6

Freedom in Sartre's Phenomenology: The Kantian Limits of a Radical Project

Sorin Baiasu

Sartre's phenomenological existentialism can be safely characterized as a philosophy that places significant emphasis on freedom.[1] The view of freedom Sartre's phenomenological existentialism is standardly taken to promote is radical.[2] As he often presents his views on freedom by contrast with Kant's, his existentialism is regarded as a radicalization of the Kantian account. Given this radical character, Sartre's existentialism is often criticized for its implausible implications. In what follows, I would like to challenge this reading. I focus on one perspective from which freedom is usually discussed and from which existentialism appears as radical. I show that the implausible implications usually attributed to existentialism do not in fact apply. I argue that this reading of Sartre is plausible, since he does object to some aspects of Kant's account. By objecting to Kant in this way, he seems to intend to take his own account in a different, more radical direction. Nevertheless, while Sartre's view of freedom is radicalized by comparison to Kant's, this is the case in a different sense than that standardly attributed to him and without the implausible implications for which he is usually criticized.

S. Baiasu (✉)
Keele University, Keele, UK
e-mail: s.baiasu@keele.ac.uk

C. D. Coe (eds.), *The Palgrave Handbook of German Idealism and Phenomenology*, Palgrave Handbooks in German Idealism,
https://doi.org/10.1007/978-3-030-66857-0_6

In the next section, I distinguish between two claims concerning the radical character of Sartre's account of freedom, I identify their radical implications, and I present some objections formulated to these interpretations or to recent accounts of freedom that are taken to follow in Sartre's footsteps. The third section explains why such a reading is plausible by focusing on a particular objection Sartre formulates to Kant's account of freedom and on the solution he proposes. The fourth section undermines this reading by challenging one of the claims of radicalization. I show that the implausible implication usually criticized does not follow from the account of freedom advanced by phenomenological existentialism. Before I conclude, I show in the fifth section that Sartre avoids the implausible radicalization of Kant's view of freedom considered here, but his view of freedom makes some significant steps away from Kant's account.

The Kantian Background of Sartre's Existentialism

How radical is Sartre's account of freedom? Following a standard reading, quite radical, since it manifests this character both for its conception of what is usually termed "negative" freedom and for its account of autonomy or "positive" freedom. Thus, on the one hand, existentialism is understood as defending a libertarian view of freedom, according to which to be free is to be free from all constraints.[3] This "negative" understanding of freedom—an understanding in terms of what freedom is *not*, namely constraints—suggests Sartre is an incompatibilist, who regards freedom and determinism as incompatible and defends a view of moral agents as free.[4] Call the version of existentialism which asserts this view of negative freedom "libertarian existentialism."

On the other hand, however, Sartrean phenomenological existentialism is also regarded as taking a moral agent's autonomy to the extreme. On this reading, existentialism considers a moral agent as autonomous in the sense of being self-governing. This understanding of freedom is "positive," since it specifies how a moral agent should be in order to be free (to wit, self-governing), rather than presenting it negatively by specifying how the agent should not be (namely, constrained). To govern oneself or others, however, we need a law or set of laws, and existentialism is read as departing from a conception of these laws as merely adopted by the agent in virtue of their validity; instead, on this reading, Sartrean phenomenological existentialism regards a moral agent as autonomous when she authors the standards through

which she governs herself.[5] Such standards are not formulated independently from the individual agent, who would then merely adopt them and self-legislate accordingly; instead, they are created by the moral agent, who then governs herself through them. Call this version of existentialism, which asserts this interpretation of positive freedom, "constructivist existentialism."

It is important to note that the negative and positive notions of freedom do not refer to different things—different types of freedom; instead, they are two related ways of presenting the same thing—a free agent. In fact, we can see that the constructivist existentialist view of freedom as autonomy is a particular case of the libertarian existentialist view of freedom presented negatively. If freedom is defined negatively as the absence of constraints, then one particular case of this view of freedom is the case of a free agent who is not constrained in her self-governance by a law or laws formulated independently from her, but who constructs her own standards and governs herself through them. Existentialist autonomy, on this reading, is the result of removing the constraint of a law of self-governance formulated and justified independently from the individual agent.

Interestingly, both these readings of phenomenological existentialism can be seen as historically motivated by the attempt to radicalize a Kantian account of freedom. Moreover, as we will see, in discussing his account of freedom, Sartre often compares his views with Kant's. Although Kant is sometimes read as a libertarian about the metaphysics of freedom, he is by no means a standard libertarian, since he accepts that our actions, while in one sense possibly free from the constraints of the laws of nature, are, in a different sense, determined by these laws.[6] By rejecting this deterministic dimension of the Kantian view of freedom, phenomenological existentialism is read as an extreme version of libertarianism.

Moreover, Kant's account of autonomy defends a view of freedom as self-legislation, but the law through which this process of self-governance is to take place—the Moral Law—is a law which is not authored by the individual agent, but is a structure of practical reason and is justified as normative without reference to aspects of individual agency. By rejecting the status of this law as independent from the individual agent as individual, phenomenological existentialism seems to take to an extreme the Kantian view of autonomy.

Both the libertarian and the constructivist existentialisms, as radical conceptions of freedom, end up with some implausible implications. According to the libertarian account, we are free from all constraints, including the constraints of the laws of nature, and, hence, we can always act freely. Yet this leads to the implausible implication that we are free even in the

contexts of very oppressive institutions; moreover, in the context of institutions which generate systemic injustice, the assumption that we are radically free renders any emancipatory discourse superfluous.[7]

According to the constructivist account, we are free whenever we act in accordance with the standards we construct, although it is unclear whether such a process of construction could not be conditioned or even determined by certain internalized norms; furthermore, even assuming that the process of construction is free from such constraints, we end up with the implausible suggestion that acting in accordance with arbitrarily constructed standards would have to be accepted as free agency.[8]

The first charge relies on a reading of Sartre's view of freedom as involved in the traditional metaphysical debate between free will and determinism. Sartre does talk about this debate but frames it in different terms: not as a dispute between a claim to the existence or non-existence of uncaused action, but of an action without motive. The traditional reading might be explained by the choice, in the most popular translation of *Being and Nothingness*, of the French "*motif*" by "cause."[9] For instance, Sartre's claim (at EN 435–36, 480) is not that no factual state can by itself *cause* an act, but that it cannot by itself *motivate* it. Moreover, Sartre does not simply talk about freedom but distinguishes between various aspects of free agency. On Sartre's account actions are always motivated by reasons, and while actions may always be ontologically free, politically they may well be constrained and may need a discourse of emancipation for the individual's liberation.[10]

As already mentioned in the first section, in what follows, I would like to focus on the second charge of implausibility. I will eventually argue that Sartre's account of autonomy is not as radical as standardly interpreted, but accepts the main limitation of the Kantian account that he was taken to remove, although there are aspects of the Kantian account with which he parts company. It follows that Sartre's existentialist phenomenology avoids the implausible position of the constructivist existentialist, although it departs from Kant's position in the attempt to draw more consistently some of the conclusions of the Kantian project. In this way, Sartre's phenomenological existentialism paradoxically turns out to be closer to Kant's account of freedom than he is to the accounts of some contemporary existentialist Kantians.

Now, one question we can begin with is why Sartre has been interpreted as defending a constructivist version of existentialism. Answering this question will be the task of the next section.

Sartre as Constructivist Existentialist

In the second part of *Being and Nothingness*, entitled "Being-for-Itself," Sartre discusses freedom in the context of a section devoted more generally to "Temporality"; there, in part B ("The dynamics of temporality") of the second sub-section ("The Ontology of Temporality"), he contrasts his view of freedom with Kant's account of the spontaneity of a person. Thus, he claims:

> It would be useless to remind us of the passages in the *Critique* where Kant shows that a non-temporal spontaneity is inconceivable, but not contradictory. It seems to us, on the contrary, that a spontaneity which would not escape from itself and which would not escape from that very escape, of which we could say, "It is this," and which would allow itself to be enclosed in an unchangeable denomination — it seems that such a spontaneity would be precisely a contradiction and that it would ultimately be the equivalent of a particular affirmative essence, the eternal subject which is never a predicate. (EN 171, 188)[11]

This is a complex passage, which needs considerable unpacking. For our purposes, however, it will suffice to provide an outline of Sartre's objection in order to see what motivates the interpretation of his account of freedom as a version of constructivist existentialism.[12] We can perhaps already grasp the gist of Sartre's criticism: the general objection is that Kant provides a contradictory account of freedom as non-temporal spontaneity. Without yet mentioning the Kantian context of this objection, we can already outline the argument Sartre advances. For him, change presupposes time. A non-temporal spontaneity, therefore, should be unchanging. As unchanging, spontaneity is viewed as possessing a character that cannot be modified. More precisely, not even a self-generated modification of spontaneity would be viewed as conceivable. Yet the idea of spontaneity is that of self-motivated possible manifestation unrestricted by any pre-set direction as given by a fixed character. Hence, the Kantian idea of a non-temporal spontaneity is contradictory. The free agent understood as non-temporally spontaneous is congealed in an "unchangeable denomination," that of "the eternal subject which is never a predicate."

Although Sartre clearly points here to a contradiction in Kant's idea of freedom as spontaneity, the details of some of the claims are unclear without a brief discussion of Kant's account. The reference seems to be clearly to the discussion of the third antinomy in the *Critique of Pure Reason* (A445–51/B473–79).[13] This is part of Transcendental Dialectic, the second division of Transcendental Logic,[14] where Kant examines the contradictory claims

at which we arrive by following logically the claims of reason. On Kant's account, an epistemic agent has the capacity to be affected by the world through her sensibility, the capacity to cognize the world through the understanding (which unifies under concepts what the sensibility senses through being affected by the world), and the capacity to unify this cognition by reason. In the attempt to unify this cognition, reason leads us to make claims about the world as a whole.

We may claim, for instance, that everything in the world, including ourselves, is determined by the laws of nature and there is no freedom. In this case, we are hard determinists. We may equally well claim that not everything in the world is determined by laws of nature and there is freedom. Such a claim is made by libertarians. The two claims are contradictory or antinomical—hence, the name of this contradiction. It is possible for equally rational claims to be contradictory, because these are claims about an object we cannot experience—namely, the world as a whole. Claims about objects which we cannot experience are speculative. Such claims generate controversies between speculative metaphysicians, claims between which we cannot adjudicate. Kant thinks that the existence of such contradiction and controversies is a "scandal" and he is, of course, in a certain sense right (Bxxxiv).[15]

Kant's attempt in the *Critique of Pure Reason* is to deal with this scandal by setting reason on "the secure path of a science" (Bvii). He thinks that his transcendental idealism can show that both antinomical claims concerning freedom are correct, but from different perspectives. In this way, the contradiction in which reason entangles itself is dissolved. His solution to the third antinomy, at least in its very general outline, is well known. From the perspective of what we experience, the perspective of phenomena, everything in the world is determined in accordance with the laws of nature. However, from the perspective of things as they are in themselves, independently from our experience, the world is not determined by the laws of nature and we are free. Hence the thesis ("[t]he causality according to laws of nature is not the only causality, … it is necessary also to assume a causality through freedom") and the antithesis ("[t]here is no freedom, but everything in the world occurs solely according to laws of nature") are no longer contradictory.

This background enables us to understand better Sartre's argument above. Since Kant regards space and time as structures of sensibility, things in themselves, including ourselves regarded as such, are non-temporal. Moreover, since "cause" is a category of the understanding, through which we constitute in our experience the world of phenomena, again, an agent, considered as a being in itself, will be spontaneous, independent from any constraints imposed by natural laws. Sartre is slightly confused when he

claims that Kant's freedom is inconceivable [*inconcevable*], but not contradictory. According to Kant, freedom is not cognizable [*erkenntbar*], but it is conceivable or thinkable; in fact to be thinkable is to be free from contradiction for him.[16]

If we understand freedom, with Kant,[17] as the property of a being to which I attribute effects in the world of sense, then there is a way to see freedom as non-contradictory, despite the fact that, with Kant, we view all effects in the world of sense as subject to the laws of nature. More exactly, it is sufficient to accept another kind of causality, that of the spontaneous agent considered in itself, in order to be able to see the effects produced by this alternative causality, even if they are regarded phenomenally, as attributable to the agent, rather than to the laws of nature. The freedom of the agent is, in this way, non-contradictory. It is, however, not cognizable, since, for Kant, cognition requires both concepts and intuitions, and we can have no experience (and hence sensible intuitions) of the agent, regarded as in itself.[18]

The contradiction in the Kantian notion of freedom introduced by the third antinomy is removed by Kant's solution to the antinomy. Yet Sartre continues to regard the Kantian conception of freedom as contradictory. This is because he takes Kant to give a particular determination to freedom, "an unchangeable denomination," a "particular affirmative essence," which makes it impossible even "to escape from itself," let alone "escape from that very escape." We will see shortly how on the dominant interpretation of Sartre's account of freedom, the free agency can be seen as able not only to escape from itself, but also to escape from that very escape. These two conditions are taken by Sartre to be implications of an account of freedom as spontaneity. Yet Kant's account seems unable to meet them: If freedom has such an affirmative essence, as "the eternal subject which is never a predicate," then it cannot escape itself, since to escape itself is equivalent to abandoning this affirmative essence and abandoning this essence is equivalent to becoming something else than freedom (since by definition the essence of something is that without which that thing cannot exist as the thing it is).

To see why Sartre reads in Kant a notion of freedom with an affirmative essence as the eternal subject which is never a predicate, it is sufficient to reflect on the status of the free agent. An agent is regarded as free from the perspective of things as they are in themselves. As such, a free agent is not cognizable and, hence, not predicable—it is similar to an eternal subject, which is never a predicate.

Sartre does not suggest that the Kantian account of freedom should be altogether rejected. He regards Kant's account of freedom as spontaneity as correct, but he thinks Kant's account of spontaneity must be modified. As

we have seen, the modification must allow the free agent to be genuinely spontaneous, and Sartre seems to suggest an alternative notion of spontaneity as autonomy: "*This* spontaneity should be allowed to define itself; this means both that it is the foundation not only of its nothingness of being, but also of its being and that simultaneously this being recaptures it to fix it in the given" (EN 171, 188 — translation slightly amended). If the spontaneous being is self-defining, then there is no affirmative essence, which would congeal the supposedly free agent and would undermine its spontaneity. This view of the genuinely free agent as an autonomous spontaneity seems to be the response Sartre provides to the problem he identifies in Kant's solution to the third antinomy.

It is easy to see now why Sartre's existentialism is read as a version of constructivist existentialism. Sartre does suggest that freedom is best conceived of in terms of autonomy, and he clearly sees this autonomy as radical. The agent is self-defining, the "foundation not only of its nothingness of being, but also of its being," a spontaneity which can "escape from itself" and can also "escape from that very escape" (EN 171, 188). The view of the agent as autonomous in the Kantian sense, as governed by a law the individual agent adopts, but which is not formulated or defined by the agent herself, since it is shared necessarily by all rational agents, would be a view of a spontaneity that may escape itself (escapes its individuality and governs itself according to a law, which is her law insofar as she is a rational being), but cannot further escape this very escape (the law cannot be defined by the agent herself, but can only be adopted as valid).

While it may be now understandable why Sartre is sometimes read as a constructivist existentialist, I have claimed that, in fact, this is not an accurate reading. I have claimed that, in fact, Sartre's phenomenological existentialism is closer to Kant's account of freedom than it is to the constructivist version of some of today's existentialist theories. Nevertheless, there is one respect in which Sartre's phenomenological existentialism is more radical than a Kantian account of freedom and specifically than a Kantian account of autonomy. To show these will be the purpose of the next two sections.

Sartre's Account of Freedom and Its Kantian Limitations

The implausible implications of a constructivist existentialist account of freedom follow from the specific view of autonomy held by this account. In particular, this view of autonomy requires that the standard through

which the moral agent governs herself be her creation. This requirement must be understood in a specific sense. Imagine a Kantian account of action, according to which human agency is centered on maxims, and maxims are created or devised by agents depending on circumstances and personal features.[19] A Kantian account of autonomy requires that these maxims be objective principles. To determine whether a maxim (which is a subjective principle of action) is an objective principle, we employ the Categorical Imperative.[20] For instance, universalizable maxims are morally permissible and may even be required as categorical imperatives. Hence, acting on such maxims would be autonomous on the Kantian account. Yet it would not be on the constructivist existentialist one, although maxims are created by the agent. This is because the principles of action of the autonomous agent depend normatively on the Categorical Imperative. This, however, is not created by the agent, but it is a structure of practical reason of all moral agents.

To show that the implausible implications of constructivist existentialism do not follow from Sartre's account of freedom, I need to show two claims. Firstly, I will defend the claim that, on Sartre's phenomenological existentialism, not any standard of action devised by the agent is an expression of the agent's autonomy. Secondly, I will argue that the condition which must be met in order for a standard to be an expression of an agent's autonomy is not some arbitrary condition. On the contrary, it is a normative criterion which Sartre takes to be normative for all moral agents. For the first claim, it is sufficient to focus on Sartre's distinction between pure and impure reflection. For the second, we can examine his view of authenticity. On the basis of this value he distinguishes between standards of action which are expressive of freedom as autonomy, and those which are not.

I started the last section with a long quotation from the second part of *Being and Nothingness* entitled "Being-for-Itself." Being-for-itself is a central concept of Sartre's early philosophy, a concept which Sartre still uses in his later philosophy to explain some of the new terminology he introduces there.[21] The notion is part of Sartre's distinction between being-for-itself and being-in-itself, usually simplified to the distinction between the for-itself and the in-itself.[22] These represent two regions of being, corresponding approximately to individual conscious beings and inanimate beings—a person is sometimes referred to by Sartre simply by "for-itself" and an inanimate being by "in-itself." Now, for Sartre, the for-itself's reflective consciousness, whether pure or impure, is one among several forms of consciousness of the for-itself. It is a form of self-consciousness, but relies on a more fundamental form of self-consciousness, called pre-reflective

self-consciousness. Moreover, reflective consciousness is a positional type of consciousness. Hence, before introducing the distinction between the for-itself's pure and impure reflective consciousness, we need to examine three other distinctions: between immediate and positional consciousness, between consciousness and self-consciousness, and, finally, between reflective and pre-reflective self-consciousness.

Consider first the distinction between immediate or non-positional consciousness and positional consciousness. When I am perceiving something, I may be conscious of that thing in virtue of an immediate consciousness that I have with regard to it. For instance, most of the time, we are surrounded by countless sounds, which we perceive, but which we do not get to take as the focus of our attention. Immediate consciousness of such sounds is not positional—it will not present its object as coming from a determinate place or as bearing particular properties. Hence, immediate consciousness is not able to make judgments with regard to the world, because it does not posit the object that it perceives. It may, however, subsequently focus on the object and posit it.

Consider now the situation when I am judging something, such as the quality of the paper of the book I am reading. Although, in this instance, I am positing the page and I am involved in the activity of judging the quality of its paper, I may not be reflectively conscious of doing this; on the contrary, I may be absorbed by my activity to the point of forgetting about myself. The focus of my activity is, therefore, the transcendent object of my attention (the page), not my activities, what I am doing.[23]

This leads to the distinction between consciousness and self-consciousness. Self-consciousness, as the name indicates, takes itself as an object. When I am positing my own conscious activities, I am self-conscious and, moreover, I am reflecting on myself. By contrast, when I am positing an object distinct from myself without also positing myself as doing this, I am simply conscious of that object. Finally, consider the distinction between reflective and pre-reflective self-consciousness. Although in the example above, I am trying to establish the quality of the book's paper, I am not reflectively aware of doing this. Yet I cannot say that I am not conscious of myself as doing this. In my mind the thought that it is me who is doing such-and-such a thing does not occur, but I can always explain what I am doing, if someone asks me.

However, I can only explain now what I have just done, if what I have done is inscribed somehow in my memory as having been done by me. But this is not possible, if I have not been somehow aware of myself as having done such-and-such a thing while doing it. Hence, one can be reflectively or pre-reflectively aware of oneself as doing something depending on whether

or not one is positionally aware of oneself as doing that. According to Sartre, any form of consciousness is at the same time pre-reflectively conscious of itself. Hence, my positional consciousness of the book in front of me is also pre-reflectively conscious of itself as such a consciousness. This pre-reflective self-consciousness is distinct from reflective consciousness (consciousness of self) and, to mark this difference, Sartre calls it "consciousness (of) self" (EN liv, 20). Reflective self-consciousness presupposes a positional consciousness of oneself, whereby one's self is posited as an object (an "I" or ego). By contrast, pre-reflective self-consciousness or consciousness (of) self, as pre-reflective, cannot be posited, since it would then have to become an object for a positional consciousness and, hence, simply part of a reflective process.

Now, for Sartre, there is a constitutive similarity between self and value (EN 117, 131). The self is for the for-itself what a value is for our actions. We act in pursuit of certain values—some more distant, such as long-term projects and ideals, some more quickly achievable, such as more immediate goals or the satisfaction of certain desires or needs. Pursuing a value implies bringing about states of affairs which realize that value and without which current states of affairs are perceived as lacking in that respect.

Similarly, the for-itself lacks something in relation to the self and constantly attempts to eliminate this lack and achieve the fullness of the self. Through reflection I can determine what I am lacking, what I need to achieve in order to obtain the fullness of self. Self-sufficiency is the stage where the for-itself achieves independence, since it becomes its own foundation: It becomes at the same time an in-itself, which is not lacking, since it simply is, and a for-itself, which alone is able to make sense of things, such as lack, sufficiency and value. This is why Sartre sometimes calls this ideal of a self-sufficient, full self, the "for-itself-in-itself."

We can now examine the distinction between pure and impure reflection. Impure reflection posits the consciousness reflected-on as an object having the ontological structure of an in-itself. This, however, is not a legitimate process, since the consciousness reflected-on, as consciousness, is at the same time a for-itself. Impure reflection attempts to cognize the consciousness reflected-on, that is, to take it as a given in-itself, and not simply to recognize it as already revealed.

The consciousness reflected on by impure reflection is thus transformed into an object of psychic life. For instance, in the attempt to reach self-sufficiency, the for-itself may try to overcome its lack through certain actions. Through these actions the for-itself tries to change the world, which seems to be lacking. In trying to identify the missing part of its self, the for-itself's impure consciousness will reflect on the for-itself and will perceive its lack

(and the value of the actions which try to change the world to overcome this lack) as a transcendent unity. Instead of seeing the consciousness reflected-on with its desires, purposes, and hopes as part of consciousness and, hence, the result of freedom, impure reflection turns them into objective entities which transcend consciousness and determine it. The unification performed by impure reflective consciousness posits a particular end of the action as given, rather than as chosen. Every act is the result of a synthesis of different consciousnesses reflected-on. Impure reflective consciousness turns them into transcending unities.

> By acts we must understand all synthetic activity of the person, that is, every disposition of means with a view to ends, not as the for-itself is its own possibilities but as the act represents a transcendent psychic synthesis which the for-itself must live. For example, the boxer's training is an act because it transcends and supports the for-itself, which, moreover, realizes itself in and through this training. (EN 185, 202)

Hence, impure reflection synthesizes and objectifies the result, losing sight of the for-itself, whose possibilities are in this way turned into necessities. Therefore, on the level of the act constituted as an object by impure reflective consciousness, my freedom is simply that of realizing the already given ends.

On the level of impure consciousness the person's sensible desires, emotions, and beliefs are objects that are causally determined by other objects. Hence, one does not see oneself as actually realizing an end or as attempting to realize it, but as determined (causally led) to attempt to realize it. Impure reflection and bad faith are very closely related. Impure reflection is reflection in bad faith, and bad faith is an attitude brought about by impure reflection. In other words, impure reflection and bad faith are mutually constitutive. Sartre explicitly denies the deterministic interpretation of the person on the level of impure reflection. The very idea of transforming a psychological state into an object of reflection presupposes the freedom of the for-itself to make explicit and modify reflectively a consciousness (of) acting into the object of positional consciousness. Hence, by making use of freedom in order to present an account of agency as devoid of freedom, "this reflection is in bad faith" (EN 187, 201).

As against impure reflection, existential psychoanalysis reveals "the ideal meaning of all human attitudes," and the fact that this ideal meaning—as a particular value—is ultimately contingent (EN 646, 690). Sartre claims that existential psychoanalysis is in fact practiced by "many men" even before they have learned "its principles." It is, in other words, a practice which need not have its method explicitly presented in order for a person to engage

in it. Since it is a means of escaping bad faith, Sartre also calls it a "moral description" (EN 645, 690).

Existential psychoanalysis aims to realize that catharsis which transforms impure reflection into pure reflective consciousness, since existential psychoanalysis attempts to provide the person with that degree of self-consciousness which, as we will see in the next section, enables her to become aware of her fundamental project. The free choice underlying this fundamental project shows that an action or a psychic state is not determined but is only conditioned by the world.

Hence, a preliminary conclusion we can draw here is that not any projects through which a person defines herself can be considered as ethically good. If such projects are not regarded as grounded in a person's freedom, then their value is not ethically good. Moreover, what drives this criterion for the moral evaluation of a person's project is the value of authenticity, which is seen as valid for all moral agents. These are the two claims I formulated at the beginning of this section as sufficient to demonstrate that Sartre's phenomenological existentialism is not constructivist, in the radical and implausible sense mentioned at the beginning of this chapter. Hence, in this respect, Sartre is closer to Kant than to some contemporary existentialist constructivists.

I have said that another aim of this chapter is to argue that Sartre's phenomenological existentialism does attempt to radicalize Kant's moral theory, but it is not in the sense in which it is usually interpreted to do so. To show this will be the aim of the next section.

Sartre's Moral Theory as a Radicalization of Kantian Ethics

In what follows, my claim will be that Sartre's and Kant's moral theories are structurally similar, but there is a sense in which Sartre introduces into his account some radical aspects. In particular, the second-order standard on the basis of which we can evaluate the moral choices of the agents is in Sartre's case a second-order value, not a law. Moreover, for Sartre, the standards which structure a person's agency are not principles, but values. Whereas in Kant rules of action are guided by the principles the person adopts, in Sartre they are guided by the values they pursue.

On Sartre's account, every action has an end, which can be seen as the instantiation of a more general project. For instance, if I plan to run a marathon, I must also plan to run a few miles every day, and this was the

end or purpose of my running this morning. Moreover, for Sartre, even this more general project of running a marathon can be understood on the basis of a more profound project:

> The problem indeed is to disengage the meanings implied by an act—by every act—and to proceed from there to the richer and more profound meanings until we encounter the meaning which does not imply any other meaning and which refers only to itself. This ascending dialectic is practiced spontaneously by most people; it can even be established that in knowledge of oneself or of another there is given a spontaneous comprehension of this hierarchy of interpretations. (EN 457, 535)

To see this more clearly, consider the following example that Sartre discusses in *Being and Nothingness*: I start out on a hike and after several hours my fatigue increases until it becomes painful; I may then either try to resist it and go on walking, or I may give up and let myself fall down beside my knapsack. Here and in general, each of my acts is comprehensible in terms of its end: "I place my knapsack down in order to rest for a moment" (EN 460, 537). The importance of resting at that point is comprehensible from the perspective of a more general end—perhaps that of wanting to have enough energy left to enjoy the evening properly, where enjoyment is one of the more general ends of my actions. Alternatively, I could go on walking in order to overcome this limit, as part of a general project of undertaking physical training, where enjoying an evening, rather than going directly to bed, is no longer part of my end.

The "ascending dialectic" through which Sartre thinks we "disengage the meaning implied by an act—by every act" is part of the cathartic process of existential psychoanalysis presented in the previous section. As we have seen, the aim of this process is to provide a moral description through which the for-itself can escape bad faith. It is this process practiced by many persons even before they have learned its principle or, as he says above, it is "spontaneously practiced by most people."

One important objection Sartre formulates in relation to Kant's ethics concerns the distinction between imperatives and values. More exactly, the key distinguishing feature that Sartre identifies in the comparison between imperatives and values is their normative force. In order for a value to determine my will to act, I have to choose its end as my possibility. This end is normative ("to be realized") insofar as I chose it; otherwise, it does not determine my will and cannot motivate me.

By contrast, in order for an imperative to be a moral imperative, and hence to be normative, an actual choice of its purpose as my possibility is

not required. In short, values require actual choice, whereas imperatives do not. What is more, as Sartre acknowledges, even if I tried to choose an imperative, I could still not choose its purpose as my possibility, since its normative force goes beyond my actual choice of a purpose: "[E]ven if all our desires were conformed to our obligation and as a consequence served the pure Will, there would still remain an underlying duality that is the source and ground of all the others" (CM 256, 267). Nevertheless, what Sartre identifies here as a distinction between imperatives and values, and between his account and Kant's account is not so much a difference between two moral theories and type of norms, as a metaethical distinction between a purpose (or a principle), as regarded by a descriptive account of action, and a purpose (or a principle), as presented by a prescriptive account of action. Hence, in order to distinguish between values and imperatives, Sartre makes use of the distinction between a descriptive and a prescriptive account of action.

Thus, when he claims that the end represented by a value is "to be realized," if I actually choose it as my possibility, he formulates the necessary and sufficient condition for a person to act in order to reach, against certain adverse circumstances, an end. This is because, when I choose an end, I regard the end as to be realized and, hence, I act to bring it about. By contrast, when he asserts that the purpose of an imperative is "to be realized" independently of, and even against, my actual choice of it as my possibility, he formulates a condition for how a person ought to act.

For Kant, whether or not a person has adopted a certain principle for the determination of the will and has acted on that principle in a certain situation cannot conclusively help us determine whether the action ought to have been performed. We only learn that the action has been possible, in the sense of being physically possible, but this feasibility is, of course, a necessary condition for both permissible and impermissible principles.

According to Sartre's account, in order to determine whether the end of a value is "to be realized" by a person, we should simply determine whether or not the person has chosen that end; actual choice is in this case both a necessary and a sufficient condition for the normativity of an end. Yet, again on Sartre's account, actual choice is no longer either a necessary or a sufficient condition for the purposes of the normativity of categorical imperatives.

A categorical imperative formulates an obligation and can even claim to obligate a person who chooses to do precisely the opposite of what the categorical imperative commands. The problem with this, however, is that a value or purpose or end will only have normative force if it can still impose an obligation, when they have not actually been chosen as ends of actions. Without such a distinction, whatever value a person happens to adopt, it must be

accepted as morally valid. Taken as such, this conclusion encourages again an interpretation of Sartre as a constructivist existentialist about autonomy. It suggests that, on Sartre's account, when a person adopts a standard (a value) and governs herself in accordance with that standard, she is autonomous. Yet, as we have seen, it is not the case that whatever end we adopt, as an end of our action, is also a morally valid end. As Sartre himself acknowledges, values that are chosen in bad faith cannot be morally valid. But since a value which has actually been chosen may not be morally valid, then actual choice cannot be the sufficient condition of moral normativity. And since a value which has not been actually chosen may be morally valid, then actual choice is not even a necessary condition of moral normativity.

This is exactly the situation with Kant's categorical imperatives too: the morally valid imperatives impose an obligation the person ought to follow but cannot determine the person to actually choose it and follow it. For Kant, it is important that this standard be chosen by the person freely because it is the right thing to do. This, however, seems to be also Sartre's requirement—that the standard proposed be followed freely by the person to whom it is proposed, and, hence, be chosen because it is considered to be right.

We end up with two surprising results: an unexpected similarity between Kant's and Sartre's accounts of action and, at the same time, an unexpected difference—the result of Sartre's radicalization of Kant's metaethical views. Thus, for both of them, actions can be described by rules and have ends that represent the projects people have when they act. These projects can also be understood as general policies of action or maxims. Projects and maxims have several rules of action under them and they guide action; hence, they are second-order rules. In turn, these second-order rules seem to be regulated by third-order rules (the categorical imperative, for instance, or the value of authenticity).[24]

Kant's emphasis on rules of action, principles, or maxims and the Categorical Imperative shows that he favors a rule-centered view of action and practical philosophy. By contrast, Sartre's discussion of value, projects, and the higher-order value of authenticity suggests a purpose-centered ethical theory and philosophy of action. There is therefore a radicalization of the Kantian ethical project in Sartre's phenomenological existentialism: the relative rigidity of rules, principles, and imperatives is replaced by the relative flexibility of values, projects, and the ideal of authenticity. Various rules, principles, and imperatives can illustrate the correspondingly same values, projects, and ideal of authenticity. The latter still provide guidance for the agent, and the ideal of authenticity is justified by Sartre as a moral criterion,

rather than being formulated by individual agents. It is not here that the radicalization of the Kantian project can be found in Sartre, but in his emphasis on a value- or purpose-centered ethics, as opposed to Kant's rule-centered practical philosophy.

Conclusion

For both Kant and Sartre, freedom is a central value, and the project of an ethical theory or practical philosophy guided by this value is a shared one. This is not surprising: They belong to a tradition of Western modern philosophy, in which human agents adopt as normative standards those standards they freely acknowledge to be valid. Sartre is standardly interpreted as a radical philosopher of freedom, who takes the moral agent to be free from any constraints, including natural laws, relatedly to be the bearer of an absolute responsibility for what happens in the world, and to be autonomous in the constructivist sense of self-legislation by the standards she herself creates. Given that Sartre often formulates his views on freedom by contrast with Kant's account and on the basis of objections to a Kantian account, he is usually regarded as advancing a radicalized version of a Kantian account of freedom.

In this chapter, I have examined some of Sartre's objections to Kant and the way he thinks he can overcome the Kantian problems, in order to explain why he is standardly interpreted as a constructivist existentialist. I have then presented some fundamental aspects of his philosophy of action and ethics, focusing on the central concepts of self, value, possibility, facticity, and reflection. We have seen that impure reflection is a mode of consciousness that presents the world and in particular an agent's values as set-in-stone, already given demands the agent must pursue. By contrast, pure reflection makes evident the freedom of the individual to commit to specific projects in an authentic mode and to pursue these in virtue of their ethical value. This higher-order value of authenticity, which functions as a test of the moral goodness of an agent's projects, shows that the account of autonomy Sartre's phenomenological existentialism advances is not a version of constructivist existentialism. The moral criterion of authenticity is shared by all agents, and it is not a criterion they create and legislate themselves according to it; it is, however, a criterion they adopt due to its validity.

Secondly, we have seen that Sartre's philosophy of action includes some interesting similarities when compared with Kant's account. For Kant, agency is structured by maxims. Agents perform actions whose rules are derived from

maxims the agent is committed to. Ethical actions are those actions that are derived from maxims that are at the same time principles, that is, maxims that are valid for all agents who are in the same situation and have the same morally relevant features. Whether or not such maxims are valid is decided in accordance with the Categorical Imperative. Similarly, Sartre takes an agency to be structured by projects, which have purposes that the agent tries to realize through her actions. The more specific purposes of these actions are therefore derived from the projects' purposes. Ethical actions are performed with a view to the realization of ethical projects, and whether or not a project is ethical is decided in accordance with the value of authenticity.

This comparative analysis reinforces the conclusion that Sartre is not a constructivist existentialist à la Korsgaard, but offers an account of freedom and in particular autonomy, which is closer to Kant's account. Yet the same analysis also makes evident a significant difference between Kant's and Sartre's account: the difference between a rule-centered and a purpose-centered practical philosophy. This indicates further that, in fact, Sartre's ethical theory does include an element of radicalization, but perhaps one whose significance has been overlooked by commentators. Sartre still regards an agent's autonomy as made possible by a moral criterion, whose validity is independent from the individual agent, just as Kant does. Yet, unlike Kant, Sartre questions the need for an ethics of rules, principles, and imperatives, and proposes, at least in his early writings, an ethics of purposes, projects, and values, an ethics guided relatively more flexibly by normative standards.

This, however, is only part of the picture of Sartre's existentialist phenomenology. While commentators increasingly discuss the continuity between the early and the later Sartre, they are still to notice the change in Sartre's conception of values and imperatives in the later writings.[25] Although a thorough discussion of this aspect goes beyond the scope of this chapter, research undertaken so far[26] shows both a move closer to Kant, but also a significant way in which Sartre does modify Kant's ethics in the direction of a philosophically more convincing critical ethics.[27]

Notes

1. According to Jonathan Webber, for instance, "as originally defined by Simone de Beauvoir and Jean-Paul Sartre, existentialism is the ethical theory that we ought to treat the freedom at the core of human existence as intrinsically valuable and the foundation of all other values" (Jonathan Webber, *Rethinking Existentialism* [Oxford: Oxford University Press, 2018], 2). Whereas regarding

freedom as central for Sartre's existentialism is not surprising, the interpretation of existentialism as an ethical theory is, I think, more problematic, given Sartre's failed attempts to publish an ethical theory (which is, of course, not to deny the ethically relevant discussions in Sartre's published work, de Beauvoir's *The Ethics of Ambiguity* (trans. Bernard Frechtman [New York: Citadel Press, 1964]) and Sartre's approval of Francis Jeanson's *Sartre and the Problem of Morality* (trans. Robert V. Stone [Bloomington, IN: Indiana University Press, 1980]), as an accurate elaboration of the existentialist ethics implicit in his own works), as well as the significance of the posthumously published or yet unpublished works on or related to ethics (for instance, *Cahiers pour une Morale* [Paris: Gallimard, 1983]; translated as *Notebooks for an Ethics*, trans. David Pellauer [Chicago: University of Chicago Press, 1992]. Hereafter: CM). For conventions concerning the referencing system, see note 11.

2. As we will see in the next section, Sartre's account is read as radical from the two main perspectives freedom is usually discussed.
3. According to Stefanie Grüne, for instance, "for Sartre, in order to maintain that a human being is free it must also be the case that neither his actions nor his choices are causally necessitated by anything" (Stefanie Grüne, "Sartre on Mistaken Sincerity," *European Journal of Philosophy* 11, no. 2 [2003]: 151).
4. Grüne attempts to place Sartre in the contemporary context of the debates between libertarianism, determinism, and compatibilism. This is problematic, since Sartre seems to reject explicitly a libertarian account of freedom and even the debate between libertarianism, determinism, and compatibilism, whereas Grüne contrasts compatibilism with Sartre's account. For a similar interpretation of Sartre, as defending a radical libertarian view of freedom, see also Susan Wolfe (*Freedom Within Reason* [Oxford: Oxford University Press, 1990], 65). Elsewhere I show why this is not the appropriate framework for Sartre's account of freedom and I place his discussion in the appropriate conceptual context (Sorin Baiasu, "Existentialist Freedom, Distorted Normativity, and Emancipation," *Oñati Social-legal Series* 5, no. 3 [2015]: 874–94).
5. This view of autonomy is present also in Christine Korsgaard's account of the sources of normativity and is considered an existentialist aspect of her work. This is well illustrated in the discussions of her position by Gerald Cohen and Thomas Nagel: "In Christine Korsgaard's ethics, the subject of the law is also its author: and that is the ground of the subject's obligation — that *it* is the author of the law that obliges it" (G. A. Cohen, "Reason, Humanity and the Moral Law," in *The Sources of Normativity*, ed. Onora O'Neill [Cambridge: Cambridge University Press, 2000], 170); "[t]hough she accepts the Kantian argument that freedom implies conformity to law, Korsgaard departs from Kant in holding that the content of the law depends on something else, namely our conception of our practical identity. This distinctly unKantian, rather existentialist idea is the heart of her position. It introduces a strong element of contingency and therefore relativism, because depending on how we conceive of ourselves as reflective beings, the law may be egoistic, nationalistic, truly universal, or just

plain wanton" (Thomas Nagel, "Universality and the Reflective Self," in *The Sources of Normativity*, ed. Onora O'Neill [Cambridge: Cambridge University Press, 2000], 203–4). In her reply, Korsgaard acknowledges the existentialist element (Christine Korsgaard, "Reply," in *The Sources of Normativity*, ed. Onora O'Neill [Cambridge: Cambridge University Press, 2000], 237). One implication of this paper is that, against standard readings of existentialism, Sartre turns out in fact to be closer to Kant than to Korsgaard.

6. This leads to what is sometimes taken to be a paradox of Kantian libertarianism. I discuss this paradox and claim to solve it elsewhere (Sorin Baiasu, "Free Will and Determinism: A Solution to the Kantian Paradox," in *The Concept of Will in Classical German Philosophy*, ed. Manja Kisner and Jörg Nöller [Berlin: De Gruyter, 2020], 7–27).
7. For instance, according to Thomas Anderson, "Sartre's views of freedom could lead to a quietistic or Stoical ethics. If human reality is freedom and human freedom is total, absolute, and unlimited, if all situations are equivalent in freedom, then there is no reason to change the concrete conditions in which humans live, even if they appear terribly oppressive" (Thomas C. Anderson, *Sartre's Two Ethics: From Authenticity to Integral Humanity* [Chicago: Open Court, 1993], 25). Associated with this implausible view of negative freedom there seems to be also an implausible notion of responsibility: "What we know as the world is the conglomerate of human projects; there 'is no nonhuman situation.' For example, there are no innocent victims of war. Any war we are in is one we deserve. Human consciousness causes the world to be as it is, and so it is entirely responsible for the world" (Jeanson, *Sartre and the Problem of Morality*, 234–35).
8. Charles Taylor criticizes this aspect of existentialism (Charles Taylor, "What's Wrong with Negative Liberty," in *Philosophy and the Human Sciences*, vol. 2 [Cambridge: Cambridge University Press, 1985]).
9. Sarah Richmond's recent translation renders "*motif*" by "reason," which I think it is much less confusing: Jean-Paul Sartre, *Being and Nothingness: An Essay on Phenomenological Ontology*, trans. Sarah Richmond (London: Routledge, 2018).
10. I have discussed the first charge elsewhere ("Existentialist Freedom, Distorted Normativity, and Emancipation"). I must add that, for the purpose of the comparative discussion of Kant and Sartre, I draw on aspects of a more detailed analysis previously published (Sorin Baiasu, *Kant and Sartre: Re-discovering Critical Ethics* [Basingstoke: Palgrave Macmillan, 2011]).
11. Jean-Paul Sartre, *L'Être et le néant: Essaie d'ontologie phénomènologique* (Paris: Gallimard, 1973); translated as *Being and Nothingness: An Essay on Phenomenological Ontology*, trans. Hazel E. Barnes (New York: Philosophical Library, 1956). Hereafter: EN. For Sartre, abbreviations will be followed by page number of the French edition used. Then, for both Kant and Sartre, the page for the English translations will be given. Emphases are in the original.
12. For a more detailed discussion of this aspect of Sartre's existentialism, see Baiasu, *Kant and Sartre*, esp. ch. 4.

13. Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Werner S. Pluhar (Indianapolis: Hackett, 1996). References to the *Critique of Pure Reason* will follow the A (first edition), B (second edition) convention. In references, abbreviations will be followed by the volume and page number from Kant's *Gesammelte Schriften* (Berlin: vols. 1–22 Preussische Akademie der Wissenschaften; vol. 23 Deutsche Akademie der Wissenschaften zu Belin; vols. 24—Akademie der Wissenschaften zu Göttingen).
14. This represents Part II of the *Critique of Pure Reason* with Part I being devoted to Transcendental Aesthetic.
15. One context for Kant's discussion here is given by scepticism; ancient radical sceptics in the Pyrrhonian tradition used to claim that, for every dogmatic claim, they can defend, with an equally strong argument, the contradictory claim. For them, the result of reflecting on this antinomy was not, as it was for Kant, a scandal; on the contrary, they thought that, in this way, they can suspend judgment [*epoché*] and achieve mental tranquillity or unperturbedness [*ataraxia*], both of which refer to their ideal of the good life (Sextus Empiricus, *Outlines of Pyrrhonism*, trans. Benson Mates [New York: Oxford University Press, 1996], Book 1, §4). It is in this practical sense that Kant's suggestion of a scandal generated by antinomies seems exaggerated and only correct in part.
16. "Now … from this second standpoint, I cannot … cognize freedom as the property of a being to which I attribute effects in the world of sense…. Nevertheless, I can still *think* freedom. I.e., at least my presentation of freedom contains no contradiction, if we make our critical distinction between the two ways of presenting (sensible and intellectual)" (Bxxviii).
17. See quotation in the previous note.
18. In the second *Critique*, Kant does see freedom as a practical postulate (4:133). Moreover, he thinks that the postulates of practical reason give objective reality to the ideas of speculative reason, so to freedom too. However, Kant specifies that the postulate is of freedom "considered positively" and that the objective reality of the ideas does not "expand theoretical cognition" (4:133). This explains why in the first *Critique* Kant claims that we cannot have theoretical cognition of freedom considered negatively, but can only show that it is not a contradictory idea. On a more detailed discussion of these aspects, see my "Kant's *Rechtferigung* and the Epistemic Character of Practical Justification," in *Kant on Practical Justification: Interpretive Essays*, ed. Mark Timmons and Sorin Baiasu (New York: Oxford University Press, 2013), esp. Section III.
19. This is a reading that can be found in Otfried Höffe's "Kants kategorischer Imperativ als Kriterium des Sittlichen," in *Zeitschrift für philosophische Forschung* 31, no. 3 (1977): 354–84.
20. I follow here the increasingly common convention of capitalizing this expression, in order to distinguish between the moral criterion on the basis of which maxims are evaluated and the objective moral principles (categorical imperatives, such as "Do not lie!"), which are the results of such an evaluation.

21. For instance, in the *Critique of Dialectical Reason* (Jean-Paul Sartre, *Critique de la raison dialectique* (précédé de *Questions de méthode*). Tome 1: Théorie des ensembles pratiques (Paris: Gallimard, 1960); translated as *Critique of Dialectical Reason*, Vol. 1: Theory of Practical Ensembles, trans. Jonathan Rée, ed. Alan Sheridan-Smith (London: Verso, 2004), 285n1, 227n68.
22. The distinction is in fact fourfold: Apart from the for-itself and the in-itself, significant are also the concepts of the for-others and the for-itself-in-itself. For the purpose of this chapter, I only need to briefly mention the distinction between the first two terms, and between the first two and the last.
23. The object is transcendent, because it is not part of my consciousness. It is an object of my consciousness, beyond it and, hence, transcends it.
24. Moreover, there seems to be an even higher-order rule, which represents the sum of the person's ethical attitudes and which Sartre calls the global project and Kant, the disposition. A discussion of these goes beyond the scope of this chapter, but a comparative analysis can be found in Baiasu, *Kant and Sartre*.
25. This is noticeable particularly in the unpublished notes for the canceled Cornell Lectures. Discussions of these can be found in Juliette Simont, "Autour des conferences de Sartre à Cornell," in *Sur les Écrits Posthumes de Sartre; Annales de l'Institut de Philosophie et de Sciences Morales*, ed. Pierre Verstraeten (Bruxelles: Éditions de l'Université de Bruxelles, 1987); and Pierre Verstraeten, "Imperatifs et Valeurs," *Sur les Écrits Posthume de Sartre: Annales de l'Institut de Philosophie et de Sciences Morales*, ed. Pierre Verstraeten (Bruxelles: Éditions de l'Université de Bruxelles, 1987).
26. See my discussion of this topic in Baiasu, *Kant and Sartre*, esp. §80.
27. I started to work on some of the ideas of this paper when I was a visiting researcher at the University of Vienna, as part of the ERC Advanced Grant "Distortions of Normativity." I would like to thank the project's PI for the opportunity and Keele University for a period of research leave which enabled me to do research abroad.