even when it's deterministic and mechanistic*, precisely because, for Kant, causation is always inherently governed by logical and mathematical principles and structures, which are themselves inherently rational. This means that Kant’s treatment of reasons and causes is *neo-Aristotelian*, and neither Cartesian-dualistic nor materialist (whether non-reductive or reductive).

¹⁷ See Hanna (2006, chs. 2 and 5; 2002; 2000a).

¹⁸ Davidson (1997).

And that is what I've called *Kant's Neo-Aristotelian Natural Power Grid*.¹⁹ Above all, according to this Kantian/neo-Aristotelian picture, rational causation guided by Kant's PSR, under its basic modes (iii) and (iv), is always *manifestly real, not noumenal*, and in particular rational causation under basic mode (iv) can be directly cognitively accessed by both empirical and pure intuitions (aka essentially non-conceptual cognition), *beyond* the reach of the natural sciences.²⁰

Finally, as I also mentioned in section 1, this logical-normativity-argument, in turn, yields a specifically and uniquely Kantian solution to the well-known circularity or reflexivity problem with respect to the PSR, which consists in asking

- (i) whether the PSR can be applied to itself, or not, and
- (ii) if it can be applied to itself, how this self-application can obtain without vicious circularity or paradoxical reflexivity.

The specifically and uniquely Kantian solution that derives from the logical-normativity-argument says, **first**, that insofar as the PSR is understood in terms of analytic or logical sufficiency and reasonhood, it *cannot* be applied to itself. This is because the PSR in that sense *is the categorically normative condition of the possibility* of all specifically human reasoning and rationality whatsoever, hence it is *transcendentally presupposed* by all such reasoning and rationality, and therefore the PSR in the sense of analytic or logical sufficiency and reasonhood simply does not fall *within* the domain of items *to which* the PSR properly applies. To use language reminiscent of Wittgenstein's *Tractatus*, the PSR in that sense constitutes *the enabling boundaries or limits* of all human reasoning and rationality. Or to use language reminiscent of Tarski's semantic solution to the Liar Paradox via a hierarchy of languages, the PSR in that sense belongs to *the higher-level "meta-language" that provides logico-semantic grounding for the lower-level "first-order language"* of human reasoning and rationality. Hence, **second**, the question of *self-application* never arises for the PSR in this sense, and thus the problems of vicious logical circularity or paradoxical reflexivity never arise for it in this sense either.

5. Conclusion

It should be evident by now that in an illuminatingly important way—perhaps surprisingly!, given the seemingly philosophically antiquarian, esoteric, and Scholastic mien of the PSR as a philosophical concept—understanding the fourfold root of Kant's principle of sufficient reason *is the key to understanding Kant's Critical and post-Critical philosophy as a whole*.²¹

¹⁹ See Hanna (2018b).

²⁰ See also Hanna (2019).

²¹ I'm very grateful to the participants in and organizers of the *Principle of Sufficient Reason Conference* at the University of Hamburg, 12-14 March 2019, especially Magali Roques and Sonja Schierbaum, for the opportunity to present an earlier version of this material and for their extremely helpful critical comments on that presentation.

BIBLIOGRAPHY

- Anderson, R.L. (2015) *The Poverty of Conceptual Truth: Kant's Analytic/Synthetic Distinction and the Limits of Metaphysics* (Oxford: Oxford University Press).
- Benacerraf, P. (1981) 'Frege: the Last Logician', in P. French, et al (eds.), *The Foundations of Analytic Philosophy*, Midwest Studies in Philosophy 6 (Minneapolis, MN: University of Minnesota Press), pp. 17–35.
- Davidson, D. (1997) 'Actions, Reasons, and Causes', in A. Mele (ed.), *The Philosophy of Action* (Oxford: Oxford University Press), pp. 3-19.
- Frege, G. (1953) *Foundations of Arithmetic*, trans. J.L. Austin (Evanston, IL: Northwestern University Press).
- Hall, B. (2014) *The Post-Critical Kant: Understanding the Critical Philosophy Through the Opus Postumum* (London: Routledge).
- Hanna, R. (2019) 'Kant's B Deduction, Cognitive Organicism, the Limits of Natural Science, and The Autonomy of Consciousness', THIS JOURNAL and THIS ISSUE, pp. 29-46 above.
- Hanna, R. (2018a) 'How to Mind the Gaps: On Oliver Thorndyke's Kant's Transition Project and Late Philosophy', *Critique* (2018), available online at URL = <https://virtualcritique.wordpress.com/2018/09/05/review-oliver-thorndykes-kants-transition-project-and-late-philosophy/>.
- Hanna, R. (2018b) 'Kant's Neo-Aristotelian Natural Power Grid: On Kant and the Laws of Nature', *Critique* 2018, available online at URL = <https://virtualcritique.wordpress.com/2018/10/21/new-work-on-kant-iii-kants-neo-aristotelian-natural-power-grid-on-kant-and-the-laws-of-nature/>.
- Hanna, R. (2017) 'Richer Than You Think: Kant, Conceptual Truth, and the Discursive Structure of the Manifest World', *Studi Kantiani* 30:115-122, also available online at URL = https://www.academia.edu/34141203/Richer_Than_You_Think_Kant_Conceptual_Truth_and_the_Discursive_Structure_of_the_Manifest_World_Forthcoming_in_Studi_Kantiani >.
- Hanna, R. (2016) 'Directions in Space, Non-Conceptual Form, and the Foundations of Transcendental Idealism', in D. Schulting (ed.), *Kantian Nonconceptualism* (London: Palgrave Macmillan), pp. 99-115
- Hanna, R. (2006) *Kant, Science, and Human Nature* (Oxford: Clarendon/Oxford University Press).
- Hanna, R. (2002) 'Mathematics for Humans: Kant's Philosophy of Arithmetic Revisited', *European Journal of Philosophy* 10:328-353.

Contemporary Studies in Kantian Philosophy 4 (2019): 47-71.

Hanna, R. (2001) *Kant and the Foundations of Analytic Philosophy* (Oxford: Clarendon/Oxford University Press).

Hanna, R. (2000a) 'The Inner and the Outer: Kant's 'Refutation' Reconstructed', *Ratio* 13:146-174.

Hanna, R. (2000b) 'Kant, Truth, and Human Nature', *British Journal for the History of Philosophy* 8:225-250.

Hanna, R. (1998) 'How Do We Know Necessary Truths? Kant's Answer', *European Journal of Philosophy*, 6:115–145.

Kauark-Leite, P. (2018) 'Kant and Scientific Explanation Beyond Mechanical Causation', in V. L. Waibel, M. Ruffing, and D. Wagner (eds.), *Natur und Freiheit: Akten des XII. Internationalen Kant-Kongresses*, pp. 1261-1267.

Leibniz, G.W. (1989) 'The Principles of Philosophy, Or, The Monadology', in G.W. Leibniz, *Philosophical Essays*, trans. R. Ariew and D. Garber (Indianapolis, IN: Hackett), pp. 213-225.

Longuenesse, B. (2001) 'Kant's Deconstruction of the Principle of Sufficient Reason', *Harvard Review of Philosophy* 9:67–87.

Quine, W.V.O. (1961) 'Two Dogmas of Empiricism', in W.V.O. Quine, *From a Logical Point of View* (2nd edn., New York: Harper and Row), pp. 20–46.

Thorndyke, O. (2018) *Kant's Transition Project and Late Philosophy* (London: Bloomsbury).