BÉATRICE LONGUENESSE AND NED BLOCK *VIDE* KANT: ON A NOVEL KANTIAN APPROACH TO THE PHILOSOPHY OF MIND

Ekin Erkan

ABSTRACT: Understanding, for Kant, does not intuit, and intuition—which involves empirical information, i.e., sense-data—does not entail thinking. What is crucial to Kant's famous claim that intuitions without concepts are blind and concepts without intuitions are empty is the idea that we have no knowledge unless we combine concepts with intuition. Although concepts and intuition are radically separated mental powers, without a way of bringing them together (i.e., synthesis) there is no knowledge for Kant. Thus Kant's metaphysical-scientific dualism: (scientific) knowledge is limited to the world of phenomena—the world of appearance—while the thing-in-itself is unknowable because there is no intuition that can correspond to the it. This paper sets to cull Béatrice Longuenesse's recent work on the first-person 'I' and put forth a novel Kantian approach to the first-person reference of mental states, working in the tradition of P.F. Strawson, Rudolf Carnap, and Wilfrid Sellars, while offering an empirical study of deafferentation to ground the thesis that the binding of representations is separate from phenomenal consciousness. Accordingly, we stake a kind of self-consciousness vis-à-vis the Fundamental Reference-Rule qua apperception that, while intimately connected to consciousness of one's own body, apperception is nevertheless distinct from it and is, moreover, the condition for any use of 'I'. We compliment neuroscientist Oliver Sacks' research with Ned Block's recent work on "no-post-perceptual cognition" and attentional variation to couch Kant's schema in perceptual psychology. We navigate Kant's work on intuitions (viz. sensation, perception, and the empirical side of the cognitive faculties) and indirect realism/weak representationism first via Sellars' naturalization of Kant's inaccessible thing-in-itself, before challenging Sellars' functionalist and inferentialist conception of perception and discursive intentionality. We ultimately endeavor to conceptualize the limit-conditions regarding our having reportable knowledge about our access to percepts and concepts.

KEYWORDS: Kant; Philosophy of mind; Strawson; Sellars; Reason; Noumena

I

NATURALIZED NOUMENA: MORALITY AND MENTALITY IN
STRAWSON, SELLARS, AND LONGUENESSE

I

In the broadest possible sense, realism can be understood as either
the metaphysical thesis that there exists a world independently of our feelings,
sensings, thoughts, beliefs, and attitudes or as the epistemological thesis that the
structure of such a mind-independent world, granting its existence, is knowable
to some extent. Idealism, the polar opposite of realism, can be understood as the
thesis that the external world is fundamentally mind-dependent. Kant's
transcendental idealism, taking up the fundamental principles behind Descartes'
rationalism and the empiricism of Locke, Berkeley, and Hume, can be seen as
splitting the divide between realism and antirealism by offering the best of both
worlds: self-consciousness reveals that we are fundamentally embodied subjects
wherein the immediate objects of perception correspond to ideas of the
perceiving subject. *Pace* Descartes' skepticism, Kant takes this thesis to its
conclusion: we have no direct knowledge of reality in-itself—as, following Kant's
account, *appearances* are not extrinsic properties of things in themselves but mere
representations—and have epistemic access to only our mental representations,
from which we infer the existence of an independent reality. However, a problem
with Kant's position is that it is unclear how it fits into the physical world; in
particular, it is unclear how to naturalize Kant's inferential conception of
"binding mental representations" which are percepts that correspond to external
objects, while retaining Kant's conception of the noumena, a concept that
depends on practical reason and seems to register outside the spatio-temporal
domain of transeunt causality.

There are two options to pursue this strategy: weak naturalism and strong
naturalism. The former entails both that everything is objective, physical, or
material, and that everything subjective, mental, or conscious is explicable
according to laws of nature (however we define these laws). The latter involves
the added claim that not only everything subjective, mental, or conscious, but
also that everything ideal, normative, and formal is explicable according to laws
of nature. In naturalizing Kant's noumena, while serving the telos of a physicalist
philosophy of mind, must we altogether abandon the metaphysical and ethical

concept according to which we, as causally determined agents, can conceptualize our thoughts, actions, and beliefs as beholden to self-determination? Furthermore, can we naturalize Kant's schema in the strong sense, such that the coordination of sensations into percepts of particular objects is not understood as a faculty of the transcendental power of imagination but via a model that takes into account the neural correlates of perceptual experience?

II

Consequently, this paper aims to purpose a strong naturalist reading of Kant vis-à-vis, and in service of, the contemporary philosophy of consciousness qua perception. The choice of which naturalist rendering to pursue, without compromising Kant's schema thus becomes relevant. Contemporary philosophers of mind like Frank Jackson may attempt to navigate the problem of explaining how the mind fits into the physical world by underscoring the diaphanousness of act-object structure—objects as we perceive them correspond to and track the external world, where to have an experience is to stand in a relation of awareness to an object with properties that determine the *kind* of experience undergone. Jackson endorses diaphanousness, whereby "accessing the nature of the experience itself is nothing other than accessing the properties of its object" and espouses *strong representationalism*, according to which the nature of experience entails accessing and exhausting the external world. For Jackson, that experience is diaphanous (or transparent) is a thesis about the phenomenology of perceptual experience: the properties that make an experience the kind of experience it is are the properties of the object of experience.[1] On the other hand, a physicalist such as Ned Block may espouse "weak representationism," separating phenomenal consciousness from mental representation and maintaining that while perceptual experiences are (perhaps essentially) representational, they may also have phenomenal or qualitative properties that cannot be reduced to their representational contents. On this second view, conscious experience may involve primitive forms of internally determined phenomenal representation that cannot be defined in functional-

[1] Torin Alter and Sven Walter, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (Oxford, Oxford University Press: 2007), 5.

externalist terms, as certain experiences may have built-in/internally-determined spatial representational contents—Block understand minds in terms of our neural machinery, highlighting how our internal biological makeup (specifically, its electrochemical character) is essential to conscious phenomenology. Block's view offers a successor-concept to Kant's conception that the senses and intellect serve separable albeit related roles, the former playing a passive view in receiving stimulations from the environment and the latter generating novel representations under the executive control of the perceiver in binding representations together. For our purposes, following Kant, it is not a live option to endorse the Cartesian interactionist dualist position taken up by those like David Chalmers who, despite being neutral about various dualist positions, grants epiphenomena causal validity *while* mixing physical and nonphysical events in a single causal chain.[2]

As such contemporary rifts regarding the metaphysics of mind demonstrate,

---

[2] As demonstrated by Jackson's knowledge argument, Jackson first thought there was some factual knowledge about color experience that Mary lacks in the black and white room and that what she lacks does not amount to just a set of recognitional or imaginative abilities. When Jackson came to consider the factual knowledge that Mary possess in the black-and-white room as including facts about what color experiences represents he turned to strong representationism (note: I choose this term, "representationism," and not "representationalism," as I am taking after Ned Block's distinction). Invoking G.E. Moore, Jackson contends that the most phenomenally salient features of perceptual features are its act-object structure and diaphanousness. Jackson's endorsement of the ability argument is via strong representationalism, with this argument's primary contribution being the characterization of phenomenologically salient objects of experiences as intentional objects, rather than instantiated properties that experience acquaints us with; accordingly, all differences in phenomenal character of these experiences can be fully explained as differences in their intentional objects or representational contents. As far as perceiving red is concerned, according to the ability argument, what is distinctive about the experience of red for the first time is that it represents an external property red, rather than some other color property. Janet Levine summarizes Jackson's use of strong representationalism adeptly, remarking that: "if what is distinctive about the experience of red is that it represents the (external) property red (rather than some other color property), and if that experience meets certain further functionally specified conditions that distinguish it from beliefs or thoughts about red, and if the representation relation is some sort of causal-informational relation, then it seems as though everything that's distinctive about the experience of red can be deduced a priori from the physical facts. Moreover, if all the introspectively accessible distinctions among different types of experiences are, as strong representationalists contend, exhausted by their representational contents, then it seems more plausible to hold that the only thing that Mary acquires when she leaves her room is the ability to 'recognize, remember, and imagine the state' in question." See: Janet Levin, "Representational Exhaustion" in *Blockheads!*, eds. D. Stoljar and A. Pautz (Cambridge, MA: MIT Press, 2019), 251; Frank Jackson, "Mind and illusion" in *There's Something about Mary*, eds. P. Ludlow, Y. Nagasawa and D. Stoljar, (Cambridge, MA: MIT Press, 2004), 421–442;

one may be a metaphysical or epistemological realist about mind-dependent entities or properties. More broadly, epistemological realism presupposes metaphysical realism, but not vice versa. For one may claim that although we can minimally stipulate that there is a mind-independent world, we cannot know how this world is constituted. Kant's position is paradigmatic in this regard: he postulates that beyond the fabric of appearances lies the "thing-in-itself" which he terms "the noumenon"—the idea of a mind-independent reality required to theoretically explain what appearances are appearances *of*. Kant thereby distinguishes between our representings of things insofar as they appear mediated by our senses and intellect, and the reality of things conceived outside of any correlation to experience or experiencers. Kant's noumenon is at once a "postulate of reason," introduced to explain the source of our objective representings, and that which is always beyond the aesthetic and intellectual powers which make up such representations. Accordingly, to claim objective knowledge of the in-itself constitutes a transgression to the limits of pure reason, an illegitimate transposition of what belongs to experience into a world beyond it. Insofar as we must infer its existence, the in-itself remains for-us a "thought-object" (*ens rationis*)—a "boundary concept" without an object—corresponding to no empirical intuition and thus having only a negative or regulative use. As Kant says,

In the end, however, we have no insight into the possibility of such 'noumena' and the domain outside of the sphere of appearances is empty (for us), i.e., we have an understanding that extends farther than sensibility problematically, but no intuition, indeed not even the concept of a possible intuition, through which objects outside of the field of sensibility could be given, and about which the understanding could be employed assertorically. The concept of a noumenon is therefore merely a boundary concept, in order to limit the pretension of sensibility, and therefore only of negative use.[3]

Kant's idealist successors would question the motivations behind the "rational necessity" accorded to the noumenon, interrogating the conditions of intelligibility at stake in speaking about a reality that lies beyond the limits of

---

[3] Immanuel Kant, *The Critique of Pure Reason*, trans. Paul Guyer (Cambridge, UK: Cambridge University Press, 1998), A254—255/B310—311

experience and also challenging the skeptical conclusions that followed from the critical method. Paradigmatically, Hegel's speculative idealism attempts to overcome transcendental critique (and its aspiration to delineate the conditions of possible experience) in the name of an "immanent critique" apposite to the ideal of "absolute knowing." In doing so, Hegel makes the difference between how things appear for-us and things as they are in-themselves intrinsic to the way in which consciousness relates to its object. Consequently, Hegel argues, consciousness is itself nothing but the difference between the object as it is in-itself and as it is for-itself. For already in empirical perceiving consciousness implicitly distinguishes between the knowing or concept of an object insofar as it appears for-itself and the truth of the object as it is in-itself. While this difference is made explicit by the phenomenological observer, it is necessarily constitutive of consciousness, for the difference between how things are and how they appear is required to make intelligible how the experience of error, nevermind knowledge, is possible.[4] Hegel's "phenomenological science" takes as its own object the being for-consciousness of the in-itself, making explicit the way which consciousness implicitly distinguishes between knowing and truth, where the latter functions as a conceptually structured normative standard that provides the measure for the former, rather than as an epistemological limit indexing a transcendent and ineffable beyond:

Since consciousness provides itself with its own standard, the investigation will be a comparison of consciousness with its own self; for the distinction just made falls in it. In consciousness, one moment is for an other; in other words, consciousness in general has the determination of the moment of knowledge in it. At the same time, this "other" is to consciousness not only something for it; it is also something outside this relationship or in itself: the moment of truth. Therefore, in what consciousness within its own self designates as the in-itself or the true, we have the standard by which consciousness itself proposes to measure its knowledge.[5]

This way of observing the subject matter is our contribution; it does not exist for the consciousness which we observe. But when viewed in this way the

---

[4] Robert Brandom, *A Spirit of Trust* (Cambridge, MA: Harvard University Press, 2019), 70—72.

[5] G.W.F. Hegel, *Phenomenology of Spirit*, trans. A.V. Miller (Oxford: Oxford University Press, 1978) [section]84.

sequence of experiences constituted by consciousness is raised to the level of a scientific progression […] A moment which is both in-itself and for-us is thereby introduced into the movement of consciousness, a moment which does not present itself for the consciousness engaged in the experience itself. But the content of what we see emerging exists for it, and we comprehend only the formal aspect of what emerges or its pure emerging. For consciousness, what has emerged exists only as an object; for us, it exists at once as movement and becoming.[6]

Initiating an adjacent and revisionary historical trajectory, the neo-Kantian recoil against Fichte, Schelling, and Hegel in the late 19th Century worked to rekindle the scope of the transcendental epistemological investigation into the limits of experience in the wake of new advances in the natural and social sciences, against the perceived metaphysical excesses of speculative idealism. This recasting of the transcendental project would be exacerbated in twentieth century attempts to radicalize rather than to overcome critique—paradigmatically, in the phenomenological and post-phenomenological tradition following from the works of Husserl and Heidegger. Although the analytic tradition was also inspired by a recoil against the perceived rise and influence of Hegelianism—particularly the British Hegelians—its early proponents were likewise hostile to the psychologistic tendencies of neo-Kantianism. As Robert Brandom argues, abjuring their youthful idealism, for Moore and Russell, the "idealist rot" had already set in with Kant, leading then to a rehabilitation of the empiricist tradition, avoiding the anti-realist swamp of the "dangerous oxbow of German Idealism" altogether.[7]

Only in the works of P.F. Strawson, Carnap, and (later) Sellars, would analytic philosophy retrieve the legacy of German Idealism and interrogate the epistemological and metaphysical assumptions guiding its empiricist envelopment, at times even anticipating the arrival of its "Kantian phase." Reforming the constructive tenets of neo-Kantianism in light of the great genealogical and historical critiques of Western philosophy emblemized in the works Marx, Nietzsche, and Freud—the great "masters of suspicion" of the 19th

---

[6] Ibid., [section]87

[7] Robert Brandom, "From German Idealism to American Pragmatism – and Back" in *Kant Und Die Philosophie in Weltbürgerlicher Absicht*, eds. Margit Ruffing et al. (Berlin: De Gruyter, 2013), 107.

Century—blossomed an orientation of thought proposing to examine forms of mediation beyond the aesthetic and cognitive faculties explored in Kant's representational account of subjectivity. This took into consideration the practical, social, and cultural determinations that comprise our collective historical being, progressively dislodging transcendental critique from its "psychologistic" or even phenomenological envelopment. At the same time, as experience was revealed to be underdetermined by ever more inescapable and impersonal forms of mediation, the idea of ineffable or unmediated being was likewise radicalized, progressively leading philosophy to a crisis that exacerbated the skeptical conclusions to which the critical method was led. Thus unspooled novel approaches to naturalizing Kant.

Chiefly drawing from Strawson's and Sellars' respective projects for a "naturalized Kantianism," we here present a constructive rejoinder integrating two vectors—one ethical and the other metaphysical—while presenting a defense of transcendental epistemology as propaedeutic to ontology, reconciling the latter via the prospects of structural naturalism. We begin by tracing Strawson's ethical naturalization of Kant vis-à-vis compatibilist determinism, as articulated in Strawson's oft-quoted paper "Freedom and Resentment."[8] We then move on to Sellars' realist epistemology and its naturalist vim, with an interest in his critique of immediate sensory experiences and perceptual representation, as well as Sellars' epistemic upgrading of propositional inferences.

The latter section of our analysis reinvigorates the naturalized Kantian spirit of Strawson and Sellars in Béatrice Longuenesse's recent work on Kant. Surprisingly little scholarship has been published that situates Longuenesse beyond the fold of a Kantian interlocutor but, as we shall demonstrate, Longuenesse does not merely interpret Kant but offers new metaphysical renderings. Via Longuenesse, we derive a kind of self-consciousness that, while intimately connected to consciousness of one's own body, is nevertheless distinct from it and is, moreover, *the condition* for any use of the first-person "I."[9] *Vide* Freud, Longuenesse also couches Kant's moral "I ought to" and the Categorical

---

[8] P.F. Strawson, "Freedom and Resentment" in *Freedom and Resentment and Other Essays* (Abingdon: Routledge, 1974), 1—28.

[9] Béatrice Longuenesse, *I, Me, Mine: Back to Kant and Back Again* (Oxford: Oxford University Press, 2017).

Imperative's compulsive nature in emotion, rather than reason.[10] For Kant, understanding does not intuit and intuition does not entail thinking. What is crucial to Kant's famous claim that *intuitions without concepts are blind* and *concepts without intuitions are empty* is the idea that we have no knowledge unless we combine concepts with intuition. Although concepts and intuition are radically separated mental powers, without a way of bringing them together (i.e., synthesis) there is no knowledge. Thus Kant's metaphysical-scientific dualism: (scientific) knowledge is limited to the world of phenomena—the world of appearance— while the thing-in-itself is unknowable because there is no intuition that can correspond to the thing-in-itself. Drawing from Oliver Sacks' research on polyneuropathy and Ned Block's distinction between phenomenal- and access-consciousness, we consider Kant's work on intuitions (conscious, objective representations of an object, event, etc.) and the vagaries of a putatively epistemically foreclosed in-itself as applied to questions of consciousness.[11] Threading the philosophy of mind qua metaphysical realism, we ultimately endeavor to conceptualize the limit-conditions regarding our having reportable knowledge about our access to percepts and concepts.

III

Kant's expressed worry in the Third Paralogism of Personality in the First Critique is in developing a concept of personhood that, given an empirically accessible entity conscious of its own identity at different times, is "necessary" *and* "sufficient" for practical use:

> Meanwhile, the concept of personality, just like the concepts of substance and of the simple, can remain (insofar as it is merely transcendental, i.e., a unity of the subject which is otherwise unknown to us, but in whose determinations there is a thoroughgoing connection of apperception), and to this extent this concept is also necessary and sufficient for practical use; but we can never boast of it as an extension of our self-knowledge through pure reason.[12]

Kant's solution of "necessary and sufficient for practical use" is not given

---

[10] Béatrice Longuenesse, *The First Person in Cognition and Morality* (Oxford: Oxford University Press, 2019).

[11] Ned Block, "On a Confusion about a Function of Consciousness," *Behavioral and Brain Sciences* 18, no. (1995): 227–87.

[12] Kant, *CPR*, A365-6.

in the Third Paralogism, but in the solution to the Third Antinomy of Pure Reason:

> By freedom in the cosmological sense, on the contrary, I understand the faculty of beginning a state from itself, 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 [….] Freedom in the practical sense is the independence of the power of choice 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.[13]

Kant concedes that if our existence were only that of spatiotemporal entities belonging to the sensible world, subject to universal causal laws, then there would be no room for metaphysical freedom—that is, for being an uncaused cause, a cause posited *sui generis* that has the power to start a causal chain without being itself determined to do so. Referencing the *Groundwork for Metaphysics*, published one year before the second edition of the *Critique of Pure Reason*, Kant notes in the concluding remarks to the Paralogism of Pure Reason:

> Suppose there subsequently turned up—not in experience but in certain (not merely logical rules but) laws holding firm a priori and concerning our existence—the occasion for presupposing ourselves to be legislative fully a priori in regard to our own existence, and as self-determining in this existence; then this would disclose a spontaneity through which our actuality is determinable without the need of conditions of empirical intuition; and here we would become aware that in the consciousness of our existence something is contained a priori that can serve to determine our existence, which is thoroughly determinable only sensibly, in regard to a certain inner faculty in relation to an intelligible world (obviously one only thought of). (B430–1.)

The notion of person that is defined in the major premise of the Third Paralogism is described by Kant as conceiving of "a being that can be thought of in every respect, and consequently even as it might be given in intuition."[14] There is also a minor premise here: that using the 'I' in 'I think' means expressing consciousness of oneself in unity over time.[15] Thus the major premise of the Third

---

[13] Ibid., A533–4/B561–2; cf A549–58/B577–86.

[14] Ibid., B411-2

[15] "I am an object of inner sense and all time is merely the form of inner sense. Consequently, I relate each and every one of my successive determinations to the numerically identical Self in all time, i.e., in the

Paralogism, dealing with the concept of "personhood," has two features: i) the unity of apperception (the unity of consciousness of contents of mental states, which makes it possible for the 'I think' to accompany them);[16] ii) the numerical identity through time of a spatiotemporal living being who possess mental states.

Kant designates that having a will that is determined under the moral law is what makes us "persons" in the moral sense—i.e., individuals capable of self-determination; "pure and practical reason," in turn, makes us accountable for our actions. For Kant, the notion of person that is "necessary and sufficient" for practical use includes not solely the "psychological" notion (that a person is an entity that is conscious of what is their own identity at different temporal moments), but also a "moral" notion—a person is a conscious being that has a "rational will," a faculty of desire determined under the moral law, which is the Categorical Imperative. Where Hypothetical Imperatives present the practical necessity of a possible action as a means to achieving interests which one desires, Categorical Imperatives index actions that are, of themselves, objectively necessary, without regard to any other end.[17] The necessity for such a noumenal "border-concept" is directly tethered to Kant's recognition that if everything we do were determined according to solely universal causal laws, we would deserve neither reward nor punishment.[18] Kant, drafting compatibilism, seeks to

---

form of the inner intuition of myself. On this basis the personality of the soul must be regarded not as inferred but rather as a completely identical proposition of self-consciousness in time, and that is also the cause of its being valid a priori. For it really says no more than that in the whole time in which I am conscious of myself, I am conscious of this time as belonging to the unity of my Self, and it is all the same whether I say that this whole time is in Me, as an individual unity." Ibid., A362.

[16] The unity of apperception makes it possible for the 'I think' to accompany mental states; the 'I' in 'I think' expresses the consciousness of oneself (that is to say, the thinking of oneself) as remaining identical throughout the holding together of one's representations. See: Longuenesse, *I, Me, Mine*, 146-152.

[17] Immanuel Kant, *Groundwork for the Metaphysics of Morals*, trans. Allen W. Wood (New Haven: Yale University Press, 2002), [section]2

[18] Galen Strawson similarly argues for the impossibility of moral responsibility with the following basic argument: 1) You do what you do in the circumstances in which you find yourself because of the way you are. 2) If you are going to be ultimately held responsible for what you, you ultimately have to be responsible for the way you are. 3) You are not ultimately responsible for the way you are and, therefore, C) you cannot be ultimately held responsible for what you do. More specifically, who we are is the result of genetic determination and early experiences (nature and nurture); since we cannot create ourselves like an uncaused cause (*causa sui*), responsibility is completely out of our control, tethered entirely to luck. Nonetheless, he concedes that the epiphenomenal feeling of free will is something we are locked in to. Galen Strawson, "The Impossibility of Moral Responsibility," *Philosophical Studies* 75, no. 1/2 (1994): 5-24.

demonstrate that the causal determination of every event by temporally antecedent conditions is compatible with the idea of certain events having another kind of cause, belonging to the noumenal sphere and exempt from the temporal condition, thus "acting" freely. Here, Kant confronts the problems introduced by determinism. As it relates to moral freedom and responsibility, moral law—universally legislating through all its maxims—is, for Kant, the *ratio cognoscendi* of freedom: "our consciousness of the moral law we find in ourselves just is our consciousness of our own freedom."[19] However, the concept of person developed by Kant in the Third Paralogism is not sufficient for practical use; it is necessary but not a sufficient condition for the practical use of reason because while it engages with the psychological facet—that a person is an entity that is conscious of their identity at different times as the entity they are—it does not deal with the moral aspect (i.e., the "rational will"). It would appear that, in the first *Critique*, Kant falls prey to his own paralogism, as no objective determination of myself as a thinking and willing entity or a claim to be an uncaused cause can be determined from the use of 'I' in "I ought to."[20]

There is, indeed, a severing that Kant exacts between nature and freedom, insofar as the noumena is linked to pure practical reason. A fundamental difficulty is brought to bear by the relation between moral law, which depends on the faculty of reason alone, and actions that unfold in the sensible world and are causally necessitated; this speaks, more broadly, to the distinction in The Architectonic of Pure Reason of the duality of two distinct concepts of philosophy for Kant: *Schulbegriff* (the scholastic concept) and *Weltbegriff* (the cosmological concept). The former deals with the object of science and technical study, and the latter the cosmological or universal concept, which is concerned with the relation of all knowledge to the essential aims of human reason. A notion of a person, in the moral sense, as an entity belonging in the empirical world does transpire in Kant's posterior work, where consciousness of being a person is deemed equipollent to consciousness of unconditional maxims. For Kant, metaphysical freedom is linked to the moral "I ought to" which is a byproduct of being conscious of one's own self-determination, a first-person standpoint on

---

[19] Longuenesse, *I, Me, Mine*, 153.

[20] Ibid.,

oneself that offers no insight into any kind of objective causal determination, i.e., a claim to be an uncaused cause.

In the *Critique of Practical Reason* Kant further develops the position that he began to develop in the Third Antinomy of the First Critique; where the latter allotted a notion of obligation via "transcendental freedom," in the Second Critique, Kant argues from the fact of moral obligation to the noumenal reality of moral freedom.[21] For Kant, the Categorical Imperative's normative grip, in being linked to reason itself, has an innate psychological obligation—*it just is* and, therefore, we feel its compulsion. Elaborating on this putative objective necessity, and the link between reason, freedom, and moral responsibility, is the onus for any philosopher naturalizing Kantian freedom. A second-order demand is explaining, then, why and how a society's moral system can be reformed if Categorical Imperatives are unconditional and an extension of the in-itself. Kant falls into the same fallacious inference he accuses Descartes and rationalist metaphysicians of in his first *Critique*, because he sets up a necessary condition for understanding oneself as morally accountable for one's actions, but this does not belong to an objective standpoint re: causal determination but to the subjective and universally stipulated standpoint about agency of actions; thus, Kant participates in an equivocation, his middle term being "self-determining."[22]

Kant, in linking these two problems by way of the unity of mental states, provides a path for the philosophy of mind to keep in mind two registers that inform one another—i) self-legislating action and determined causality; ii) phenomenal consciousness and the binding of representations. The Kantian, or neo-Kantian, doctrine of "two standpoints" compatibilism proffers a position where the compatibilist thinks that facts about the physical world and facts about human action occupy different realms, "standpoints," or conceptual schemes, and, therefore, cannot be brought into conflict with one another. Correspondingly, claims that we are free and that the world is deterministic cannot be made to contradict. Strawson's endeavor in "Freedom and Resentment" begins here. Strawson culls Kant by naturalizing moral freedom as independent from the contingencies (e.g., history or circumstances) that the truth

---

[21] Immanuel Kant, *Critique of Practical Reason*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1997), [section]5/30.

[22] Longuenesse, *I, Me, Mine*, 158.

of determinism might imply.

Evoking Kant's notion of the "rational will," Strawson offers that we have a natural, nonrational (not *ir*-rational) commitment to engaging in characteristically interpersonal relationships. Observing our actual practices, Strawson's descriptive metaphysics is directed at the integration of a complex of conceptual resources within our commonsense realistic conception of experience; thus metaphysical freedom involves an epistemological confrontation with coherence theories of justification. The quality of others' wills towards us matters to us and are manifest in their behavior such that we put some set of demands on the quality of others' wills; accordingly, we will react in certain ways when those demands are violated. Participant reactive attitudes include "gratitude, resentment, forgiveness, love, and hurt feelings."[23] Strawson also speaks of what he calls an "objective attitude,"

> The objective attitude is not only something we naturally tend to fall into in cases [of] abnormalities or immaturity. It is also something which is available as a resource in other cases. The objective attitude is not only something we naturally tend to fall into in cases [of] abnormalities or immaturity. It is also something which is available as a resource in other cases, too . . . we can sometimes look with something like the same [objective] eye on the behavior of the normal and mature. We have this resource and can sometimes use it—as a refuge, say, from the strains of involvement; or as an aid to policy; or simply out of intellectual curiosity.[24]

According to Strawson, the "objective attitude" extends the logic of *exemption* from moral demands. The distinction between *exemption and excuse* is critical to Strawson's picture of responsibility: in addition to our personal reactive attitudes are impersonal or vicarious analogues. Cases of *excuse* are those in which we suspend or modify reactive attitudes due to error about the quality of the will; we might include an agent who is innocently ignorant, acting in accident, or acting unintentionally. In cases of *excuse*, we exculpate the injury in question but do not question the quality of the person's will; the usual moral demands stay in place— this one act is excused. However, there are also cases of *exemption*. We encounter this in the case of small children, people suffering from dementia, or those with forms of other serious mental illness. We might view such people as equipped

---

[23] Strawson, "Freedom and Resentment," 5.

[24] Ibid., 10.

with mental attributes such that we are disposed not to indulge in our reactive attitudes of resentment, approbation, etc., since they are incapable of ordinary interpersonal relating. Thus, in cases of *exemption*, rather than react with the corresponding reactive attitudes, we do not take a moral view of others, adopting the *objective attitude* and exempting them from the usual demands.

As mentioned above, the *objective attitude* can also work as a "resource," as we sometimes shift from reactive to objective attitude even in cases in which the will is neither "immature, diseased, nor in extreme or unusual circumstances."[25] These are instrumental scenarios of emotional effort, as demonstrated in scenarios of emotive disengagement from the stresses of involvement. Considering that we *can* adopt this "resource," the skeptic, an incompatibilist about freedom and responsibility, asks could or should the acceptance of the determinist thesis lead us always to look on everyone exclusively this way, where the acceptance of determinism entails the decay or repudiation of participant reactive attitudes?[26] For the unconvinced pessimist, a metaphysics of morals could enumerate an argument that starts from claims about the nature of moral requirement or moral demand and reach the conclusion that moral demands require a form of control, possibility, originality, or spontaneity, which is ruled out by the truth of determinism. Thus, it would follow that we would universalize the *objective attitude*, applying the "resource" indiscriminately. Constructing a response, Strawson considers two naturalizing trajectories.

The first is the (simple) Humean interpretation. This rendering follows from Strawson's analysis that, because determinism is an entirely general thesis true of everyone at all times, its truth would not show that we are not morally responsible since we are not capable of abandoning our reactive attitudes given our psychological limitations. Thus, we are ineluctably freighted with treating each other as if we are morally responsible. Strawson writes,

> According to Hume the naturalist, skeptical doubts are not to be met by argument. They are simply to be neglected…. because they are idle; powerless against the force of nature, of our naturally implanted disposition to belief. This does not mean that Reason has no part to play in relation to our beliefs concerning matters of fact

[25] Pamela Hieronymi, *Freedom, Resentment, and the Metaphysics of Morals* (Princeton: Princeton University Press, 2020), 10.

[26] Strawson, "Freedom and Resentment," 12.

and existence. It has a part to play, though a subordinate one.[27]

Just as Hume, appealing to Nature, posits two commitments—the existence of body and the general reliability of inductive belief formation—Strawson's Humean interpretation conceives of our moral framework as a "given." Recall Hume's refrain that our beliefs in the general reliability of induction are not grounded beliefs and, simultaneously, not open to serious doubt—rather, they are outside our critical and rational competence because they define the area in which that competence is exercised. Similarly, once we are "inside" a moral framework, questions of justification have free rein. Granting the Humean claim that (only) certain "framework" commitments are unconditionally immune to questions of justification, we can challenge *which* commitments belong to the framework and why those, and only those, need no justification. This leads to the unsatisfactory conclusion that if someone lacks the ability to do or change something, then their continuing in the status quo neither calls for nor permits justification.[28] Thus Strawson's preferred rendering, the broadly Wittgensteinian interpretation:

> Wittgenstein distinguishes between those matters—those propositions—which are up for question and decision in the light of reason and experience and those which are not, which are, as he puts it, 'exempt from doubt.' …We do not, for example, find in Wittgenstein any explicit repetition of Hume's quite explicit appeal to Nature….Above all, there is, in Wittgenstein's work…the distinction between 'what it is vain' to make a matter of inquiry, what 'we must take for granted in all our reasonings,' as Hume puts it, on the one hand, and what is genuinely matter for inquiry on the other.[29]

This reading proffers that Strawson forwards a *conceptual* line of thought wherefore we can neither support nor call into question the *whole* of a practice using notions that are, themselves, constituted by this practice. We cannot ask whether our moral practices themselves are appropriate, right, or fair, just as we cannot sensibly call the game of baseball, itself taken as a whole, "fair" or "foul." According to the broadly Wittgensteinian thought, it makes no sense to question or to criticize a practice as a whole in terms that owe their meaning to that

---

[27] P.F. Strawson, *Skepticism and Naturalism: Some Varieties* (Abingdon: Routledge, 1987), 14—15.

[28] Hieronymi, *Freedom*, 81.

[29] Strawson, *Skepticism and Naturalism*, 15—16.

practice.[30] For Wittgenstein (specifically, the Wittgenstein of *On Certainty*) society, as nature, establishes the framework. What is critical about Strawson's displacement of nature for society is that the latter implies a social process that occurs by way of "our learning, from childhood up, an activity, a practice, a social practice—of making judgements, of forming beliefs."[31] The crypto-propositions within this social framework reflect the general character of the practice itself, forming a frame within which the judgments we actually make hang together in a coherent modality due to the competence of a community of social agents.

Strawson arrives at a position he calls "social naturalism," whereby the reasoning behind our commitments need no justification in the face of determinism because they are linked to both our natural psychological limitations and the role the commitments play.[32] Furthermore, since it is true of everyone, we can know in advance that being determined is not a reason to exempt and the objective attitude. Unlike Kant, Strawson does not think that our moral standards, themselves, are intractable. Moral reason is diachronic and, in turn, our reactive attitudes develop and can be contested via reflection. Reflection allots our standards of regard progressive development and, more broadly, proffers a criteria for re-identifying the concepts of objects, allowing for amendments within the objective world; Strawson's scheme is subject to the limitation that change must be consistent with the possibility of applying those concepts and criteria in experience.[33] The unconditional nature of Kant's moral "ought" is reconciled by Strawson's metaphysics of morals—despite standards of regard do not exist in space and time, the status of ideas become sovereign-as-process, with time rendered sensible as we become conscious of the in-itself thanks to the incompleteness to our ethical system.

For Strawson, participating in our ordinary reactive attitudes means that we do not simply take an external evaluative view of ourselves and others, instead we judge the participants *themselves* rather than their desirability or utility;

---

[30] Ibid., 6.

[31] Ibid., 20.

[32] The "acceptance of the existence of body and of the general reliability of inductive belief-formation [….] is the groundwork; and its source is unambiguously identified. These unavoidable natural convictions, commitments, or prejudices are ineradicably implanted in our minds by Nature." Ibid., 19.

[33] P.F. Strawson, *Bounds of Sense* (Abingdon: UK, 1975), 90.

however, as Thomas Nagel remarks, try as we might, the "external view" impels itself upon us while we resist it: "[o]ne way this occurs is through the gradual erosion of what we do by the subtraction of what happens."[34] Reflection, poised within mankind's scientific project, helps us better understand the terms of our relating—it reveals to us the principles *actually* at work. It is not that we should conclude that we have been in error all along about the nature of our freedom when we are tasked with revising our reasons for exempting in the face of novel scientific/ethical discoveries. Rather, via rational reflection, we can refine standards, improve practices, and supply rules. Kant did not offer a fluid picture of interpersonal relating, and Strawson remedies this as discoveries illuminate new facts about our ways of relating—if we discover an inconsistency, we have discovered a point at which reason qua reflection has revealed that we are wrong about the principles at work in our practice. Moral principles, construed along this naturalist axis, are of our own making.

In *Individuals* Strawson recalls Kant's argumentation about the mutuality of self and world, concluding that any apperceptive discursive intelligence is engaged in representing a world and, thus, necessarily represents the contents of that world "categorically"—that is, as systematically interrelated and as existing independently of their being represented. In *The Bounds of Sense*, Strawson carves a sceptic ones again, this time challenging us to reconstruct a public world of physical objects and events via nothing but private sense-data to evince justified belief in the objective world by working outwards—in contradistinction to concepts of sense-data, concepts of objects are necessarily compendia of causal law. Strawson's objective realism returns to the question of how limitations between order and regularity must be reflected "in the character of our concepts themselves."[35] "Freedom and Resentment" applies the logic of Strawson's *constitutive realism* to a social scheme. More broadly, Strawson's naturalized Kantianism offers that, although our independence from the physical world does not appear when we consider ourselves as merely empirical subjects, it is nonetheless true of us as we are *in ourselves*.

Just as identification occupies the evaluative fulcrum for Strawson's

---

[34] Thomas Nagel, *Mortal Questions* (Cambridge, UK: Cambridge University Press, 1979), 38.

[35] Strawson, *Bounds of Sense*, 90—91.

progressive picture of morality, his causal picture of objective reference offers a theory of differentiating physical particulars via contextual, demonstrative relations. We will see how Gareth Evans eventually draws from this interpretation of Kant, particularly Strawson's linking perceptual representation of physical entities in conceptualizing objective realism. As Tyler Burge notes, however, this also opens the door for considerable conceptual error, reducing the problem of explaining "minimum conditions on experience of objective reality to the problem of explaining necessary conditions on our conception of the relation between perceptions and their objects an aspect of the problem of explaining our conception of objectivity."[36]

IV

Wilfrid Sellars's reality theory of knowledge, imbued with a process monist ontology, proffers a notion of judgment via an inferentialist framework. Like Strawson, Sellars emphasizes sensory re-afference, where reasoning does not involve solely the innate structures in the mind but a space of dialogical interactions—a logical "space of reasons" of justifying or being able to justify one's stance. Sellars' "naturalized Kantianism" avows the irreducibility of epistemological explanation, while reconciling it with prospects for a genuine materialism. Such is the basis of Sellars' transcendental materialism, which distinguishes the epistemic and ontological dimensions involved in accounting for the cognitive and sensory mediations between subject and world. Sellars simultaneously resists the correlationist delimitation of knowledge to appearances and absolutizing the sensible as much as the Platonist solution of reifying the direct ontological purchase of formal vocabularies against natural languages and the endowments of the sensible. Sellars' project thus endeavors to define a hierarchy of cognition within a representational and realist register, accounting for how the intelligible dimension of thought in its conceptual and structural dimensions abstracts itself from intuition. This means explaining how the intelligible develops from the sensible faculties as part of an informational dynamics through which nature is progressively disclosed to thinking. The representational relation between thinking and being is neither to be understood in terms of a semantic relationship between conceptual states nor in terms of a

---

[36] Tyler Burge, *Origins of Objectivity* (Oxford: Oxford University Press, 2010), 161.

qualitative resemblance between propositions and the world. Rather, Sellars argues, the logical order structurally constitutes a nomological double or map of its environment, such that an isomorphism or homomorphism can be drawn between the expressions of a language or theory and the basic constituents of the world as progressively represented by empirical science. This entails distinguishing between proprieties in the conceptual order and those of the propositional order, without thereby falling into a dualist picture that renders the relation between the two unintelligible.

Denying a representationalist theory of the relation between thinking and being is not to deny that representation is a phenomenon to explain. Sellars helps us conceive of the possibility of avoiding the metaphysical excess and skeptical consequences of Kant's transcendental project without absolutizing the conditions of experience into the mechanisms of nature as a whole, nor relinquishing the project of positive epistemology or ontology as a whole: to be worthy of a belief, individual perceptual beliefs hold warrant only insofar as they infer further beliefs about sensory experience and a sign's reliability. Sellars' reconstruction of Kant via Peirce involves the resuscitation of the concept of noumena within a pragmatic register, refashioning noumenal reality as a regulative ideal which orients the historicity of science as a self-correcting enterprise, but which allow us to assess how, diachronically speaking, theories converge toward a specific *limit*. This project not only re-orients the formal sciences but engages in a dialectic by virtue of which we arrive at a self-conception of ourselves and nature—one that includes how we try to explain our place in nature. As Strawson notes, while notions such as "universal, concept, type" exist solely as objects of thought, "thought, thinking, is certainly a natural process, something that takes place in nature, has causes and effects."[37] Like Strawson, Sellars prods Kant's noumena into experience, with science progressively getting closer to the mapping of a mind-independent reality—a regulative ideal in relation to which we can put ideas in concession. Thus, "[t]he conceptual distinction between what is, and what isn't but yet is possible, is relevant to and relative to each concrete system."[38]

---

[37] Strawson, *Skepticism and Naturalism*, 90.

[38] Wilfrid Sellars, "Concepts as Involving Laws and Inconceivable without Them," *Philosophy of Science* 15 (1948): 287–315.

One of Sellars' most cited contributions to the analysis of the "Science of Structures" is his critique of the Myth of the Given.[39] The "given" is the immediate presentation of sense-data, is neither a product of inference nor presupposing any other knowledge while offering indefeasible empirical information the physical environment. Sellars is adamant that the privileged status of "any one history" must "be a status which stems from outside the Conceptual Realm and which consequently cannot be penetrated by the a priori Science of Structures"[40]—within our "Conceptual Realm" no concrete system or family of concrete systems has a privileged position. Thus, Sellars' critique entails that there is no "pre-theoretical" foundation of knowledge waiting for us once we have "bracketed" naturalist or metaphysical commitments. However, Sellars goes further than this anti-foundationalist gesture, arguing that the vocabulary of subjective experience is a conceptual achievement, and one which is epistemologically dependent upon our capacity to describe the world and its constituents in an objective mode. The "manifest image of man in the world" as an intentional agent that mediates its overt behavior by way of experiential episodes thereby constitutes a development over what Sellars names the "original image," in which man as much as nature were conceived in broadly animistic and behaviorist terms, while lacking psychological or phenomenological vocabulary.[41] The "manifest image of man in the world" is a byproduct of postulating inner episodes as the mediating causes of overt action that we use to explain human behavior. These psychological, internal episodes are *introspectible* and remain *observable* to the agents that bear them. Acts of introspection are thus understood analogically to *extrospectible* perceptual acts.

The emergence of the manifest image which inaugurates the "framework of persons" is coincident with a "de-personalization" of the rest of the world.[42] In this regard, Sellars argues, "the manifest image" is already a kind of "scientific image," insofar as it proceeds by theoretically introducing new kinds of entities:

---

[39] Wilfrid Sellars, *Empiricism and the Philosophy of Mind*, (Cambridge, MA: Harvard University Press, 1997).

[40] Ibid., 314.

[41] Wilfrid Sellars, "Philosophy and the Scientific Image of Man" in *In the Space of Reasons*, eds. Kevin Scharp and Robert Brandom (Cambridge, MA: Harvard University Press, 2007), 374.

[42] Ibid., 378.

internal episodes. The manifest image seeks to explain behavior qua mind by attending to the relations between two different kinds of observable states of affairs: extrospectible physical events and introspectable psychological episodes. In contrast, the emergence of the scientific image depends on the postulation of unobservables as causes for the relations between observables. The arrival of the scientific image proper from the manifest image involves an additional step, flattening man onto nature by depersonalizing man itself, assimilating operations to an objective framework of physical-causal explanations that dispense with the "folk-psychology" of intentions altogether. While the scientific image emerges from the concepts elaborated within the manifest image, Sellars insists, it also presents itself as a rival image, such that to the extent that the manifest image does not survive, "man himself" will not survive.[43]

In reconciling this transcendental naturalist and realist project with a historical conception of scientific cognition, the distinction between the epistemological and metaphysical orders of explanation also reveals a dialectical intrication. For the scientific and manifest images are diachronically coordinated, such that the former appears as an elaboration of the latter:

> [S]cience is continuous with common sense, and the ways in which the scientist seeks to explain empirical phenomena are refinements of the ways in which plain men, however crudely and schematically, have attempted to understand their environment and their fellow men since the dawn of intelligence.[44]

Sellars' progressive self-correction—which contemporaries of Sellars such as Brandom and McDowell have tethered to the Hegelian project of diachronic practices and concepts, respectively—is analogous to that which Strawson similarly took up with his consideration of compatibilist determinism (re: freedom and morality) qua interpersonal relating. This tradition emphasizes assessing reasons for actions and thought by naturalizing the relation and reasoning between self-consciousness and an object-referent. Brandom preserves some basic, non-modal, purely sensory-perceptual notion of observation and McDowell, even when acknowledging that our perceptual capacities are fallible, insists that this generic fallibility is compatible with distributive infallibility

---

[43] Ibid., 386.

[44] Wilfrid Sellars, *EPM*, 183.

regarding non-defective perceptual episodes.[45]

The view that reasons which support knowledge must be propositional, wherein reason is fundamentally a faculty of a priori principles, recalls Kant's Transcendental Deduction. Kant's strategy was to argue that, in postulating that synthetic a priori judgements are possible, we offer the claim that empirical objects fall under pure concepts of the understanding, which is justified because objects of experience would not be presented as objects in the first place unless the manifold of perceptual input from which we derive our representations of objects in the world had been ordered according to a priori concepts of the understanding. It is a priori true that those concepts are veridical for objects of experience; in any given case, we have to rely on empirical investigation to determine which empirical objects fall under which categories. Via the Transcendental Deduction, we know a priori that all empirically given objects fall under one or the other of each "title" of the categories (quantity, quality, relation, modality).

Reflection can help us better understand the terms of our relating, revealing to us the principles actually at work. Tasked with revising our reasons for exempting in the face of novel scientific/ethical discoveries we, via rational reflection, can refine standards, improve practices, and supply rules. Sellars typologizes the different kinds of transitions invoked by cognition in a robust manner, so as to account, for example, for which kinds of language-entry transitions obtain besides perceptual responses to sensory signals and direct somatic stimulation. Many of these capacities involve the kinds of capacities associated with memory and the imagination already identified by Kant. Just as it possible to abstract the functional role of a given concept from its specific syntactic-linguistic expression, we can abstract the generic functional capacities involved in cognition from their modes of implementation. This leads to a more nuanced way to understand the "conditions of experience" that Kant associates with our aesthetic and intellectual faculties as part of a natural and cultural history, complicating the relationship between the empirical and the transcendental. For while the generic function of "the space of reasons" is defined

---

[45] John McDowell, "A Sellarsian Blind Spot" in *Sellars and his Legacy*, ed. James O'Shea (Oxford: Oxford University Press, 2016), 100–116.

a priori, with invariant capacities across all material mediums, languages, and cultures, it is also the product of contingent selections in natural history, subject also to indefinitely varied cultural-historical reconfigurations:

> The idea that this logical space is an evolutionary development, culturally inherited, is an adaptation rather than a rejection of Kant's contention that the forms of experience are a priori and innate … his conception of the forms of experience was too narrow, and that non-formal patterns of [material] inference are as essential to the conceptual order as are the patterns explored by formal logic.[46]

Just like the concepts, vocabularies, and theories that furnish the mind are to be historicized, so the "forms of intuition" that furnish the bedrock of our sensory receptivity must be understood in a dual sense. For while in one sense these designate a functional set of invariant capacities by virtue of which a system represents its environment, in another sense they comprise a set of system-specific modalities which, while invariant across a species or system's classes, are nevertheless the products of natural evolutionary history. "Transcendental" determinations cannot simply be reduced to empirical or contingent-historical ones; in turn, we must separate two levels of transcendental determination pertaining to the forms of experience: (1) *Hard transcendental constraints* pertaining to the invariant and necessary functional capacities required as a priori conditions for cognition, irrespective of its forms of implementation in specific material-linguistic mediums. (2) *Soft transcendental constraints* pertaining to the invariant implementation of these generic functional capacities in a given class of cognitive systems that are contingent products of natural-cultural history, and which may be subject to variation, revision and intervention.

For Sellars, perception is defeasible, as reasons are construed as necessarily propositional and perception is non-propositional—it plays no normative role in knowledge. Thus, all epistemic warrant derives through propositional inferences from propositional attitudes. Causally prompted by non-conceptual inputs, perception is a linguistically mediated capacity for Sellars, in which such inputs are incorporated into a discursive economy, becoming liable to inferential uses. To say of any episode that it counts as an instance of knowledge is not to describe it in an empirical mode or to index its causal origin, but to place it in a

---

[46] Wilfrid Sellars, "Phenomenalism" in *Science, Perception and Reality* (Atascadero: Ridgeview, 1963), 90.

justificatory network of implications, by virtue of which a cognitive system enters the logical "space of reasons" of "justifying and being able to justify what one says."[47] Presupposing the receptivity of sensation via non-linguistic inputs as a necessary condition, to perceive something "as something" involves binding labeling responses to sensory stimuli, producing counterfactually robust inferential relations of incompatibility and consequence:

> For while one does not have the concept of red until one has directly perceived something as red, the coming to see something as red is the culmination of a complicated process which is the slow building up of a multi-dimensional pattern of linguistic responses (by verbal expressions to things, by verbal expressions to verbal expressions, by meta-linguistic expressions to object-language expressions, etc.) the fruition of which as conceptual occurs when all these dimensions come into play in such direct perceptions as that this physical object (not that one) over here (not over there) is (rather than was) red (not orange, yellow, etc.) [...] the direct perception of physical objects is mediated by the occurrence of sense impressions which latter are, in themselves, thoroughly non-cognitive [...] this mediation is causal rather than epistemic.[48]

For Sellars, just because perceptual reports are non-inferentially prompted in relation to non-conceptual sensory inputs, they are not exempt from epistemic challenge. Indeed, as Sellars shows, we cannot via mere sensory inspection ascertain the causal structure of any material particular, nor of any of its manifest features. However, this does not entail that its manifest features are merely occurrent rather than dispositional causal characteristics.[49] Observational concepts which are used to make non-inferential perceptual reports are every bit as liable to assessment and revision as any theoretical concepts. While observational concepts "ground" empirical knowledge insofar as they articulate conceptual responses to environmental data coming from sensation, this does not mean that those concepts are "given" in the sense of foundational items that are epistemically dependent or apodictic. The difference between the observational and the theoretical is methodological rather than ontological: concepts that only have theoretical, inferential uses may eventually acquire observational ones,

---

[47] Sellars, *EPM*, 169.

[48] Wilfrid Sellars, "Phenomenalism," 90—91.

[49] Kenneth Westphal, Kenneth Westphal, *Kant's Critical Epistemology: Why Epistemology Must Consider Judgment First* (Abingdon: UK, 2020), 222.

without implying that what the concepts designate or describe changes.

> [T]o reject the myth of the given is not to commit oneself to the idea that empirical
> knowledge as it is now constituted has no rock bottom level of observation
> predicates proper. It is to commit oneself rather to the idea that even if it does have
> a rock bottom level, it is still in principle replaceable by another conceptual
> framework in which these predicates do not, strictly speaking, occur.[50]

Tracing the porosity between the observable and the unobservable is crucial to understanding how conceptual construction guides the historical development of Sellars' so-called "scientific image," and its development from the "manifest image." As Sellars put it, the behavior of representational systems must exhibit the functional structure of what he calls "Humean" orientation systems, which associate their representings in accordance with logical rules for conjunction (co-presence), disjunction (negation), and quantification (measurement).[51] This navigational scheme fulfills the general functions of reference and characterization required for individuating represented items.

In its functional and pragmatic basis, Sellars accounts for representation as more primitive than consciousness or intentionality, even if representing is augmented by the acquisition of sensing and language. Sellars indeed provides us with a meaningful way to understand a sapient system's discursively mediated use of theoretical reasonings to describe their worlds and themselves. However, we reject Sellars' claim that reasons are the only source of epistemic warrant and that perceptual beliefs without reasons cannot be epistemically warranted—warrants for perceptual belief consist partially in an individual's being in perceptual states which reliably figure in the formation of true beliefs, even if they are not reason-giving:

> An animal with visual perceptions of, and as of, bodies cannot 'make sense' of the
> reference. It need not have reasons. It need not know any criterion for being a body:
> it is enough that the animal be able to discriminate and track bodies by visual

---

[50] Wilfrid Sellars, "Scientific Realism or Irenic Instrumentalism: A Critique of Nagel and Feyerabend on Theoretical Explanation" in *Boston Studies in the Philosophy of Science Vol. II*, eds. Robert Cohen and Max Wartofsky (Springer: New York, 1965): 353.

[51] Wilfrid Sellars, "Mental Events," *An International Journal for Philosophy in the Analytic Tradition* 39, No. 4 (1981): 325—345.

perception.[52]

Animals, children, and many adults lack reasons for their perceptual beliefs but are often warranted in having them epistemically entitled to them. To be in a contentful state just is to be in a state of a certain specific representational kind. Perceptual representational content sets veridicality conditions or, more specifically, accuracy conditions—"accuracy" distinguishes perception's conditions from truth conditions, which are propositional. The fundamental point about perceptual states is that they, and hence their representational contents, can be accurate or inaccurate, within some range of approximation. "Egocentric indexes," representing particulars that are not perceived, mark the present time, the origin or anchor of a spatial framework, or the believer—thus a believer's referring depends on the structural framework of the perception or perceptual belief.[53] Consequently, epistemological warrants ought not be confined to reasons—perceptual content grounds much propositional content.

With regard to Sellars' inferentialist explanation of discursive intentionality, his stripe of functionalism concerns itself with the role that mental states play as part of a system of relations. Sellars' essential break with Kant comes in noticing that the introduction of such theoretical postulates need not imply existential commitment to an ontologically distinct domain of entities or "things" in relation to the observational domain, which would split noumenal reality from phenomenal appearing. Sellars deals with attitude-ascriptions and how they bolster the ability to explain the conditions for the realization or failure of epistemic and volitional acts that bind agents to their environments, anticipating the postulation of theoretical unobservables as the causes of observables. Similarly, Kant's postulation of the noumenal "thing-in-itself" begins by introducing a theoretical "thought-entity"—knowledge which is only inferentially licensed, on whose basis one explains the origin of observable phenomena. The naturalized noumena uncloaks itself in influencing our reasoning: diachronic in both Strawson's conception of ordinary reactive attitudes and Sellars' conceptual licensing. However, as mentioned above, Sellars makes a

---

[52] Tyler Burge, "Psychological Content and Egocentric Indexes" in *Blockheads!*, eds. Adam Pautz and Daniel Stoljar (Cambridge, MA: MIT Press, 2019), 46.

[53] Ibid., 71-79.

grave error, conflating sensory registration with perception and, therefore, hastily dismissing the epistemological warrant of non-propositional representation. Sellars fallaciously presumes that representation depends on the knowledge of conditions for representation. Longuenesse remedies Sellars' erring by conceding epistemic warrant to the perceptual input(s) from which we derive our representations of objects while, simultaneously, prioritizing the more fundamental role of ordering- and binding-representations according to a priori conditions via the synthesis of apperception—the synthesis that yields experience.[54]

V

Reading Kant, Béatrice Longuenesse's metaphysical stipulation is that consciousness in the rational unity of our thinking is *more fundamental* than consciousness of our proprioceptive body—being attentive to the rational unity of content(s) in one's thinking is what makes it possible to assess the standpoints from which we initially formulate, and then arrive at, shared universal conclusions. Two dichotomies transpire: *singular/universal* and *bodily/rational*. What is radically individual in what we assert of ourselves is what is true of us as an entity individuated in space and time—existence as a material organism. However, with the apperceptive "I think" and the moral "I ought to" we assert the exercise of capacities that, by principle, we share universally. It is not that when we use the first-person "I" it is not indexical, for it still refers to an individuated entity: with "I think," if I am correct to say that there are "users" of "I," such that "I" am the thinker of "I think," then "I" is still individual. Yet I am also asserting something conceptually universal—for Strawson, this universality is tethered to our ordinary inter-personal relationships and reactive attitudes but, for Longuenesse, it will be expounded through the developmental and metapsychological structure undergirding the moral "I ought to."

For Kant, there is a fundamental difference between the self-consciousness proper to the thinking subject in the course of their thinking and the consciousness of themselves as an embodied object in the world. It follows that what Kant means by *consciousness of oneself as a thinking subject* is not and cannot, as Quassim Cassam claims, be reduced to consciousness of oneself as a physical

---

[54] Kant, *CPR*, A176/B218–A181/B224.

entity, despite it *is* intimately connected with the consciousness one has of one's own body (viz. proprioception).[55] In the Strawsonian tradition, Kant's claim that the self can be known as object, but not as subject, has been criticized for neglecting the essential role of the body/embodiment. Strawson provided individual representationalist arguments for thinking that perception constitutively depends for representing physical reality on propositional thought—that objective empirical representation of physical, environmental particulars cannot stand on its own, among an individual's representations, and must be derived from and embedded in other sorts of representations available in the individual's psychology. Contra Sellars, Strawson saw perceptual demonstrative reference as based on perception of physical objects, not sense data. Just as Cassam argues that awareness of oneself as a physical object is necessary to ground self-consciousness, Evans, following Strawson, proposes that the body is necessary to ground any referential use of "I," including the self-ascription of mental states.[56]

Drawing from Wittgenstein's *Blue Book* and Sydney Shoemaker's reinterpretation, this tradition involves bifurcating two uses of "I": "I" as object and of "'I" as subject."[57] Following Wittgenstein, all uses of "I" as object depend on identity statements (e.g., "I have a tattoo on my forearm"). Here, knowing the predicate to be true of someone is not, *ipso facto*, knowing it to be true of oneself, the current believer of the thought or speaker of the corresponding sentence, because an intermediate premise that consists in a statement of identity must be justified on objective grounds (i.e., "The person with the tattoo on their forearm = I"). With judgments in which "I" is used as subject (e.g., "I think"), Wittgenstein posits the self-ascription of psychological predicates. Following this use of "I," no recognitional capacity, intermediate premise, or criterion of identification is in play in order to determine whom the predicate is true of, unlike "I" as object.

According to Shoemaker's amendment, judgments in which "I" is used as subject are characterized not by the fact that they are not about a particular embodied person but by the fact that they are immune to error through

---

[55] Quassim Cassam, *Self and World* (Oxford: Oxford University Press, 1997).

[56] Gareth Evans, *The Varieties of Reference* (Oxford: Clarendon Press, 1982).

[57] Ludwig Wittgenstein, *The Blue and Brown Books* (Oxford: Blackwell, 1958); Sydney Shoemaker, *The First Person Perspective and Other Essays* (Cambridge, UK: Cambridge University Press, 1996).

misidentification relative to the first-person pronoun. This involves inside-knowledge: knowledge of predicates one asserts to be true of oneself on the basis of a kind of information not available as the basis for knowledge of others' states.[58] Perceiving by way of thinking is here *a state of awareness*—i.e., epiphenomenal qualia, the "what-it's-like-for-the-subject-of-thinking."[59] Evans also upholds that *the body* is necessary to ground any referential use of "I," including the self-ascription of mental states—"I-thoughts," or self-consciousness about oneself as the subject of thought and action, require the self-ascription of some of our bodily states to be so immune. Evans' interpretation propounds the Fundamental Reference Rule, according to which the essence of self-consciousness is self-reference: "'I' is a word or concept that refers […] to the author of the thought or the speaker of the sentence in which 'I' is being used."[60] G.E.M. Anscombe's work on reflective self-consciousness defends the view that "I" never has referential use. For Anscombe, "I" is neither a name nor another kind of expression whose logical rule is to make reference at all.[61] That is, "I" is a non-referential term that predicates the reflexivity of the activity of thinking—espousing Anscombe's position, "I"-thoughts are not thoughts *about* a particular object named by "I" or to which "I' refers; rather, they are unmediated examples of reflective consciousness (of states, actions, movements, etc.) of the body.

Burge shows, contra Evans and Strawson, that genuine spatial content and singular reference via the first-person "I" are distinct. Egocentric indexes occur in perception but first-person concepts, which are constitutively associated with propositional thought, do not. Evans accounts for the proper epistemic requirements in using "I" elides the activity of combining representations—how the referent "I" is made available to cognizing subjects. Whether "I" is a use of "I" as subject or "I" as object, Kant's Transcendental Unity of Apperception (TUA) enumerates a common ground: "the unity of self-consciousness that makes possible both our synthesizing representations into conceptualizable wholes […]

---

[58] Sydney Shoemaker, "Self-Reference and Self-Awareness," *The Journal of Philosophy* 65, no. 19 (1968): 555—567.

[59] Thomas Nagel, "What Is It Like to Be a Bat?," *Philosophical Review* 4 (1974): 435–50.

[60] Longuenesse, *I, Me, Mine*, 23.

[61] G.E.M. Anscombe "The First Person," in *Self-Knowledge*, ed. Quassim Cassam. (Oxford: Oxford University Press: 1994): 140–59.

and our ascribing thoughts to ourselves in the proposition "I think."[62] Countering the Cartesian cogito argument of "self" as thinking-identity, Longuenesse defends that in cases of "I think" nothing is necessary to competently use the first-person pronoun "I" aside from mastery, implicit or explicit, of the Fundamental Reference Rule (the "thinker-rule").

On one hand, "I" refers to the producer of the thought. However, the predicate attributed to "I" produces a kind of consciousness of self that is the basis of making a statement where "I" is indicated, which references embodied consciousness ("I am running"), thinking ("I think the proof is valid"), or both ("I *see* a magnolia"). When I say, "I think this proof is valid," there is nothing beyond my being engaged in that thought that makes the predicate valid.[63] Consequently, *all* uses of "I," even uses of "I" as object, depend at least in part on the kind of information on which the uses of "I" as subject depend. This infallibilism regarding cognitive justification ties epistemology to metaphysical realism, proscribing any first-order appeal to logically contingent empirical premises other than those pertaining to first-person reports of appearances to oneself and, simultaneously, constructing justification outside of access internalism by prioritizing first-person cognizance, awareness, and the relevant justificatory factors to any predicate at issue.

Longuenesse here challenges Sellars, who contends that the actual existence of individuals is privy to recognizance solely because our sensory states, our thoughts, and our language are structured by functional counterparts to individuals, their attributes and our experiences of them.[64] Consider uses of "I" as subject (e.g., "I have tooth-ache," "I am angry," "I see red"), where the content of the relevant state is conceptually articulated into propositions, and the thinker of the propositions thereby thinks of themselves, the thinker and speaker, as the entity of which the relevant predicate is true. The concept "I" presupposes the exercise of the capacity for unifying and conceptualizing mental contents, as outlined by Kant's TUA, which demonstrates how the analytic unity of apperception is necessary for any empirically determinate self-conscious

---

[62] Ibid., 31.

[63] Longuenesse, *First Person in Cognition*, 3-33.

[64] Wilfrid Sellars, *Science and Metaphysics: Variations on Kantian Themes* (London: Routledge & Kegan Paul, 1978), 30.

experience. For Longuenesse, the very exercise of this capacity is conceptually expressed in the proposition "I think." Satisfying the reference-rule for "I" calls for nothing more than being, in any instance of the use of "I," *awareness of one's being engaged in thinking.* This is the case despite, in many uses of "I," the predicate that is self-ascribed is a predicate referring to some bodily property.

Consciousness of oneself as thinking is, as a matter of empirical fact rather than as a matter of a priori argument, intimately connected to awareness of one's own body. This is a lesson we draw out from Kant's Transcendental Deduction of the Categories and the Paralogisms of Pure Reason. In the Transcendental Deduction, Kant develops an analysis of the type of self-consciousness grounding the proposition "I think" and the role of "I" in that proposition.[65] Parrying with Descartes' cogito argument from the Second Mediation, Kant's Paralogism of Substantiality criticizes what he takes to be the fallacious inference by which rationalist metaphysicians support their claim that the referent of "I" in "I think" is a soul, a thinking substance.[66] In the Paralogism of Simplicity, Kant criticizes the inference that thinking, like an action, cannot be composite. Thought under the first-person standpoint, the entity "I" is conceived to be 'simple' (indivisible) in virtue of consciousness' stringing together intentional thoughts as logically connected.[67] From the objective "third-person" standpoint, the entity referred to by "I" may be divisible—a composition of metaphysical subjects. As Longuenesse notes, the subjective "I" expresses normative commitment to the unity of thinking (its connectedness and consistency) is the representation of "I" by which I refer to myself in thinking, a universal condition, "for any thinking being, of being engaged in the activity of thinking."[68]

In the Paralogism of Personality, Kant criticizes rationalists for, again, engaging in a fallacy of equivocation.[69] For rationalist metaphysics, as thinking beings we are immediately aware of our own numerical identity at different times and, as such, we are persons. Kant argues that we are aware of our own numerical identity at different times not in virtue of mere thinking (thinking "I think"), but

---

[65] Kant, *CPR*, A341/B399-40.

[66] Ibid., A349-351.

[67] Ibid., A352-361; B407–8.

[68] Longuenesse, *I, Me, Mine*, 131.

[69] Kant, *CPR,* A362–366.

insofar as we, as thinking beings, are capable of consciousness of our continued existence as spatiotemporal, empirically given, embodied entities. Accordingly, Kant offers resources for developing a positive notion of persons as embodied entities endowed with the unity of apperception. Yet this third-person objective standpoint is embedded in the more fundamental first-person standpoint. Notably, this is not the conclusion Kant draws from his criticism but one Longuenesse teases out and, thus, the center of our Kantian infallibilism: Longuenesse denies that the consciousness of the numerical identity of any entity, including oneself, is possible for us other than by relying on criteria for identifying and re-identifying ourselves as that entity in space and time.

Longuenesse's Kantian infallibilism can be naturalized. Vis-à-vis empirical cases such as Oliver Sacks' "disembodied lady" and scenarios of deafferented patients, we see the order of logical connection when it comes to the relationship between proprioceptive consciousness of one's own body and the consciousness of oneself *as a self*, an entity that counts as the referent of "I" whenever "I" is used by that entity.[70] Consider Christina's case:

> Following an antibiotic treatment in preparation for a relatively minor surgery, Christina started gradually losing the sense of her own body. She could stand only if she looked down at her feet and hold things only if she kept her eyes fixed on her hands. She reported that she could not feel her body any more. It was as if the parietal lobes of her brain were not processing the sensory information from her body. In fact, tests showed that the parietal lobes were working, but they had no information to work with. She […] in a state of total proprioceptive deficit, going from the tips of her toes to her head. The diagnosis was one of acute polyneuritis, not the one known as the Guillain-Barré syndrome, which affects motricity, but a version that affected the proprioceptive fibers only. Christina undertook to compensate for the loss of proprioception by using vision to monitor the movements of her body. As a result, her vision automatism and reflexes were increasingly integrated and fluent.[71]

Christina, having lost all sensations of movement and position-sense below the face area, reports being unable to control her movements and thus less "in

---

[70] Oliver Sacks, *The Man Who Mistook His Wife for a Hat, and Other Clinical Cases* (New York: Touchstone, Simon & Schuster, 1998).

[71] Longuenesse, *I, Me, Mine*, 33.

her body."[72] The loss of proprioception requires Christina to acquire information from vision and use it to control and correct her movements. Christina's experience of self-location is not the third-person embodied experience as a physical entity; she experiences herself as a bearer of a point of view that can locate her body in space, for it cannot locate itself. The self-ascription of visual experience and self-location is immune to error through misidentification relative to the first-person pronoun, her body experienced as hers only through the mediation of deliberate evaluation of its position as an object she *sees* rather than feels. When she stops monitoring her own posture/body, she becomes slumped, a ragdoll-like state. When Christina refers to herself as "I," she engages with a kind of unity that allows her to string together perceptual information, engage in inferential reasoning, and keep rational control over her actions. Thus, despite Christina has lost, along with her proprioception, the "fundamental organic mooring of identity"—the "body-ego"—she nonetheless uses perception as an "I."[73]

Thus the "I" in the first-person "I think" serves to bind objective representations to come up with concepts, combined in judgments, connected in inferential patterns. Christina's perceptual representation is subsumed under the concept of the first-person "I." Following Kant, perceptual knowledge, however elementary, depends on mental capacities having complex a priori structures. Intuitions, as percepts are conceptualized, synthesized to yield experience. The transcendental unity, a process on which justification of oneself *as* "I" depends, is "the work of imagination, a blind though indispensable function of the soul, without which we would have no cognition at all, but of which we are seldom even conscious;"[74] notably, there is no phenomenal consciousness of this activity. Since there is no intentional consciousness—the facet of consciousness that confers an object to phenomenal reports, i.e., that what *this* phenomenal consciousness is consciousness *of*—of this synthesizing activity of the imagination, it is unlike both the higher-order mental activity that differentiates intentional objects under concepts and the "what it is like" of lower-order phenomenal

---

[72] Sacks, *The Man Who Mistook His Wife for a Hat*, 42—52.

[73] Ibid., 52.

[74] Kant, *CPR*, A78/B103.

consciousness. Indeed, we are conscious of this binding activity mostly through its results.[75]

There are, thus, two orders of consciousness: i) a fundamental qualitative "what it is like" where consciousness is a property of representational states; ii) a higher-order consciousness, "a representation that another representation is in me," by way of which we represent the identities and differences of objects in space and time, where these representational states can, themselves, become objects of representation, insofar as we are conscious of having them.[76] As Longuenesse notes, "[t]his higher-order consciousness depends on the merely qualitative consciousness … the 'what it's like for the subject of the representation' character of sensations."[77] This distinction between a representation's being conscious by virtue of being qualitatively or phenomenally present and its being conscious by virtue of being conceptualized is comparable to Ned Block's distinction between phenomenal and access consciousness.

Whether cognition has its own phenomenology of thought or whether the phenomenology of cognition is entirely perceptual remains central to Block's query. For Block, perception is not linguistic in form but iconic (i.e., picture-like) and, thus, we have to understand how a picture-like representation could be a conscious representation. Phenomenal consciousness, or "P-consciousness," as Block calls it, is experience and what makes a state phenomenally conscious is that there is something "it is like" to be in that perceptual state.[78] Originally, Block considered access consciousness, or "A-consciousness"—the aspect of consciousness that plays a role in cognitive life (thinking, noticing, etc.) using the perceptual information of P-consciousness for reasoning and control of action—in virtue of one's having the state, if a representation of its content is: i) inferentially poised for use as a premise in reasoning, ii) poised for rational control of action, and iii) poised for "rational control of speech."[79] Access consciousness was

---

[75] Longuenesse, *I, Me, Mine*, 181—182.

[76] Immanuel Kant, *Lectures on Logic*, trans. J. Michael Young (Cambridge, UK: University of Cambridge Press, 1992), 33.

[77] Longuenesse, *I, Me, Mine*, 180.

[78] Ned Block, "Concepts of Consciousness" in *Philosophy of Mind: Classical and Contemporary Readings*, ed. David Chalmers (Oxford: Oxford University Press, 2002), 206—19.

[79] Block, "On a confusion," 231.

accounted for partly in terms of poise and partly in terms of attention. Thus, P-consciousness without A-consciousness can exist, as when we recognize that we had been aware all along of the humming of the refrigerator only once it stops.

After Burge's "modal objection"—which argues that it is objectionable that any sort of consciousness could be thought of as dispositional in the way that Block suggests—Block eventually dropped the poise definition in favor of a model more akin to the "global workspace theory of consciousness" which requires representation rather than either the mental state or the subject who is in the state to have the features mentioned in the previous definition.[85] Notably, however, Block does not endorse Bernard Baars' global workspace model tout court. Block originally theorized access consciousness such that for any subject S and any psychological state X of S, this state X is an access conscious state of S if and only if X is poised for use by S in the rational control of S's thought and action. Block's revision offers that for any subject S and any psychological state X of S, X is an access conscious state if and only if X consists in a representation that is globally broadcast for free use in the rational control of thought and action. Indeed, this notion of "broadcasting" is ineliminably tethered to Baars' theory of communal processing and Dennett's notion of consciousness as cerebral celebrity, where information is global not only along the functional register but also anatomically, where cognitive information is distributed throughout the cortex and other brain areas. However, Block also brings to bear the representational nature of intention—the central nexus of this broadcasting being world-involving, affordance-detecting, and action-guiding representation. What is critical is that a state, S, is access conscious if it consists in a *representation* that has this sort of global cognitive property. Following Block's replacement, what is critical is not that a representation *can* be broadcast for free use but that it *is* actually being used by some system and, thus, is linked to picture-thinking: A-consciousness may necessarily entail p-consciousness, however, if, as Block has said elsewhere, for a percept to have a certain phenomenal character is for it to have a certain representational content.

With this revision, A-consciousness refers to the global availability of the information in an experience and is dispositional (and dispositionally distributed) while P-consciousness is occurrent. The question arises where, as a result of this replacement, A-consciousness can occur in a zombie: e.g., a philosophical zombie

lays their hand on the stove and the panging heat, distributed as described above, causes the zombie to turn their hand away from the oven, all the while not feeling the "what its likeness" of pain. Considering the *possibility* of representations that are poised for report but that are not phenomenally conscious and so not conscious in any conceptual sense, Block notes:

> What I would say now is that one has a choice in framing a definition of A consciousness. One can adopt a pure information-processing notion and face the fact that since it will categorize a zombie functional duplicate as conscious, it will not accord with many people's intuitions about the term. Or one could in effect require P consciousness for some states of an individual who has an A conscious state, thereby framing a notion that fits the ordinary role of the sense of consciousness that emphasizes function.[80]

Regardless of which version one chooses, it is clear that while A-consciousness may functionally be linked to percepts, it is not equipollent to picturing; rather, it is *poise guided by intentional relating* (intentional relating captures information-registration without necessitating attention or awareness). Although A-consciousness originally was analogous to the Kantian conception of higher-order consciousness—an intentional mental state where representations are reflected under a concept for judgment, allowing for reportability—the information-processing interpretation offers A-consciousness as a cognitive successor-concept to Longuenesse's understanding of the apperceptive "I" (behind "I think"), the egocentric index grounding the unconscious binding or synthesizing activity of the imagination. Subjective experience can be understood as a broad category and consciousness as a narrower category within it—the binding that Kant speaks of we, today, might refer to via the "global workspace" that maps onto Blocks information-processing conception of A-consciousness. It follows that content is conscious once it enters the global neural workspace, where it gains the capacity to be broadcast to a host of higher-level cognitive systems for executive control, motor control, reportability, and so on. This strategy helps us pinpoint the mechanism in Christina's brain that endows perceptual content with a set of

---

[80] Ned Block, "Poise, Dispositions, and Access Consciousness: Reply to Daniel Stoljar" in *Blockheads!*, 539.

properties considered central to the perceptual unity necessary for first-person cognition's reference rule. Researchers in cognitive psychology and neuroscience have identified a distributed communication protocol integrating a hierarchy of (domain-relative) modular processors with high-level processors endowed with long-distance interconnectivity (viz. cortical neurons with long-range axons in the prefrontal cortex and parietal regions).[81] This networked mechanism located in the prefrontal region of the brain is the engine behind first-person cognition, performing the function of conscious monitoring and control.

Although we, donning the physicalist stance, can point to biochemical mechanisms in the brain for phenomenal consciousness, our best information about *where* in the brain consciousness happens is that every conscious content is processed by the brain-area which processes that kind of information. We know that conscious contents of motion have to do with the actions in the Medial Temporal area of the visual cortex and likely involve reciprocal connections to the lower visual area. Similarly, conscious appreciation of faces has to do with activation of the fusiform face area in the bottom of the right temporal lobe. Yet we have no "place" or unified network for pinning down phenomenal consciousness although we do understand that P-consciousness is not an intentional property while, simultaneously, intentional differences can make a P-conscious difference.[82]

Facing the impasse that P-conscious difference meets when faced with the hard problem of consciousness, we need not mean embrace Dennett-style weak emergentist illusionism (nor need we embrace Chalmers' strong emergentist panpsychism). Of course, the existence of NCCs by itself does not demonstrate that illusionism is false—illusionists do not deny that there are conscious experiences but deny that phenomenal consciousness is real, i.e., that conscious states consist of intrinsically contentful or simple, irreducibly subjective and immediately grasped properties. For the illusionist, there are consciousness-type experiences but no phenomenal properties and, consequently, the illusionist can accept first-person reports about conscious states but not interpretations of those

---

[81] Lionel Naccache and Stanislas Dehaene, "Reportability and illusions of phenomenality in the light of the global neuronal workspace model," *Behavioral and Brain Sciences* 30 (2007): 518—519.

[82] Ned Block, "Two Neural Correlates of Consciousness" in *Consciousness, Function and Representation* (Cambridge, MA: MIT Press, 2007), 343—363.

reports that invoke a metaphysics of phenomenal properties Block, like Patricia Churchland, offers us a meaningful mode to endorse the theory of neural correlates of consciousness (NCC)—there are, indeed, promising experimental approaches to NCC that do not rely on subjective reports, such as Doris Tsao's work on macaques.[83] Following Block's program, there exist a variety of paradigms in which we can use convergent evidence involving varying degrees of access to separate out the Phenomenal from Access NCC (e.g., signal-detection theory approaches) without regressing into a "winner-take all" model of a "single NCC" paradigm. More broadly, the claim that consciousness is unobservable is the kind of folk-philosophical armchair speculation Churchland's eliminative materialism and that Block distinction between a Phenomenal NCC and an Access NCC oppose.

Insofar as the panpsychism is concerned, the panpsychist has to explain how, even if they espouse stuff-monism, their fundamental smallest unit of consciousness interacts with electrons, magnetic fields, brains, and so on. Does it dissipate when brains die? More broadly, the panpsychist has to provide an account for how its minimal unit relates to brains, what are its causal properties, and, if it does or does not have mass, whether it also has energy.

Imagine that the panpsychist thus responds with the following argument for idealist physical constructivism:

P1: Whatever we have direct perceptual awareness of is radically constructed.
P2: We have direct perceptual awareness of the physical (our bodies, environment).
C: Therefore, the physical is radically constructed.

In response, we will, keeping in mind philosophy after Kant and Sellars—which evinces that to be a direct realist about the external world is fallacious—reject P1. Our senses are not "literally direct" due to the brain's preprocessing. While our conscious experiences *are* constructed by the brain's preprocessing, these constructed experiences turn out to be rather predictable maps to the non-constructed external real world. Given that visual and auditory senses are tightly

---

[83] Doris Y. Tsao, Le Chang, and Pinglei Bao, "The representation of colored objects in macaque color patches," *Nature communications* 8.1 (2017): 1-14; "Representation of multiple objects in macaque category-selective areas," *Nature communications* 9.1 (2018): 1-16.

correlated, when we see someone hit a drum, for instance, we can reasonably be sure we are in practical contact with reality.

Following Block and the late Putnam, we have a rather robust account of a Turing-type computational theory of mind complimented by weak representationism, according to which for a subject to be intelligent is for the subject to instantiate a particular computational property. This offers a methodological solution to problem of how intelligence and rationality find a place in the physical world by realizing computational properties in physical properties, even if they are not reducible to such properties. The right computational properties are what makes Block's China-body robot and Commander Data intelligent. Block's functionalist theory of representation allows for multiple realizability—so how do beliefs and desires, or states in head that *realize* beliefs and desires, represent certain states of affairs? According to Block's long-arm functionalism, which allows for multiple realizability, the representational contents of internal physical states are determined by a combination of: i) their relations—usually causal or informational relations—to the external relations they represent, and ii) their functional role in producing outputs and interacting with each other.[84] This is contra the strong representationism of Jackson and Bill Brewer, where there is a direct mapping between the external objects of the world and these representational states. Thus, rejecting P1, even if our conscious experiences are constructed by neural preprocessing, they functionally respond to objects in the external environment reliably. Accordingly, our constructed experiences end up as reliable maps to the non-constructed external real world.

P-consciousness is experience, P-conscious properties are experiential properties, and P-conscious states are experiential states. Since P-consciousness differences seemingly make an intentional difference, it requires that we point to those empirical experiences (tastes, smells, etc.). Thus the challenge arises for the Kantian naturalist:

> [A] realist about phenomenal consciousness should not find first-person operationalism any more palatable than third-person operationalism.

---

[84] Ned Block, "Advertisement for a Semantics for Psychology," *Midwest Studies in Philosophy* 10.1 (1986): 615—678.

> Consider the reduction of phenomenal consciousness to reflective consciousness…if phenomenal consciousness is reflective consciousness, then for the pain to be phenomenally conscious is for there to be a higher-order thought about the pain, or at least some sort of cognitive state that scans it.[85]

Block distinguishes P-conscious properties to be distinct from any cognitive, intentional, or functional property.[86] P-consciousness binds us neither to epiphenomenalism nor Cartesian mind-body dualism—quite the opposite, in fact, as Block asserts that we can be a physicalist about the cognitive activity related to P-consciousness, allotting a naturalist-cum-physicalist account that well supplements Longuenesse's. Sellars' line of questioning had already opened up the question of how do beliefs and desires, or states in the head that *realize* beliefs and desires, represent certain states of affairs? As demonstrated by the conceivability of spectral inversion, or the idea that the same object might look red to one person but green to another, phenomenal variation is independent of how things are represented as being in the world around the perceiver or the way the world itself is presented in experience.[87] Refining Sellars' line of thinking, we see P-consciousness concepts as "a matter of the causal role of the expression in reasoning and deliberation and, in general, in the way the expression combines and interacts with other expressions so as to mediate between sensory inputs and behavioral outputs."[88]

Kant's "unity of synthesis" (*Verbindung*)—the combinatory "act of the spontaneity of the power of representation"—opens the question of whether the representational format of imagery in the brain is pictorial (iconic) or descriptive (discursive).[89] Block shows that the representational format of imagery and perception in the brain is *iconic*, opposing that consciousness is *reducible* to

---

[85] Ned Block, "Review of Daniel Dennett, *Consciousness Explained*" in *Consciousness, Function and Representation*, 129—141.

[86] Cognitive = essentially involving thought; intentional properties = properties in virtue of which a representation or state is about something; functional properties = e.g., properties definable in terms of a computer program. See: Block, "Concepts of Consciousness," 207.

[87] Ned Block, "Advertisement," 93.

[88] Ibid.

[89] Kant, *CPR*, B130.

functional organization, representation, or cognitive access.[90] The phenomenal character of experience cannot be exhausted by such representational contents as seeing, hearing, smelling, tasting, and having pains; phenomenal consciousness *is* separable from cognitive access. While cognition (thought, reasoning, problem-solving) is sparse due to our working memory, actual conscious perception is rich due to phenomenal awareness—this shows that conscious perception is at least partially distinct from cognition. Such conclusion are further verified by partial-report superiority experiments, where there is phenomenally conscious information that is not necessarily cognitively accessible.[91] The separability of the neural basis of phenomenal consciousness from the neural machinery of the cognitive access that underlies reports of phenomenal consciousness becomes pellucid in Christina's and other deafferented patients' being conscious of themselves as "I" despite being disembodied and losing cutaneous sensation. Other empirical cases include change blindness, where subjects can report the details of only a few objects, but subjectively seem to see the entire visual scene in front of them. The inattentional accessibility view says that subjects may legitimately *see* features that change but fail to notice the difference because although much of the detail in representations is phenomenally registered, "it is not conceptualized at a level that allows cognitive access to the difference."[92] As indicated by Block's description of "mental oil," perceptual experiences have both a phenomenal mode of presentation and representational content, in which the phenomenal mode does not supervene on the representational content, or vice versa.[93] This props up Longuenesse's claim that there is no phenomenal consciousness or intentional consciousness characterizing the mental activity of binding representations.

Rejecting this physicalist biological picture, the mind-body dualist, however, argues for an interactionist picture. For the substance dualist, immaterial minds, outside physical space, causally interact with material objects in space. In the

---

[90] Ned Block, "Mental pictures and cognitive science," *Philosophical Review* 93 (1983): 499—542.

[91] Ned Block, "Consciousness, accessibility, and the mesh between psychology and neuroscience" *Behavioral and Brain Sciences* 30, no. 5 (2007): 481—548.

[92] Ned Block, "Phenomenal and Access Consciousness Ned Block and Cynthia MacDonald: Consciousness and Cognitive Access," *Proceedings of the Aristotelian Society* 108 (2008): 296.

[93] Ned Block, "Attention and mental paint," *Philosophical Issues* 20 (201): 23–63.

Sixth Meditation, Descartes writes:

> There is nothing that my own nature teaches me more vividly than that I have a body, and that when I feel pain there is something wrong with the body, and that when I am hungry or thirsty the body needs food and drink, and so on. So I should not doubt that there is some truth in this. Nature also teaches me, by these sensations of pain, hunger, thirst and so on, that I am not merely present in my body as a sailor is present in a ship…I am very closely joined and, as it were, intermingled with it, so that I and the body form a unit."[94]

Notably, Christina does, indeed, experience herself as being present in her own body as a sailor is present in a ship.[95] She does not remember how to feel and has no experience of being intermingled but she instrumentalizes the first-person pronoun, using the "I" as subject such that her ascription of visual experience and self-location are immune to error through misidentification. Christina does not experience herself as a physical entity but as a bearer of a point of view on her body, able to locate her body in space when the body has lost the proprioceptive ability to locate itself. This harmony of performance becomes equipollent to "consciousness of oneself as subject," with Christina's use of "I" transposed in a sensorimotor key, perceptual awareness bolstered by the reference rule, with Christina engaged in the metacognitive unity of agency and action.

Kant's Fourth Paralogism Kant outlines Descartes' two guiding questions:

1) *The Epistemic Question*—are we justified in believing that the existence of outer objects is open to doubt while only the existence of inner perceptions is absolutely certain?

2) *The Ontological Question*—are we are justified in thinking that we could exist only as a mind rather than as an entity that is inseparably mind and body?[96]

While Descartes answers both questions with an affirmative "yes," for Kant with all possible instances of thinking "I think," thinking of oneself as such requires cognizing under a concept ("I") that can only be *subject*, not predicate, in the proposition "I think." Thus, Descartes' "*cogito, existo*" argument is indubitably

---

[94] René Descartes, "Meditations on First Philosophy" in *The Philosophical Writings of Descartes*, trans. John Cottingham (Cambridge: Cambridge University Press, 1984), 67.

[95] Longuenesse, *I, Me, Mine*, 33.

[96] Kant, *CPR*, A347/B405

"directly downstream" from Kant's theory of the "I think," for "I think" is the expression of a unity of mental activity that conditions all the particular instances of 'I think p.'[97]

VI

As Filipa Foot points out in her discussion of the non-hypothetical instantiation of "should," the normative character of Kant's moral judgment does not guarantee its reason-giving force.[98] Namely, this is meant to show that Kant is mistaken in asserting a radical divide between hypothetical imperatives and categorical imperatives with respect to morality. Foot grants that some imperatives are hypothetical ("*if* I want to get to train on time, *then* I should X") and some *seem* to be categorical ("you should be charitable") but that Kant is wrong in saying that all moral demands on us must be categorical. Although moral judgments are normative, so are judgments of manner/etiquette. According to Foot, in every case it is because there is a background of "teaching," or a developmental substratum, that the non-hypothetical "should" can be used. Does this mean that we are to say that there is nothing but the relative stringency of our moral teaching behind the idea that moral judgments are categorical imperatives? Why are some imperatives more forceful than others? For Foot, this is tethered to our cultural training—recalling Aristotle's account of virtues, Foot thinks it folly to regard putatively "moral" demands as having some special force unique from etiquette. Foot's conclusion is that "morality is supposed to be inescapable in some special way … this may turn out to be merely the reflection of the way morality is taught."[99]

Longuenesse's naturalist solution, recruiting Freud, allots a developmental account that grounds psychological compulsion while still engaging with inner mental structuration, as per Kant's initial concern. Moving from theoretical cognition to practical cognition, Longuenesse appeals to Freud's developmental account of mental life to examine how our use of "I" in the moral "I ought (to)" is premised on a type of self-consciousness that has both an individual and a universal aspect—this is not an ontological claim but an epistemological claim.

---

[97] Longuenesse, *I, Me, Mine*, 82.

[98] Fillipa Foot, "Morality as a System of Hypothetical Imperatives," *The Philosophical Review* 18, no. 3 (1972): 305—316.

[99] Ibid., 311.

Furthermore, Kant's view of the structure of our mental life, grounding the use of the first-person "I" in "I think' and in the moral "I ought to," finds a descendant in Freuds 'ego' and 'super-ego.' Our guiding query concerns the relation between the radically individual referent of "I" and the claims to universal validity of those first-person moral models exercised in "I (morally) ought to X."

Freud's genealogy of both ego and superego contributes to our understanding of the combination of particular and universal claims carried by our use of "I" in the moral "I ought to." For Freud, there is a connection between organizing the contents of mental events according to logical rules—organization proper to the ego—and the ability to cognize in the first-person. For Freud, the ego's organization (*das Ich*) is paramount for indexically using the word and concept *Ich* ("I"). This position is strikingly similar to Kant's, for whom unifying our perceptual inputs according to logical rules is a necessary condition for thinking in the first-person. Contra prudential and instrumental imperatives, morality, in the case of the Categorical Imperative, is universalizable because self-legislation has a universal foundation with a standpoint shared by (all) human-beings-cum-rational-beings. This capability of a universal standpoint in cognition allows us to access the particular reasons we may have to act in one way rather than another, as well as claiming normative validity in willing.

Freud explicitly mentions Kant's Categorical Imperative in two instances, remarking that "[a]s the child was once under a compulsion to obey its parents, so the ego submits to the categorical imperative of its super-ego" and "Kant's categorical imperative is the direct heir of the Oedipus complex."[100] For Freud, the reasoning behind Kant's categorical imperative being heir to the Oedipus complex amounts to saying that its categorical nature has *emotional* roots and that the unconditional demands of morality rest upon emotional life, which also sustains the development of the ability to assess, endorse or reject justifications of moral commands. According to Freud's genealogical claim, determinate structure is not provided by reason/its universal principle but the structure of developmental interdictions and renunciations, which are subsequently internalized. Despite this internalization, its original structure, which provides for

---

[100] Sigmund Freud, "The Ego and the Id," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud, vol. 19* (London: Hogarth, 1923), 49; "The Economic Problem of Masochism" in *The Standard Edition of the Complete Psychological Works of Sigmund Freud* (London: Hogarth, 1924), 167.

the context of what is originally categorical in norms of behavior, remains. Acquiring knowledge of facts of the matter about the world (the ego's "reality principle") counters the repeated "spinning" of fantasies born from earlier, uncontrolled traumatic experiences.

For Freud, the ground level of moral attitude is emotion (stationed in the id and its superego) and is a feature of mental life imposed through development from childhood through the stages of adulthood. Regardless of if the *origin* of morality is in the id and its outgrowth, the superego, as the "engine" of the id, is enrolled in the task of internalizing the features of the ego into the mental life of the young child. On Freud's account, it is the proper work of the ego to transform our emotional attitudes—the moral attitude more than any other—into reasoned attitudes. Freud naturalizes practical reason with a developmental account of the structure of mental life; Freud's ego provides a developmental story for just the unity of apperception grounding the use of "I" in "I think." Similarly, Freud's metapsychological account of the unconscious component—'super-ego'—and its compulsive power provides a developmental story for the conflicted structure of mental life that grounds Kant's use of "I" in the moral 'I ought to.' Thus, Longuenesse argues that Freud can naturalize Kant, illuminating the origin of the categorical nature of morality as tethered not to reason but *emotion*.

Both the Kantian model and Freud's genealogical model posit overriding personal interest in favor of the categorical commands of obligations. For Freud, the enlargement of moral concerns runs in parallel with the human infant's development vis-a-vis internalization of norms that trump self-interested rules of instrumental agency, resulting in an integration of norms into the rational positions concerning the world one inhabits. This depends on three factors: i) original internalization of norms; ii) current social context; iii) perceptive expansion of our moral compass according to ego's reality principle. These conditions subsume Strawson's conception of "minimal morality," Strawson's transcendental condition for the existence of any social organization, which moves from the existence of society to the satisfaction of the conditions for it vis-à-vis the typical observance of a minimal set of rules. Focusing on analytic transcendental arguments regarding issues of concept-possession, Strawson's purview focuses almost exclusively upon our concepts and their interrelations— as Kenneth Westphal remarks, "the strongest conclusion Strawson can justify

pertains to how we must 'conceive' or 'think of' our experience, how we must 'take' objects to be, or how we must perceive them 'as' physical objects and events."[101] Unlike Strawson's interpersonal relating qua minimal expectation, however, Freudian self-consciousness is *not* directed at ourselves/self-perception but at *ourselves in the world*. Thus, we argue, the moral "I ought" illuminates a combination of the particular standpoint premised on individual emotional biography (Freud) and rationally endorsed binding (Strawson). Freud, reminding us of the archaic roots of morality in human psychology, voices agreement with Kant's position that the moral attitude is made necessary by the fact that conflicts between egoistic interest and categorical moral commands arise necessarily, and Strawson's espousing revision through restless groping, the moral 'I ought' "engaging each of us individually and all of us rationally with the same urgent normative demand."[102]

VII

Strawson correlates the diachronicity of the scientific picture to tractable moral concepts via the objective and reactive attitude. Disambiguating reference-dependence, Sellars' work on the metaphysics of mind evinces that, even if our observation reports about the world are typically caused by sensory impressions, the empirical concepts and judgments made on their basis are not by default about sensations. The Strawsonian-Sellarsian lesson, despite both philosophers' shortcomings, proffers characterizing conceptual behavior in terms of socially instituted attitudes and generating a dialogic and social space of interactions between multiple agents who, themselves, are developmentally situated.

As Freud relates to Kant, Longuenesse draws out the imagination's role in producing and binding images in what Freud calls the "pre-conscious" and thus the "ego," where images and their associations are tethered with words and thus available for judgment. Most critically, the Freudian subject can come to know—noninferentially and nonobservationally—the repressed content of its unconscious even though the thought is not *phenomenally conscious*. Just as Freudian presentiments and repressed memories can be unconscious, facing

---

[101] Kenneth Westphal, *Kant's Critical Epistemology*, 67.

[102] Longuenesse, *First Person in Cognition*, 64.

phenomenological blockage, so is A-consciousness. Freudian unconsciousness does not require P-unconsciousness, but it does require A-unconsciousness, as A-consciousness indicates the cavernous substratum of where traumatically-licensed phenomena (e.g., repressed mental images) are suspended until they are culled by, say, psychoanalytic therapy.

The super-ego, the repository of an individual's moral standards, is, for Freud, incredibly plastic during childhood-to-adolescent development, internalizing external demands. The ego is the part of the psyche in which mental processes are organized in such a way that "what the world is like" finds representation in the mind and is projected through the perceptual-consciousness system; the super-ego (or the "ego-ideal") is the unconscious aspect of the ego that sources much of our social adaptation and, therefore, is the source of moral imperatives. The unconditional character of morality is, for Kant, originally grounded in pure reason affecting the faculty of desire; for Freud, it is originally grounded in the raw emotion that binds us to the (authority-)figure from whom we have learned the rules of our socialization. In its metaphysical probing, Longuenesse's naturalized Kantianism offers the intimate connection, and distinction, between *consciousness of oneself in thinking* and *consciousness of one's own embodied existence* without appealing to the noumenal realm.

Ekin Erkan
CUNY Graduate Center
ee2447@columbia.edu