

Kant's Biological Theory of Freedom

Robert Hanna

University of Colorado at Boulder, USA

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 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. (*CPR* A534/B562)¹

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 overcome 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. (*CPR* A802-803/B830-831)

The capacity *for desire* is the capacity to be by means of one's representations the cause of the objects of these representations. The capacity of a being to act in accordance with its representations is called *life*. (*MM* 6:211)

THE HUMAN BEING AS A BEING IN THE WORLD, SELF-LIMITED THROUGH NATURE AND DUTY. (*OP* 21: 34)

If one accepts classical physics, free will must apparently be explained as being *compatible* with determinism. The only alternative to compatibilism, if sense is to be made of free will, would be to postulate that the laws of physics do not have universal application and the human free will can cause things to happen contrary to those laws. It might be suggested that Kant found a third alternative, but if so it is one I am unable to understand.

--David Hodgson²

I. Introduction

Kant was the first post-Newtonian philosopher to attempt to face up directly and fully to the basic philosophical problems of free will and Universal Natural Determinism. Prior to the 18th century, philosophers had always addressed issues about free will in the context of either Fatalism or Universal Divine Determinism. And other 18th century post-Newtonian philosophers focused almost exclusively on trying to provide a *phenomenology* of free will as opposed to a metaphysics of free will.³ Furthermore,

neither pre-18th century philosophers nor other 18th century post-Newtonian philosophers had clearly framed the free will problem *both* as a problem about explaining the possibility of free will in a universally determined natural world *and also* as a problem about the compatibility or incompatibility of free will and Universal Natural Determinism. Otherwise put, I think that Kant was the first philosopher to see clearly and distinctly that Incompatibilism is consistent with Hard Determinism and Metaphysical Libertarianism *alike*, both of which are mistakenly committed to a deep philosophical mistake—the failure to distinguish fundamentally between appearances and things-in-themselves—which in turn leads to a logico-metaphysical antinomy, and inevitably hides the real philosophical significance of Compatibilism. Kant was unique in trying to address *both* the metaphysics and the phenomenology of free will in the post-Newtonian context of Universal Natural Determinism, *and also* the classical Compatibilism vs. classical Incompatibilism dilemma. In this paper, I will focus on explaining and defending a new interpretation of what Kant calls “transcendental freedom,” an interpretation which I call *Kant’s Biological Theory of Freedom*

II. Kant’s *Theories* of Freedom

Kant’s theory of *transcendental* freedom is his metaphysics of free will. Transcendental freedom is how a rational animal or person can, “from itself” (*von selbst*) (*CPR* A533/B561), be the spontaneous mental cause of certain natural events or processes. If I am that rational animal or person, then insofar as I am transcendently free, it follows that certain events or processes in physical nature are *up to me*—or to use Kant’s own phrase, *in meiner Gewalt* (literally: “in my control” or “in my power”; *CPR* 5: 94-95). So otherwise put, transcendental freedom is deep freedom of the will, or up-to-

me-ness (as it were, *In-Meiner-Gewalt-Sein*). I will argue, contrary to standard interpretations,⁴ that Kant's theory of transcendental freedom entails neither a classically *Incompatibilist* Timeless Agency theory nor a classically *Compatibilist* Regulative Idea theory, and also that it thereby constitutes what David Hodgson aptly calls a "third alternative" to the all-too-familiar and seemingly exhaustive dichotomy between Incompatibilism and Compatibilism, namely an "Incompatibilistically Compatibilist" or "Post-Compatibilist" alternative, namely what I am calling Kant's Biological Theory of Freedom.

Transcendental freedom must be contrasted with what Kant calls *practical* freedom. Practical freedom *presupposes* transcendental freedom, and can be defined in a negative way as the independence of first-order volition, or the "power of choice" (*Willkür*), from necessitation by sensible impulses (*CPR* A533/B561). But practical freedom is also necessarily equivalent to what Kant calls "autonomy": "the moral law expresses nothing other than the *autonomy* of pure practical reason, that is, [practical] freedom" (*CPrR* 5: 33). Practical freedom in this positive sense, or autonomy, is how a transcendently free person can choose or do things by means of her subjective experience or consciousness of recognizing the Categorical Imperative or moral law as a desire-overriding, strictly universal, a priori, categorically normative, non-instrumental practical reason that has both motivating *and* justifying force. The fact of this subjective experience or consciousness of autonomous agency is what Kant also calls the "fact of reason" (*Faktum der Vernunft*):

The consciousness of this fundamental law [of pure practical reason, which says: so act that the maxim of your will could always hold at the same time as a principle of universal law giving] may be called a fact of reason, since one cannot ferret it out from antecedent data of reason, *such as the consciousness of freedom* (for this is not antecedently given), and since it forces itself upon us as a synthetic proposition a priori based on no pure or empirical intuition... In order to regard this law

without any misinterpretation as given, one must note that it is not an empirical fact, but the sole fact of pure reason, which by it proclaims itself as originating law. (*CPrR* 5: 31, underlining added).

So otherwise put, practical freedom in the positive sense, or autonomy, is *moral causation*.

As I noted already, Kant was the first post-Newtonian theorist of free will to try to face up directly and fully to the two basic free will problems. It is well known to contemporary Kantians, however, that in scholarly space there exist at least two sharply distinct *versions* of Kant's theory of transcendental freedom, each of which has a fairly solid grounding in Kant's texts: The Timeless Agency Theory,⁵ and The Regulative Idea Theory.⁶

The Timeless Agency Theory adopts the classical Two World or Two Object Theory of the noumena vs. phenomena distinction and asserts that a noumenal subject is autonomous in that it has absolutely spontaneous causal efficacy, or nomological sufficiency of the self-legislating positively noumenal will, apart from all alien causes and all sensible impulses, from or for the sake of the Categorical Imperative, by causing, *from outside of time and space*, phenomenal human behavioral movements (in outer sense) and psychological processes (in inner sense) that are themselves independently necessarily causally determined by general causal laws of nature plus the settled empirical facts about the past. The Timeless Agency Theory is supported primarily by texts drawn from the *Critique of Pure Reason* (esp. *CPR* A538-558/B566-586).

By contrast, The Regulative Idea Theory adopts the neoclassical Two Aspect or Two Standpoint Theory of the noumena vs. phenomena distinction and says that we are required by our innate capacity for practical reason to believe or take ourselves to be acting morally only under the rational idea of our own practical freedom or autonomy.

The Regulative Idea Theory is supported primarily by section III of *Groundwork of the Metaphysics of Morals*.

Both The Timeless Agency Theory and The Regulative Idea Theory have some serious problems.

On the one hand, it is crucial to note that the texts which best support the Timeless Agency Theory are explicitly said by Kant to demonstrate only the bare conceivability and logical consistency of the notions of freedom and Universal Natural Determinism, and neither the *reality* nor the *real* (i.e., synthetic, non-logical, or strong metaphysical) *possibility* of freedom:

Do freedom and natural necessity in one and the same action contradict each another? And this we have answered sufficiently when we showed that since in freedom a relation is possible to conditions of a kind entirely different from those in natural necessity, the law of the latter does not affect the former; hence each is independent of the other, and can take place without being disturbed by the other.... It should be noted here that we have not been trying to establish the reality of freedom, as a faculty that contains the causes of appearance in our world of sense.... Further, we have not even tried to prove the possibility of freedom; for this would have not succeeded either, because from mere concepts *a priori* we cannot cognize anything about the possibility of any real ground or any causality. (CPR A557-558/B585-586, underlining added)

Correspondingly, the most serious problem with The Timeless Agency Theory is that it is really (i.e., synthetically, non-logically, or strongly metaphysically) *impossible*. If all phenomenal events are all independently necessarily determined by natural laws together with antecedent facts, then the noumenal causality of the will implies the non-standard causal overdetermination of phenomenal human behavioral movements in outer sense and psychological processes in inner sense. The Thesis of Non-Standard Causal

Overdetermination says:

- (i) There can be two ontologically distinct nomologically sufficient causes of the same event, such that one of the two causes is physical and one of the two causes is non-physical, and each of which can operate in the absence of the other,

and correspondingly

(ii) There can be two complete and independent causal explanations of the same event. But as Jaegwon Kim has compellingly argued, it seems entirely reasonable to hold that if there already exists a nomologically sufficient physical cause of some event, and if correspondingly a complete and independent physical causal explanation of that same event also exists, then this cause and this causal explanation together necessarily *exclude* there being any other distinct nomologically sufficient cause or distinct causal explanation of the same event.⁷ So the non-standard causal overdetermination implied by The Timeless Agency Theory, although barely conceivable and logically possible, is really (i.e., synthetically, non-logically, or strongly metaphysically) a priori ruled out.

On the other hand, it is also crucial to note that the texts which best support The Regulative Idea Theory are explicitly said by Kant to demonstrate only that “freedom must be presupposed (*vorausgesetzt*) as a property of the will of all rational beings” (*GMM* 4: 447, underlining added) and that “all human beings think of themselves as having free will” (*GMM* 4: 455, underlining added). Correspondingly, the most serious problem with The Regulative Idea Theory is that even if it is true, it just does not do the philosophical work required of the noumenal causation *vs.* phenomenal causation distinction, precisely because it does not entail either the reality or the real (i.e., synthetic, non-logical, or strong metaphysical) possibility of freedom of the will but rather entails only at best our *belief* in its reality or real possibility, which is not only ontologically deflationary but also, arguably, does not even rationally justify that belief. In fact, our belief in freedom is only a self-evident *practical* belief—in effect, a moral certainty—which according to Kant is a sufficiently warranted practical commitment that is nevertheless held on theoretically insufficient grounds:

Only in a **practical relation**, however, can taking something that is theoretically insufficient to be true be called believing (*Glauben*). This practical aim is either that of **skill** or **morality**, the former for arbitrary and contingent ends, the latter, however, for absolutely necessary ends. (CPR A823/B851, underlining added)

So this moral certainty could still be theoretically *wrong*. For all we know, and for all that The Regulative Idea Theory says, we could *still* be nothing but deterministic natural automata, or fleshy deterministic Turing machines, in the tragic situation of epiphenomenally dreaming that we are free, and with no more causal power of our own than a turnspit:

[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. (CPrR 5: 97, underlining added)

For these reasons, it seems to me that both The Timeless Agency Theory and The Regulative Idea Theory are very likely to be objectively false, whatever else we may think about the question of which theory most accurately reflects Kant’s own considered views about freedom of the will.

In contrast to these theories, as I have mentioned, I want to develop and defend Kant’s Biological Theory of Freedom.⁸ Just like The Timeless Agency Theory and The Regulative Idea Theory, The Biological Theory also has a fairly solid grounding in Kant’s texts, although it is primarily supported by texts drawn from what I like to call the “Post-Critical” period after 1787,⁹ especially including the *Critique of the Power of Judgment* and the *Opus postumum*. But philosophically it differs sharply from the other two theories in that it avoids their serious problems and also, in my opinion, is arguably quite close to being objectively true. So I think that we should prefer it both on grounds

of inference to the most rationally charitable interpretation, and also for independent philosophical reasons. Above all, however, The Biological Theory shows how transcendental freedom of the will in the occurrent sense—deep freedom or up-to-me-ness, as actually occurring—can also be a natural dynamic process. If I am correct, then this makes Kant a defender of Liberal Naturalism,¹⁰ which says that the irreducible but also non-dualistic mental properties of minded animals are as basic in nature as biological properties, and metaphysically continuous with them. In other words, rational mindedness *grows naturally* in the physical world. This Liberal Naturalism follows directly from Kant's Transcendental Idealism, which, as I have proposed in the first of these three essays, and have argued for explicitly elsewhere,¹¹ is most charitably interpreted as Weak or Counterfactual Transcendental Idealism, or WCTI, and says that a necessary condition of the manifest natural world's existence is its knowability by rational human animals, at least to some extent, *were* they to exist, and which also directly implies a necessary conformity between basic structures of manifest nature on the one hand and basic structures of rational human cognition on the other. But an even more direct way of seeing Kant's Liberal Naturalism is to recognize that human persons for him are rational human living organisms or animals whose rational mindedness is fully metaphysically continuous with their animality:

The human being, as animal, belongs to the world, but, as person, also to the beings who are capable of rights—and, consequently, have *freedom* of the will. Which ability essentially differentiates [the human being] from all other beings; *mens* is innate to [the human being]. (*OP* 21:36)

III. Transcendental-Freedom-in-Life

Kant's theory of transcendental freedom is based on his notion of "spontaneity." For him, X is spontaneous if and only if X is a conscious mental event that expresses some acts or operations of a creature, and X is

- (i) causal-dynamically *necessarily unprecedented*, in the two-part sense that
 - (ia) conscious mental events of those specific sorts have never actually happened before,
- and
- (ib) the settled empirical facts about the past together with the general causal mechanical laws of nature do not provide nomologically sufficient conditions for the existence or specific character those conscious mental events,
- (ii) *underdetermined* by external sensory informational inputs, and also by prior desires, even though it may have been triggered by those very inputs or motivated by those very desires
- (iii) *creative* in the sense of being recursively constructive, or able to generate infinitely complex outputs from finite resources,

and also

- (iv) *self-guiding*. (*CPR* A51/B75, B130, B132, B152, A445-447/B473-475)

Furthermore, spontaneity can be either *relative* or *absolute*. Relative spontaneity requires inputs to the conscious mind, whereas absolute spontaneity allows the conscious mind to generate its own outputs without any triggering inputs. For example, human a priori cognition is only relatively spontaneous, because it requires sensory inputs via empirical intuition, whereas an *intellectual intuition*, if it existed, would be absolutely spontaneous, because it could cause the objects of its thoughts to exist just by thinking them (*CPR* A19-22/B33-36, B71-72). Now according to Kant, the concept of a cause analytically entails the concept of its effect, and the general schematized pure concept of CAUSE says that something X (the cause) necessitates something else Y (its effect) in time

according to a necessary rule or law. Or equivalently, according to Kant, to say that X causes its effect Y is to say that X is nomologically sufficient for Y in time (*CPR* B112, A144/B183). Then X is a relatively or absolutely spontaneous cause of its effect Y if and only if

(1) X is nomologically sufficient for Y in time,

and

(2) X is a conscious mental event that is necessarily unprecedented, underdetermined by external sensory inputs and desires, creative, and self-guiding.

Finally, absolutely spontaneous mental causation is the same as transcendental freedom:

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, which, first, contains nothing borrowed from experience, and second, the object of which cannot be given determinately in any experience.... But since in such a way no absolute totality of [natural] conditions in causal relations is forthcoming, reason creates the idea of a spontaneity, which could start to act from itself, without needing to be preceded by any other cause that in turn determines it to action according to the law of causal connection. (*CPR* A533/B561, underlining added)

Although transcendental freedom is a particularly robust kind of *mental* causation, in the second *Critique* Kant sharply distinguishes transcendental freedom from mere *psychological* freedom:

These determining representations [i.e., instincts or motives] themselves have the ground of their existence in time and indeed in the *antecedent state*, and in a preceding state, and so forth, these determinations may be internal and they may have psychological instead of mechanical causality, this is, produce actions by means of representations and not by bodily movements; they are always determining grounds of the causality of a being insofar as its existence is determinable in time and therefore under conditions of past time, which are thus, when the subject is to act, no longer within his control and which may therefore bring with them psychological freedom (if one wants to use this term for a merely internal chain of representations in the soul) but nevertheless natural necessity, leaving no room for *transcendental freedom* which must be thought of as independence from everything empirical and so from nature generally, whether regarded as an object of inner sense in time only or also as an object of outer sense in both space and time; without this freedom (in the latter and proper sense), which alone is practical a priori, no moral law is possible and no imputation in accordance with it. (*CPrR* 5: 96-97, underlining added)

Otherwise put, psychological freedom is the subject's subjective experience or consciousness of choosing or acting without being prevented, and without inner or outer

compulsion. As Kant explicitly points out, and as Hume and Leibniz also noted in anticipation of contemporary Compatibilism, it is both logically and metaphysically possible to be psychologically free without being transcendently free. Nevertheless, as we have seen, according to Kant psychological freedom without transcendental freedom, is nothing but the epiphenomenal dream of a human “turnspit” (*CPrR* 5: 97). So psychological freedom is not a sufficient condition of transcendental freedom. Nevertheless, according to Kant psychological freedom remains a *necessary* condition of transcendental freedom. And this seems independently highly plausible. No one could be transcendently free and also at the same time undergo the subjective experience or consciousness of being prevented from choosing or acting, or of being inwardly or outwardly compelled to choose or act. Indeed, as the second Analogy of Experience explicitly shows, and as we saw in the second essay in this trilogy, psychological freedom is necessarily built into the mental representation of *any* objective causal sequence, via what Kant calls the “the **subjective sequence** of apprehension,” whose ordering is always subjectively experienced as “entirely arbitrary” (*ganz beliebig*) and not necessitated (*CPR* A193/B238).

When we ascribe transcendental freedom specifically to the will of a rational human animal or human person, then in addition to the positive factor of absolute spontaneity, which confers deep freedom or up-to-me-ness on the person’s choices and acts, and psychological freedom, which guarantees the subjective experience or consciousness of being unprevented and uncompelled in one’s choices and acts, there is also a negative dimension of freedom which guarantees the person’s choices and acts

occur independently of all “alien causes,” that is, independently of all pathological inner and unowned outer sources of nomologically sufficient compulsion:

The will is a kind of causality that living beings have so far as they are rational. Freedom would then be that property whereby this causality can be active, independently of alien causes determining it; just as *natural necessity* is a property characterizing the causality of all non-rational beings—the property of being determined to activity by the influence of alien causes. The above definition of freedom is *negative*. (*GMM* 4: 446, underlining added)

This is where practical freedom comes back onto the scene. Practical freedom presupposes but also exceeds transcendental freedom, in that practical freedom is the absolute spontaneity of the will independently of all alien causes and also independently of *all sensible impulses* (empirical desires):

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 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. (*CPR* A534/B562, underlining added)

As I mentioned above, however, this is merely a negative characterization of practical freedom. As positively characterized, practical freedom also involves the capacity for *self-legislation* in conformity with the Categorical Imperative or moral law. Or in other words, practical freedom is necessarily equivalent with autonomy (*GMM* 4: 440-441, 446-463).

It may seem, on the face of it, that there is and should be no direct connection whatsoever between the person’s absolutely spontaneous, psychologically free, autonomous will, and her actual existence in physical nature. Indeed, that is the basic idea behind the classical theory of Agent Causation, according to which the freely willing person necessarily stands *outside* the natural causal order in spacetime.¹² And Kant is

often cited as a paradigmatic defender of the Agent Causation theory—as per The Timeless Agency Theory.¹³ But in fact Kant himself explicitly asserts otherwise:

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 overcome 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. (CPR A802-803/B830-831, underlining added)

Two things fill the mind with ever new and increasing admiration and reverence, the more often and more steadily one reflects on them: the starry heavens above me [i.e., nature] and the moral law within me [i.e., freedom]. I do not need to search for them and merely conjecture them as though they were veiled in obscurity or in the transcendent region beyond my horizon; I see them before me and connect them immediately with the consciousness of my existence. (CPrR 5: 161-162, 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 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. (CPJ 5: 176, underlining added)

In other words, Kant is explicitly saying that transcendental freedom is both really (i.e., synthetically, non-logically, or strongly metaphysically) possible and real. I will now reconstruct Kant's reasoning for this perhaps surprising thesis, and in so doing, argue that his theory of transcendental freedom can be very plausibly—and philosophically most defensibly—interpreted as a biologically-based theory: in effect, as *transcendental-freedom-in-life*. As I mentioned above, I shall be drawing primarily on texts from Kant's Post-Critical period after 1787, and in particular from the third *Critique*.

Kant argues in the two Introductions and the second half of the *Critique of the Power of Judgment* that the concepts LIFE and ORGANISM, and in particular the concept of a “natural purpose” (*Naturzweck*) or living organism, are not ordinary

empirical concepts of matter, and that they invoke a type of causation which cannot be known in the classical Newtonian mechanistic physics:

For a body to be judged as a natural purpose in itself and in accordance with its internal possibility, it is required that its parts reciprocally produce each other, as far as both their form and their combination is concerned, and thus produce a whole out of their own causality, the concept of which, conversely is in turn the cause (in a being that would possess the causality according to concepts appropriate for such a product) of it in accordance with a principle; consequently the connection of **efficient causes** could at the same time be judged as an **effect though final causes**. In such a product of nature each part is conceived as if it exists only **through** all the others, thus as if existing **for the sake of the others** and **on account** of the whole, i.e., as an instrument of art, and thus represented as possible at all only as a purpose); rather it must be thought of as an organ that **produces** the other parts (consequently each produces the others reciprocally), which cannot be the case in any instrument of art, but only of nature, which provides all the matter for instruments (even those of art): only then and on that account can such a product, as an **organized and self-organizing** being, be called a **natural purpose**. (CPJ 5: 373-374, underlining added)

Strictly speaking, the organization of nature is ... not analogous with any causality that we know. (CPJ 5: 375, underlining added)

Now because the causality of living organisms is scientifically unknowable, the basic concepts of biology are merely “regulative” or “hypothetical” concepts of reason, that is, heuristic, logical-fictional, or analogical concepts employed in our everyday encounters with the natural world, and for the unification and promotion of natural scientific inquiry (CPJ 5: 369-415; see also CPR A642-647/B670-675).¹⁴ But it does not follow at all from this that on fully Kantian grounds the organismic life (in particular, the organismic life of my own animal body) cannot *also* be directly cognized by essentially non-conceptual, non-propositional, or non-epistemic/non-judgment-based means. Indeed, according to Kant in the First Part of the third *Critique*, the feelings of pleasure and pain, bodily affects including bodily desires and drives, and proprioceptive feelings, constitute “the feeling of life” (CPJ 5: 204, 278), or the feeling of embodied vitality. Furthermore, according to Kant there is an essential connection between the affective-emotional psychological life of my mind, including intentional agency, and the biological life of my own body:

[L]ife is the subjective condition of all our possible experience. (*P* 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. (*CPJ* 5: 278, underlining added)

The capacity *for desire* is the capacity to be by means of one's representations the cause of the objects of these representations. The capacity of a being to act in accordance with its representations is called *life*. (*MM* 6:211, underlining added)

So in this way, according to Kant, biological life is metaphysically continuous with the essentially embodied conscious, intentional, caring, rational human mind. Otherwise put, our essentially non-conceptual phenomenal affective-emotional consciousness in inner sense synthetically a priori (i.e., non-logically or strongly metaphysically a priori necessarily) entails embodied biological life—that is, conscious, intentional, caring, rational human beings are, synthetically a priori, also living organisms. So our rationality also grows naturally in our living organismic bodies. Kant is *not* saying that necessarily, everything that is biologically alive and is an organism, is itself conscious or minded—i.e., that necessarily, all organismic life is minded or mind, which is clearly false. Rather he is saying that *necessarily, everything that is conscious or minded is alive and is an organism*—i.e., that necessarily, all mind is alive or life, which is arguably true.

This is a crucially important point. The semantic and epistemic constraints that Kant places on teleological judgments about distal material objects in space in the context of biological science—namely, that such judgments are always “regulative” and not “constitutive”—do *not* in fact apply to the human conscious experience of living organismic embodiment, which is essentially non-conceptual or intuitional, and affective-emotional in character, and therefore essentially *not* conceptual, propositional, or epistemic/judgmental. So there is an important Kantian distinction to be drawn between

(i) teleological *judgments*, which are inherently neither directly referential nor existentially-committed, because they are inherently based on concepts and also regulative,

and

(ii) teleological *faith* or *Glaube*, via essentially non-conceptual *inner sense intuitions* or first-order phenomenally conscious states, which are inherently both directly referential and also existentially committed.

In other words, when Kant so famously remarks in the B Preface that

I had to deny **scientific knowing** (*Wissen*) in order to make room for **faith** (*Glauben*). (*CPR* Bxxx; see also A828/B856),

what he most fundamentally means is that he had to deny the possibility of *teleological scientific judgment* in order to make room for *teleological faith in his own biological life as a conscious, rational human intentional and moral agent in the manifest natural world*. In this way, even if teleological judgments are only regulative, I can still have an essentially non-conceptual, non-propositional, non-epistemic/non-judgment-based teleologically faithful *biophenomenology* that is fully constitutive.

If so, then for Kant there are real biological facts in nature. It is just that I cannot scientifically know them via physical theories that are exclusively based on naturally mechanistic principles. But I can still truly consciously feel at least some of them, precisely by consciously feeling my own essentially embodied animal life. Most importantly of all, by way of teleological faith, according to Kant I can consciously feel my own transcendental freedom:

Sensible life has, with respect to the *intelligible* consciousness of its existence, (consciousness of *freedom*), the absolute unity of a phenomenon, which, so far as it contains merely appearances of the disposition that the moral law is concerned with (appearances of the character), must be appraised not in accordance with the natural necessity that belongs to it as appearance but in accordance with the absolute spontaneity of freedom. (*CPrR* 5: 99, underlining added)

In view of these points, Kant regards empirical psychology as a constitutive and nomological yet also non-deterministic and non-mechanistic Life Science of the mind. Even though psychology contains unique “psycho-psycho” laws which strictly govern the phenomenological facts of inner sense¹⁵—which, we now recognize, must also be actual biological facts—nevertheless mental phenomena cannot be arithmetically analyzed because, as we have already seen, their merely subjective temporal ordering in inner sense is “entirely arbitrary” (*ganz beliebig*) (*CPR* A193/B238) according to the desires and choices of the conscious, intentional, caring, rational human animal. That is, the radical open-endedness of possible orderings in inner sense means that the set of all mental phenomena cannot be put into a one-to-one correspondence with the set of natural numbers, or reconstructed as Turing-computable functions of Primitive Recursive Arithmetic, or PRA. Kant’s conception of mathematics, together with the Axioms of Intuition and the Anticipations of Perception—i.e., the *mathematical* synthetic a priori principles of pure understanding (*CPR* A160-162/B199-201)—and the Analogies of Experience, show that Universal Natural Determinism as Kant understood it requires the simple primitive recursive arithmetization of causal processes in time. But given Gödel’s incompleteness theorems together with The Church-Turing Thesis, it is also arguable that mathematical truth, whether in Peano Arithmetic or in richer systems of arithmetic including enough of the axioms of Peano Arithmetic, is inherently uncomputable or non-mechanical. Thus for Kant psychological laws cannot be either deterministic or more generally naturally mechanical:

The empirical doctrine of the soul must always remain ... removed ... from the rank of what may be called a natural science proper. This is because mathematics is inapplicable to the phenomena of the inner sense and their laws... It can, therefore, never become anything more than a historical (and, as such, as much as possible) systematic natural doctrine of the inner sense, i.e., a natural description of the soul, but not a science of the soul. (*MFNS* 4: 471, underlining added)

Furthermore since embodied human mental life synthetically a priori entails biological life, it follows directly from Kant's transcendental-freedom-in-life thesis, together with Gödel's incompleteness theorems and The Church-Turing Thesis, that that there can never be a Newton, a Church, or a Turing of biological life, and also that there can also never be a Newton, a Church, or a Turing of transcendental freedom.¹⁶ So again, our psychological life, especially including our power of choice or *Willkür*, cannot be naturally mechanized,¹⁷ because instead it grows naturally and is an inherently uncomputable process.

IV. Conclusion

According to Kant's Biological Theory of Freedom, even if all the inert, non-living parts of material nature, as metaphysically described by the three Analogies of Experience, fall under the general causal laws of classical Newtonian physics, nevertheless the existence of these deterministic natural automata is *fully consistent with* and indeed is also *metaphysically presupposed by* the instantiation of an irreducibly different set of properties in the living organism that is the rational human animal. This is a set of irreducible, inherently non-mechanical, uncomputable conscious, intentional, caring, a priori, and categorically normative properties, whose precise pattern of instantiations constitutes both that animal's power of choice and also its transcendental and practical freedom of the will, or its autonomy. Therefore our rational human animal body movements are caused by our transcendental freedom, which is a non-empirical or a priori but still fully natural biological and neurobiological fact about us. Rational human animals are *not* deterministic (or indeterministic) natural automata or machines, and they therefore *are* inherently non-mechanical, non-Turing-computable living organisms of a

suitably complex kind. Indeed, in the *Critique of Practical Reason* Kant explicitly asserts that rational personhood (*Persönlichkeit*) itself is just

freedom and independence from the mechanism of nature regarded as a capacity of a being subject to special laws (pure practical laws given by its own reason). (*CPrR* 5: 87)

In this way, the difference between, on the one hand,

(i) the general causal mechanical laws of nature,

with which the categorically normative moral laws of human action are *inconsistent* when applied to one and the same event of rational human choosing or acting (*CPrR* 5: 94-95), and on the other hand,

(ii) the non-deterministic, non-mechanical one-off or one-time-only laws of rational human animal movement,

with which categorically normative moral laws are perfectly and indeed inherently *consistent* when applied to one and the same event of rational human animal choosing or acting—since both transcendental freedom and practical freedom alike require the strict *underdetermination* of a rational human animal’s choosing and acting by general causal mechanical laws together with the settled facts about the past—is itself the metaphysical core of Kant’s Biological Theory of Freedom, and what I call “Kant’s Incompatibilistic Compatibilism” or “Kant’s Post-Compatibilism.”

NOTES

¹ For convenience, I cite Kant's works infratextually in parentheses. The citations include both an abbreviation of the English title and the corresponding volume and page numbers in the standard "Akademie" edition of Kant's works: *Kants gesammelte Schriften*, edited by the Königlich Preussischen (now Deutschen) Akademie der Wissenschaften (Berlin: G. Reimer [now de Gruyter], 1902-). I generally follow the standard English translations, but have occasionally modified them where appropriate. For references to the first *Critique*, I follow the common practice of giving page numbers from the A (1781) and B (1787) German editions only. Here is a list of the relevant abbreviations and English translations:

- CPJ* *Critique of the Power of Judgment*. Trans. P. Guyer and E. Matthews. Cambridge: Cambridge Univ. Press, 2000.
- CPR* *Critique of Pure Reason*. Trans. P. Guyer and A. Wood. Cambridge: Cambridge Univ. Press, 1997.
- CPrR* *Critique of Practical Reason*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. Cambridge: Cambridge Univ. Press, 1996. Pp. 133-272.
- GMM* *Groundwork of the Metaphysics of Morals*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. Pp. 37-108.
- MM* *Metaphysics of Morals*. Trans. M. Gregor. In *Immanuel Kant: Practical Philosophy*. Pp. 353-604.
- OP* *Immanuel Kant: Opus postumum*. Trans. E. Förster and M. Rosen. Cambridge: Cambridge Univ. Press, 1993.
- P* *Prolegomena to Any Future Metaphysics*. Trans. J. Ellington. Indianapolis, IN: Hackett, 1977.

² D. Hodgson, "Quantum Physics, Consciousness, and Free Will," in R. Kane (ed.), *The Oxford Handbook of Free Will* (Oxford: Oxford Univ. Press, 2002), pp. 85-110, at p. 86.

³ See, e.g., J. Harris, *Of Liberty and Necessity: The Free Will Debate in Eighteenth Century British Philosophy* (Oxford: Clarendon/OUP, 2005).

⁴ For some recent attempts to reconstruct Kant's theory of freedom, see H. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge Univ. Press, 1990); R. Hanna, *Kant, Science, and Human Nature* (Oxford: Clarendon/OUP, 2006), ch. 8; R. Hanna, "Kant, Causation, and Freedom," *Canadian Journal of Philosophy* 36 (2006): 281-306; R. Hanna, "Freedom, Teleology, and Rational Causation," in *Kant Yearbook* 1 (2009): 99-142; H. Hudson, *Kant's Compatibilism* (Ithaca, NY: Cornell Univ. Press, 1990); D. Pereboom, "Kant on Transcendental Freedom," *Philosophy and Phenomenological Research* 73 (2006): 537-67; E. Watkins, *Kant and the Metaphysics of Causality* (Cambridge: Cambridge Univ. Press, 2005); and A. Wood, "Kant's Compatibilism," in A. Wood (ed.), *Self and Nature in Kant's Philosophy* (New York: Cornell Univ. Press, 1984), pp. 73-101

⁵ See, e.g., Allison, *Kant's Theory of Freedom*, pp. 47-53; Pereboom, "Kant on Transcendental Freedom"; Watkins, *Kant and the Metaphysics of Causality*, chs. 5-6; and Wood, "Kant's Compatibilism."

⁶ See, e.g., *Kant's Theory of Freedom*, ch. 13; and A. Wood, *Kant's Ethical Thought* (Cambridge: Cambridge Univ. Press, 1999), pp. 180-182.

⁷ See J. Kim, "Mechanism, Purpose, and Explanatory Exclusion," in J. Kim, *Supervenience and Mind* (Cambridge: Cambridge Univ. Press, 1993), pp. 237-264. For other versions of Kim's exclusion argument, some of which turn on the assumption of supervenience and the principle of the causal closure of the physical (which says that if an event has a cause, then it has a physical cause only), see also J. Kim "The Myth of Nonreductive Materialism," in Kim, *Supervenience and Mind*, pp. 265-284; J. Kim "The Non-Reductivist's Troubles with Mental Causation," in Kim, *Supervenience and Mind*, pp. 336-357; and J. Kim, *Physicalism, or Something Near Enough*. (Princeton, NJ: Princeton Univ. Press, 2005), ch. 3.

⁸ Strictly speaking, Kant's Biological Theory of Freedom is just one part of a more comprehensive interpretation of Kant's theory of freedom that I call *The Embodied Agency Theory*. See Hanna, *Kant, Science, and Human Nature*, ch. 8.

⁹ See, e.g., Hanna, "Kant, Causation, and Freedom."

¹⁰ See also, e.g., G. Rosenberg, *A Place for Consciousness* (New York: Oxford Univ. Press, 2004), pp. 8-10.

¹¹ See, e.g., Hanna, *Kant, Science, and Human Nature*, section 6.1.

¹² See, e.g., R. Chisholm, "Human Freedom and the Self," in G. Watson (ed.), *Free Will* (2nd edn., Oxford: OUP, 2003), pp. 26-37; R. Clarke, "Agent Causation and Event Causation in the Production of Free Action," *Philosophical Topics* 24 (1996): 19-48; and T. O'Connor, *Persons and Causes* (New York: Oxford Univ. Press, 2000).

¹³ See, e.g., Watkins, *Kant and the Metaphysics of Causality*.

¹⁴ See, e.g., A. Breitenbach, "Two Views on Nature: A Solution to Kant's Antinomy of Mechanism and Teleology," *British Journal for the History of Philosophy* 16 (2008): 351-369; A. Breitenbach, "Teleology in Biology: A Kantian Perspective," *Kant Yearbook* 1 (2009): 31-56; H. Ginsborg, "Kant on Understanding Organisms as Natural Purposes," in E. Watkins (ed.), *Kant and the Sciences* (New York: Oxford Univ. Press, 2001), pp. 231-258; P. Guyer, *Kant's System of Nature and Freedom* (Oxford: Oxford Univ. Press, 2005), chs. 5 and 13; and J. Kreines, "The Inexplicability of Kant's *Naturzweck*: Kant on Teleology, Explanation, and Biology," *Archiv für Geschichte der Philosophie* 87 (2005): 270-311.

¹⁵ For Kant, laws do not have to be semantically insensitive to contextual conditions or mentalistic facts in order to be necessary and strict, since they can also be *non-logically or synthetically necessary, that is, restrictedly necessary*. See Hanna, *Kant and the Foundations of Analytic Philosophy*, ch. 5. Fodor calls such psychological laws "ceteris paribus laws": see his "Making Mind Matter More," in J. Fodor, *A Theory of Content and Other Essays* (Cambridge: MIT Press, 1990), 137-159. Where Kant and Fodor would strongly disagree is that for Kant, these synthetically necessary psychological laws are wholly particular and *one-time-only* or "one-off," not general, whereas for Fodor they must be general laws.

¹⁶ See also J.R. Lucas, *The Freedom of the Will* (Oxford: Clarendon/Oxford Univ. Press, (Oxford: Clarendon/Oxford Univ. Press, 1970), chs. 24-30; and J.R. Lucas, "Minds, Machines, and Gödel," *Philosophy* 36 (1961): 112-127.

¹⁷ See also K. Westphal, *Kant's Transcendental Proof of Realism* (Cambridge: Cambridge Univ. Press, 2004), pp. 229-243.