



HAL
open science

Peculiar access: Sartre, self-knowledge, and the question of the irreducibility of the first-person perspective

Pierre-Jean Renaudie, Jack Reynolds

► **To cite this version:**

Pierre-Jean Renaudie, Jack Reynolds. Peculiar access: Sartre, self-knowledge, and the question of the irreducibility of the first-person perspective. Talia Morag. Sartre and Analytic Philosophy, Routledge, 2023, 9781138316058. <10.4324/9780429455919-5>. <hal-04576857>

HAL Id: hal-04576857

<https://hal.science/hal-04576857v1>

Submitted on 15 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Peculiar access: Sartre, self-knowledge, and the question of the irreducibility of the first-person perspective

Pierre-Jean Renaudie and Jack Reynolds

Abstract:

In the debates on phenomenal consciousness that occurred over the last 20 years, Sartre's analysis of pre-reflective consciousness has often been quoted in defense of a distinction between first and third-personal modes of givenness that naturalists reject. This distinction aims both at determining the specificity of the access one has to their own thoughts, beliefs, intentions or desires, and at justifying the particular privilege that one enjoys while making epistemic claims about their own mental states. This chapter defends an interpretation of Sartre's theory of pre-reflective consciousness that does not put him completely at odds with the naturalist critique and stresses that one's pre-reflective experience should not be understood in terms of first/third-person differentiation. This pre-reflective dimension of experience nevertheless allows a peculiar access to oneself, which grants no infallible epistemic privilege to the first-person, but lays the ground for a transformative process that makes self-knowledge possible.

It is sometimes said that philosophy begins with the Socratic injunction to “know thyself”. Whether contemporary philosophers live up to this vocation might be questioned, but the recent literature nonetheless provides useful resources to determine whether that (alleged) self-knowledge is metaphysically and epistemically different from the knowledge that we might have of other people, or of the external world. On these questions the philosophical landscape is divided, broadly between naturalist and non-naturalist perspectives. The former are in the ascendancy, notwithstanding that non-naturalist conceptions of self-knowledge are in greater accord with many of our basic pre-philosophical intuitions, which tend to assume some sort of special or unique first-personal access to our own thoughts, desires, and intentions. Phenomenological approaches to this question are caught betwixt and between here – neo-behaviorist in certain ways as we will see, but also Cartesian in others. Our engagement with the work of Jean-Paul Sartre in this chapter will pivot

around these differences, steering a middle way between naturalist positions that deny the epistemic and broader philosophical significance of any first/third-person distinction, and some other phenomenological work that deploy Sartrean ideas and insights to inflate this distinction, notably the influential writings of Dan Zahavi on “for-me-ness”, “mineness”, and first-personal givenness. While we agree with Zahavi and others that pre-reflective experience is indexed to an embodied perspective and (implicitly) to a ‘here’ and ‘now’, we argue that recognizing the non-egological anonymity of this perspective (and the agency involved), which is prior to any first and third-person differentiation that depends on language and reflection, is preferable to the sort of ‘mineness’ that Zahavi emphasises,¹ and which appears to constitute something like the ontological or metaphysical basis for privileging self-knowledge, when undertaken with the requisite phenomenological pre-cautions and training.² Through our own reading of Sartre’s work, we outline a position that can endorse the naturalist’s claim that we don’t have *privileged* self-knowledge in regard to our own mental states, but without accepting the too quickly conjoined view that the first-person perspective might be reduced to the third-person perspective, or that it is only the latter which has epistemic and metaphysical significance. In short, we argue that there is an irreducibility of pre-reflective consciousness to third-personal accounts in which any given perspectivality is rendered irrelevant (contra naturalism), and hence that there is an important asymmetry between self-knowledge and knowledge of others in terms of peculiarity of access³, but this does not entail epistemic privilege (contra non-naturalism).

1. Naturalism, self-knowledge and the first-person perspective

Accepting a naturalistic world-view constrains how we understand the Socratic demand to “know thyself”. For example, naturalist positions tend to give negative answers concerning both the metaphysical status of the self (which is more like a convenient fiction or a product of grammar), as well as any knowledge we might claim to possess regarding our ‘selves’, whether they are ultimately illusory or not in a metaphysical sense. This is because naturalism calls for continuity with the findings of science and requires the methodological exclusion of modes of knowing that are akin to “papal infallibility.”⁴ While this is undoubtedly a little vague, in practice it reflects the idea that the findings and methods of science purport to be objective and third-personal, reproducible by any other scientist.⁵ Any particular subjectivity may be relevant to the context of hypothesis formation, but is irrelevant to the context of discovery and the results produced. As a consequence, orthodox forms of naturalism deny that the distinction between the first and third-person perspectives is either methodologically or metaphysically deep. Indeed, such a position appears entailed by any allegiance to strong or exclusive renderings of *both* ontological and methodological naturalism. For naturalists who subscribe to a metaphysical thesis like Wilfrid Sellars’ *Scientia Mensura* principle (e.g. “science is the measure of all things, of what is that it is, and of what is not that it is not”),⁶ and who also endorse a methodological thesis in which any epistemic inquiry ought to emulate the methods of inquiry of certain privileged natural sciences, then any self-knowledge worthy of the name must be accrued in the same manner as other forms of knowledge, notably scientific knowledge, which is held to be paradigmatically third-personal, notwithstanding the importance of observation and experiment.⁷ Any claims concerning the special intimacy, immediacy, or epistemic authority about ourselves

will appear suspect, whether gleaned through first-personal introspection, inner perception, or that somehow accompany lived-experience.

For the naturalist, then, to know thyself is not to turn inwards, nor is it something fleetingly grasped in a pre-reflective and inchoate way through our embodied experience of a world in action, or when those “average, everyday” projects are problematized, as Karl Jaspers, Martin Heidegger, or any neo-existentialist might hold.⁸ Rather, to know ourselves we must adopt something like what Dennett calls the “intentional stance” in regard to oneself, which treats oneself as if one were another person.⁹ It is not the first-person mode of access to our experience that matters, since this is liable to confabulation and impromptu theorizing.¹⁰ Instead, what Dennett calls a third-personal “hetero-phenomenology” is required: a spectatorial stance towards testimonies and alleged first-personal knowledge and lived experiences, whether our own or others, in both cases attributing beliefs and desires in order to predict and explain behavior.¹¹ Dennett invokes the term phenomenology here because his method requires us to neutrally bracket any subjective commitment to the truth of our beliefs (in this case beliefs about ourselves), and we instead quasi-inductively catalogue reports of such beliefs, and then treat them objectively from a perspective external to them: that is, spectatorially and quasi-scientifically.¹² Even with non-reductive physicalism, which generally denies any stronger physicalist aim to reduce psychology to biology (and ultimately to physics), little credence is given to evidences derived from the first-person perspective.

Two key philosophical moves are made here. On the one hand, the first-person perspective is treated as epistemically suspect. On the other hand, it is held that even

if the first-person perspective is epistemically less reliable, in the end there is also no important asymmetry, or fundamental difference in kind, between the first and third-person perspectives. While ‘liberal’ and non-reductive naturalists contest this,¹³ the burden of proof is on those who think the first-person perspective is philosophically significant to establish that, in non-reducible fashion, and without buying into any version of first philosophy wherein we know the structures and contents of our mind in apodictic fashion and from the armchair.

It is worth noting, however, that some famous arguments against physicalism (and scientific naturalism) do invoke the first-person perspective, and also that they often transgress some central naturalist scruples. This is a vast literature that we cannot address, with Thomas Nagel and David Chalmers’ work being notable, but it is worth considering briefly Frank Jackson’s so-called “knowledge argument” regarding “black and white Mary”, and her subsequent experience of the color red.¹⁴ The central claim that Jackson derives from this famous thought experiment is that there is some kind of phenomenal experience, i.e. qualia, or at least a distinctive kind of “know-how”, when Mary suddenly experiences red for the first time, having previously been unable to have such an experience due to being either color-blind or rather cruelly stuck in a black and white room. And, of course, Mary is also a neuro-scientist and expert concerning all of the physical properties of color and their biological conditions. According to Jackson, however, there remains a first-personal experience, and/or a type of know-how, when she first perceives red that cannot be captured purely third-personally or with a catalogue of all facts and scientific (e.g. third-personal) knowledge. The “know-how” interpretation of Jackson’s thought experiment is of special relevance to us here,¹⁵ since the self-knowledge we will

discuss is akin to a kind of implicit know-how, a practical awareness that is neither accessible to introspection nor readily given propositional form. Jackson's famous claim is that the thought experiment establishes that physicalism leaves something out, since the complete physical knowledge that Mary possesses is insufficient.

Such arguments have been criticized from a broadly naturalist perspective that contests the details of the thought experiment and the philosophical force of such "intuition pumps,"¹⁶ and also because it appears to violate some naturalist scruples concerning the causal closure of the universe (that is, it looks dualist). But there are some important Sartrean additions to this story that can be made regarding the relationship between the *lived* and the *known*, the pre-reflective experience of agency that we have and reflective knowledge of that experience and its contents, as well as the complexity that they introduce into any effort at substantive self-knowledge. This can be elaborated without subscribing to a metaphysical argument for epiphenomenalism as with Jackson's own use of "phenomenology" and the first-person perspective, although it must be conceded that like Jackson's Mary argument it is not entirely clear that Sartre's account of a lived and embodied perspective that is "pre-personal" involves a new fact that is in question, nor any strictly propositional account of self-knowledge. The kind of pre-reflective experience that is central to Sartre's phenomenology is not exactly knowledge, but nor is it nothing either, wholly epistemically and methodologically inert. Rather, we are reminded of Wittgenstein's famous remark regarding sensations: "It is not a something, but not a nothing either!"¹⁷. As such, there are valid questions regarding what we can do with this quasi-knowledge without succumbing to a sort of mysterianism about somatic intentionality and lived-experience¹⁸. Many theorists of self-knowledge will continue

to refuse to count it as knowledge. For Boghossian, for example, there are three possibilities for self-knowledge: “such knowledge is either based on inference, or by a kind of looking, or it is nothing.”¹⁹ What are we to say of the kind of indirect knowledge that Sartre says we have, albeit knowledge of a sort that is in the first case mainly negative (we are not x, not y, not *just* this or that Ego or putative identity – e.g. a university professor), and then more existentially lived rather than propositionally known in positive fashion, e.g. Reynolds and Renaudie are this or that kind of parents, this or that kind of philosophers? But let us prepare the ground for this argument by explicating some key Sartrean insights regarding consciousness, self-consciousness and the Ego.

2. Sartre on consciousness, self-consciousness and the Ego

A distinction is usually made between intentional or object-consciousness and self-consciousness, depending on whether consciousness is oriented towards an object or towards itself. In *Being and Nothingness*, however, Sartre famously claimed that these two aspects of consciousness necessarily go together, and this particular claim became the first of the theses Sartre presented to the *Société française de philosophie* in a 1947 conference: “Every positional consciousness of an object is necessarily at the same time a non-positional consciousness of itself.”²⁰

With this claim, Sartre stresses the relation between these two forms of consciousness only as much as he emphasizes the necessity to distinguish them on the basis of the opposition between their positional and non-positional (orthetic and non-thetic) character: self-consciousness is an “immediate and non-cognitive relation to

oneself,”²¹ which does not need to “transcend itself to reach an object” like intentional consciousness does.²² When I perceive the book in front of me, I am not only perceptively aware of a transcendent object, I am also at the same time aware of having a perceptual experience of the book. These two aspects of the perceptual experience are strictly irreducible to one another, and need to be distinguished even though they always come intertwined. Seeing an object and being aware that we are seeing an object belong to two different kinds of experiences, oriented towards the world in the first case, and towards ourselves in the second.

Of course, many naturalists deny this distinction, instead invoking theses concerning the “diaphanousness” of experience. In short, they will deny that we do in fact phenomenologically have this second mode of consciousness (i.e. non-thetic self-awareness of seeing red that is different from the perception of a red object in the world), contending that it is either trivial or generates problems regarding an infinite regress.²³ Defenses of the position also appeal to phenomenology, to memory, and sometimes to empirical studies concerning learning.²⁴ While we are broadly on the phenomenological side of this debate, our aim is not to defend that in detail here. Rather, we seek to show that a proponent of this sort of view (and the peculiar asymmetry between self-knowledge and knowledge of others that results) need not be encumbered by theses regarding epistemic privilege that are suspect for the naturalist. Why? In short, Sartre’s commitment to what he calls the translucency of consciousness, “does not guarantee that something is clearly present in consciousness”, so we can have what Peter Carruthers’ calls the opacity of mind.²⁵ There is no guarantee that “what is in consciousness is correctly conceptualised by

us,”²⁶ since there is a disjunction between the lived and the known in regard to our experiential life.

Self-consciousness, according to Sartre, must not be mistaken with reflexive consciousness: the latter is a form of intentional consciousness that takes one’s own lived-experiences as its specific object, whereas the former is pre-reflexive and need not involve the intentional distance to the object that the act of reflexion entails. Self-consciousness is characterized as an immediate and non-cognitive form of self-awareness, thanks to which one comes to be aware of her own experience without having to posit herself as the object of her reflection. Reflection, on the other hand, is unable to give access to oneself as the *subject* of unreflected consciousness, but only as the intentional *object* of the act of reflection, i.e. the Ego in Sartre’s terminology. The Ego is the specific object that intentional consciousness is directed upon when performing reflection - an object that consciousness “posits and grasps [...] in the same act”,²⁷ and that is constituted in and by the act of reflection.²⁸ The Ego is consequently transcendent and available to the third-person perspective. According to Sartre, and here he remains close to the Husserl of *Logical Investigations*, it “is outside, *in the world*. It is a being of the world, like the Ego of another.”²⁹

This analysis seems to hinge upon a sharp distinction between two irreducible modes of presentation or ‘givenness’, following the **more phenomenological terminology used by Zahavi**,³⁰ intentional in the case of transcendent objects and prereflexive in the case of self-consciousness – a distinction from which the opposition between the first and third-person perspectives can easily be derived. The immediate, self-referential and pre-reflexive character of self-consciousness grants us a specific and

strictly first-personal access to ourselves, a first-person perspective that expresses, as Shoemaker puts it, the “distinctive way mental states present themselves to the subjects whose states they are.”³¹ In contrast, the objective relation to the world that intentional (positional) consciousness achieves, insofar as it does not involve such first-personal mode of givenness, can be characterized as an “impersonal” relation to a transcendent object:

When I run after a streetcar, when I look at the time, when I am absorbed in contemplating a portrait, there is no *I*. [...] In fact I am plunged in the world of objects; it is they which constitute the unity of my consciousness; [...] but *me*, I have disappeared; I have annihilated myself. There is no place for *me* on this level.³²

The streetcar may appear to me *in a specific way* (as “having-to-be-overtaken”, for instance). Nevertheless, it is not presented to me as an object experienced in first-person, but as a transcendent object, experienced in a way that *overrides*, so to say, the subjective aspects of consciousness: the object obliterates the personal features of conscious experience and appears as carrying a set of objective properties strictly independent from my personal experience or subjective relation to the object and fundamentally accessible to a third-person perspective. The “having-to-be-overtakenness” of the tram appears as describing objectively the way the world is and not as an aspect of my subjective experience of the world.³³

This distinction between intentional consciousness and pre-reflexive self-consciousness seems to commit Sartre to a version of the kind of theories of self-

reference that maintain a “fundamental divide” between first and third-person perspectives and constitute the target of Herman Cappelen and Josh Dever’s criticisms in their book, *The Inessential Indexical*.³⁴ Cappelen and Dever call ‘perspectival’ the idea according to which a philosophical distinction needs to be drawn between the description of the world as it objectively is and as it is subjectively experienced. According to them, this distinction is “philosophically shallow”³⁵ and unable to ground a significant asymmetry between the first and the third-person perspective. Their book presents a series of critical arguments that aim to show how the first-person perspective can always be explained in terms of – and subsequently reduced to – the third-person viewpoint.

Although Sartre is not Cappelen and Dever’s explicit target, his key distinction between (pre-reflexive) self-consciousness and (intentional) consciousness would fall under their criticism and thus, from their perspective, doom his theory of consciousness. Emphasizing this very aspect of Sartre’s phenomenology of consciousness, Dan Zahavi considers it as an account of consciousness in which “the types of self-reference available from a first-person perspective and those available from a third-person perspective are different,”³⁶ akin to the theories of self-reference developed by Castañeda, Perry, Nagel, and Cassam.³⁷ The immediate and pre-reflexive dimension of conscious experience makes one aware of what she experiences in a specific way, which does not involve the level of conceptualization and objectivation that intentional consciousness entails and does not result in a form of reflexive knowledge similar to our intentional knowledge: I experience my feeling of anger in a way that is fundamentally different from the way I experience the object of my anger, the very thing that makes me angry and that my emotional experience is

directed upon. And this remains true even when the object of my anger is myself, if for instance I reflect on an unforgiveable mistake that I made: in this case, my experience of myself as being angry and my relation to myself as the object of my anger do not coincide. Using David Chalmers' vocabulary,³⁸ Zahavi notes that "in first-personal self-reference one is not aware of oneself as an object that happens to be oneself [...]. Rather, first-personal self-reference involves a non-objectifying self-acquaintance."³⁹ Such a view accords privileged access to our own occurrent moods, thoughts, desires, based upon a first-personal acquaintance with oneself that he often refers to as 'mineness' or "for-me-ness" (a 'mineness' which we contend is not phenomenologically basic⁴⁰).

The question, however, is whether the specificity of this first-personal awareness of one's own experiences and its irreducibility to third-personal modes of objectual knowledge implies a form of asymmetry between the first and third-person perspectives that presupposes the superiority or primacy of the former over the latter. Sartre's analysis of consciousness demonstrates, as we shall see, that the irreducibility of the first to the third-person perspective does not necessarily entail that a metaphysically suspicious privilege is granted to the first-person.

The first question that needs to be asked is whether consciousness is personal at all according to Sartre, and in which sense. The entire first part of *Transcendence of the Ego* is dedicated to establishing the impersonal (or "pre-personal") character of consciousness, which stems from its non-egological structure and results directly, according to Sartre, from his demonstration of the absence of the *I* in the transcendental field. When he writes that consciousness is "pre-personal", Sartre

implies that its determination in terms of first or third-person perspectives is meaningless, and that such question can only be asked at a reflective level, i.e. once consciousness has been reflexively tied to an Ego that it becomes essentially related with.

Let's come back to the example of the streetcar. When I run after the streetcar, my consciousness is absorbed in the relation to its intentional object, "*the streetcar-having-to-be-overtaken*", and there is no trace of the I in such lived-experience. I do not need to be aware of my intention to take the streetcar, since the object itself appears as having-to-be-overtaken, and the subjective properties of my experience disappear in the intentional relation to the object. A fundamental characteristic of such intentional experiences, according to Sartre, is that the absence of the experiencing self or Ego makes them look like they were not dependent upon a perspective determined as strictly first-personal. The mode of manifestation that characterises non-reflective states makes them not only non-egological, but subjectless⁴¹, insofar as they are lived-through without any reference whatsoever to the experiencing subject (or to the fact that this experience has to be experienced *by someone*): they are so "translucent"⁴² that the subjective or first-personal dimension of the experiencing vanishes through the appearing of the object. When I perceive Pierre as loathsome, to take Sartre's example, I do not perceive my feeling of hatred; Pierre repulses me, and I experience *him* as repulsive.⁴³ Repulsiveness constitutes primarily an essential feature of *his* distinctive mode of appearing (or givenness), rather than a trait of my feelings towards him. Intentional consciousness does not only posit and objectify its intentional correlate, it does so by dissolving the subjective and first-personal aspects of the experience of the object, making it impersonal. The object then appears as

carrying in itself the determinations that I experience ‘on’ it, as with Pierre’s repulsiveness. The specific way that he appears to me is given *as* impersonal, insofar as it does not appear to me as *my* specific feeling about Pierre – a feeling of repulsiveness that should characterize my experience rather than him. My experience of the object is not primarily *mine* but is presented as objective and universalisable, so that anyone who knows Pierre could (and maybe should) see his repulsiveness just as much as I do.

The fact that the experience of Pierre’s repulsiveness is *my* experience is not relevant for the description of the experience I am having (as naturalists like Dennett and Garfield 2016 would agree): the experience is given in a way that presents it as identical to anyone else’s experience of the same object, and likely to be shared by anyone. At this level, conscious experience can paradoxically only be described as first-personal *from an external point of view* (from the point of view of someone who is not currently experiencing it, or from a philosophical perspective that does not coincide with the lived experience exactly). From the point of view of the experiencing subject, it is irrelevant to characterize this experience as “his” rather than as someone else’s.

This is why Sartre claims in *The Transcendence of the Ego* that I-statements (reflective) cannot be derived from lived-experiences (non-reflective) without raising some doubts regarding their epistemic validity:

Thus to say ‘I hate’ or ‘I love’ on the occasion of a singular consciousness of attraction or repulsion, is to carry out a veritable

passage to the infinite [...] Nothing more is needed for the rights of reflection to be singularly limited: it is certain that Pierre repulses me, yet it is and will remain forever doubtful that I hate him. Indeed, this affirmation infinitely exceeds the power of reflection.⁴⁴

The reflective switch from the experience of a certain feeling to its expression in an I-statement enables one to adopt either a first or a third-person perspective on herself, while this possibility to switch from a perspective to the other was not yet available at the non-reflective level that Sartre's phenomenology describes. The robust distinction between first and third-person perspectives that constitutes 'perspectivity' according to Cappelen and Dever's can only make sense *after* reflection has taken over one's conscious life and attached some egological content to their experiences. But then, the experience reflected upon cannot be identical with the consciousness that is reflecting: a "nothingness" slips between them, as Sartre famously argues, which enables bad faith in our own self-relation. In contrast, pre-reflective awareness in Sartre describes a primary level⁴⁵ of experience that does not provide a personal perspective at all, be it either first or third-personal. Describing one's experience as subjective or first-personal can only be relevant in retrospect (through reflection), or from the external viewpoint of another. For Sartre, then, saying that every conscious state is non-positionally and non-reflectively directed upon itself is another way of saying that pre-reflective self-consciousness is not tantamount to a consciousness *of one's self*. The distinction between first and third-person perspectives is unable to capture the *impersonal* mode of givenness that characterizes pre-reflective consciousness⁴⁶.

3. Sartre on self-awareness, self-knowledge, and bad faith

In regard to any substantive self-knowledge that might be achieved through direct self-consciousness, then, our options are limited. We cannot look inwards and discover the truth through introspection (there is literally no-thing to observe). Moreover, when we have a lived experience, and then reflect on ourselves from outside (e.g. third-personally, as an Ego), we are not strictly reducible to that Ego that is so posited. We transcend it. Or better, we both are that Ego (just as we are what the Other perceives) and yet are also not reducible to it, partly for temporal reasons, since we are not just our past and our objective attributes (in accord with some sort of principle of identity) because we are also our future projects, albeit in the mode of not being them, as Sartre puts it. This issue of irreducibility is important here, involving something akin to an ontological recognition. This irreducibility expresses, as it were, our mode of existing: we exist in the mode of a non-coincidence to our own Ego, so that we fail to know ourselves every time we simply identify with the Ego that we are. Self-reflection is necessarily delusive, insofar as it posits and grasps the Ego in the same act, and claims that it discovers the very object that it constitutes.⁴⁷

For now, however, it is worth noting that the kind of epistemic considerations that motivate Dennett's third-personal or hetero-phenomenology appear to involve a related split concerning the subject of cognition, a kind of epistemic "bad faith", a "heavy-handed sleight of hand" as Hubert Dreyfus and Sean Kelly put it in a paper of that name.⁴⁸ Recall that Dennett and many other scientific naturalists demand that we bracket away any subjective commitments or beliefs that we may have – hence Dennett's idiosyncratic reappropriation of the idea of phenomenology – but it is not to do auto-phenomenology and better attend to the things as they present themselves, but to enable a third-personal hetero-phenomenology in which we adopt the "intentional

stance” in regard to oneself and in essence treat oneself as if one were another person, a stranger even.⁴⁹ One’s experiences are given no special credulity here. In fact, to be analysed and admissible in any science of cognition, they must be framed strictly third-personally. For Dennett et al, we hence treat ourselves as an object or Id-like (to recall Sartre on Freud), even while we are conscious of this idealization and thus, on Sartre’s account, partly distanced from it.

While bad faith is inevitable in Sartre’s view, it is also important to recognize that the germ of its destruction lies within, and this begins to hint at the way in which a pre-personal perspective plays an ongoing and ineliminable role within his philosophy. Although bad faith is pervasive and able to saturate our lives, there is also a kind of optimism about Sartre’s account, in that he describes it as “metastable”, meaning liable to change.⁵⁰ This is because bad faith always remains at least partly available to us in our own lived-experience, albeit not in a manner that might be given propositional form in the same way as knowledge of an external object. In short, when I existentially comprehend that my life is dissatisfying, say, or even reflect on this basis that I have lived an inauthentic life, while I am grasping something about myself (since qua lived experience it is given differently to the recognition that others have lived a lie: it is more likely to induce anxiety), I am nonetheless not strictly equivalent or identical with the ‘I’ that is claimed to be in bad faith.⁵¹ There is a distanciation involved in coming to this recognition and the potential for self-transformation of a more practical kind, even if this is under-thematized in *Being and Nothingness*⁵². To put the point slightly differently, there is a gap here between the modalities of the “I think” and the “I can” to borrow from Merleau-Ponty and Husserl’s related discussions, or the “I think” and the “I am”, as with Sartre’s

discussion of Descartes in *Transcendence of the Ego*. And this lived 'I can' is at bottom impersonal and anonymous, albeit indexed to a given body-world relationship (it is not a view from nowhere, a God's eye view: its objectivity is given differently from that).

Of course, in the case of bad faith it is not the bad faith of just anyone at stake: rather, it is an obscure intimation (if not a direct thetic recognition) of my own bad faith and it is given at the embodied and pre-reflective level, involving an indirect self-awareness that is not equivalent to that which we have towards objects in the world or other people. What good faith would be is much less clear and much less explicated in *Being and Nothingness*, and necessarily so given his account of consciousness and self-consciousness, this negative knowledge remains a form of self-knowledge, albeit of an inchoate and imprecise kind. In addition, while it is not precise knowledge, it is nonetheless *less* liable to confabulation than knowledge of a more reflective variety when we treat the self as an Ego and reflect on ourselves as an object, and also then distinguish between first and third-personal perspectives. We may, of course, dispute Sartre's pessimism about the possibility of adequate reflective knowledge of the 'I', which is both epistemic and existential here, noting perhaps that here "perfect is the enemy of good" or reasonable self-knowledge, but on this point at least he is broadly with the naturalists. Where he differs from them is in maintaining the irreducibility of the perspectives involved and giving lived-experience an ongoing and ineliminable role.

To better see this point of difference, let us consider an example of bad faith in Sartre's work that has been comparatively neglected in the literature (except by

Joseph Catalano), but which is useful for our purposes. We are asked to imagine a situation in which two men are rowing (or hiking, in another related story that Sartre also tells). One is straining to make their efforts appear easy and effortless.⁵³ Such a gesture may have been part of their weekly rowing activities for months or years, and over this period of time it is likely to have ceased being directly given to any conscious intention on their behalf. Instead, it has become more akin to a general manner of comportment in the world. Consider Catalano's characterization of this:

For example, I am living a false idea of manliness that makes me believe that I am rowing easily and that the other is admiring my rowing. As my body moves over time, I think that I experience a smoothness of my muscles that is simply not there. I believe in my ability, just as I believe in the other's admiration. This self-deception is possible because my I is my interiorization of the way I see myself in the world before others. I see a sparkle in the other's eye that I choose to interpret as admiration for my ability at the oars, but which any unbiased observer would recognize as admiration for my effort to impress my companion. To an unbiased observer, my rowing exhibits a tension that is my effort to sustain an effortlessness that is not present. For me this tension is understood as my effort at rowing...⁵⁴

And Catalano concludes the article by referring back to this point and example:

... the effort is going to the creation of an ease-of-rowing that is simply not present. Is this self hidden from this self-deceiver? Yes and no. Yes, in the sense that he manages to conceptualize his tension as the natural effort that appears when rowing vigorously. No, in the sense that he does indeed experience a tension... But this tensed act of rowing is the self-deceiver; it is a self maintaining itself in a self-deception as it keeps from itself the possibility of change.⁵⁵

This example is interesting for various reasons. First, Sartre's view appears neo-behaviorist in a way that even seems closely related to Dennett. The rower's bad faith and lived-tension is nothing other than its expression, in the world. It is in principle available to other perceivers, evident in behavior and comportment. With the right context, we can perceive the bad faith of another in their comportment⁵⁶. Sartre's analyses also retain a role for peculiarity of access to that experience, and it is this which makes bad faith internally unstable and liable to change. But it is not that there are two objects here: the rower engaging in the act of rowing and whatever they are deliberately attending to, plus consciousness of their own act of bad faith as an object that they might also have direct knowledge of. Rather, the activity of rowing is accompanied by an implicit awareness of one's agency, but not necessarily judgments or beliefs about agency (it is more like a motor awareness rather than reflective awareness). And as with other types of know-how, it is not readily introspectible or easily made into an object of knowledge. As Ryle, Dewey, Merleau-Ponty, and many others have noted, we might know-how to ride a bike to get to work, but nonetheless find it very difficult to give appropriate third-personal directions a stranger or novice

could follow⁵⁷. Sometimes it is difficult to frame our motor-intentionality in any precise way, to give “what” and “why” answers about what one is up to.⁵⁸ In this case of bad faith, which is a kind of skill and an implicit way of dealing with the world, there might also be a tacit self-consciousness that accompanies the activity that is not about knowing one’s self qua object. It may include a peculiar lived inkling that our ratiocinations and self-conceptions are, at best, approximate, and there is a negative form of self-knowledge and awareness available in this admittedly imprecise way.

In this scenario, we have a peculiar access to bad faith in our own experiences, in contradistinction to the experience of others. Moreover, the pre-reflective experience of ourselves and the view which the other has of our rowing are irreducible to each other, without either having a general epistemic privilege. Although in the example of the rower there is a sense in which the bad faith is apparent externally and to an unbiased observer with sufficient context, it is also internally available, but not in introspection or reflection upon the self, but in the lived action. Contrary to any picture in which reflective self-knowledge is privileged in such cases, we are more liable to believe our own bad faith and self-narration that our rowing is easy and effortless than any third-personal observer is. In at least these kinds of cases, then, the epistemic privilege may well be with others, including in direct perception of others, for example in the anger of the other that is immediately and non-inferentially given,⁵⁹ but also in intentional and complex behavior like rowing while attempting (in some pre-conscious and non-thetic way) to make it look easy. With the naturalist, then, Sartre denies claims that we can adequately know ourselves as objects in any exhaustive way through self-reflection or any introspective self-seeing, but that is not the end of the matter, since there is an indirect self-consciousness and lived awareness

that persists. These pre-reflective experiences retain a kind of irreducibility and they have a peculiar but not epistemically privileged mode of access to them qua lived rather than known.

Sartre hence complicates any model in which we have direct access to our own mental states in self-knowledge and indirect or hypothetical access to those of others. We have direct access to at least some of the emotions and intentions of others. And our relationship to our own mental states is highly complex and indeed liable to confabulation and bad faith, but this does not motivate the kind of reaction to Descartes that one might find in Ryle and Dennett. Consider for example, the following remark from *Transcendence of the Ego*:

The ego gives itself as an object. Therefore the only method for coming to know it is observation, approximation, expectation, and experience. Yet these procedures, which are perfectly adequate to every non-intimate transcendent, are not adequate here, because of the very intimacy of the ego.⁶⁰

Here we have a pithy statement that encapsulates the paradox. According to Sartre, the ego is a transcendent and worldly thing, not something we might peer inwards and have knowledge of internally. And yet the 'I' is not more certain, but more "intimate", which looks like a thesis of peculiarity of access, of asymmetrical access. While the ego is transcendent it is also *not* like other transcendent objects, precisely because intimate. And this means that knowledge of the sort that we enjoy of other objects (e.g. in science and ordinary inductive reasoning) will not adequately grasp it. There

are particular difficulties involved in self-knowledge, especially for anything more complex than the occurrent recognition that I know that I believe I am in pain. After all, “reflecting on our psychology can change the objects of this reflection,”⁶¹ and, as such, there is no self-knowledge that might be final or wholly adequate. We are, rather, our projects and Sartre’s existential and moral aim, of course, is to distinguish this open-ended project of self-creation from complete bad faith, in which one refuses any ontic or empirical characterization of who and indeed what one is.

For Sartre, this complexity problematises views of self-knowledge that treat it as equivalent with third-personal knowledge, rather than testifies to them, even if both parties agree that bad faith and confabulation are endemic when it comes to ascribing characteristics to our own Ego, and even if both agree that we cannot look inwards to discover who we are. On the Ryle and Dennett story that we have glossed, the methods for finding out about ourselves are the same as those we employ in regard to others. And while we can use such hetero-phenomenological methods in regards to ourselves as well as in regard to other people (hypothesis, inference, etc.), the methods cannot be precisely the same without something akin to the “heavy handed sleight of hand” discussed above. When it comes to any quasi-scientific theory about ourselves the difficulties are redoubled; we are especially prone to epistemic failure in our own cases. The methods hence must either be different, or at least practiced with more expertise and adroitness than any kind of ordinary knowledge might have it. If there were no difference of this nature, then there would be no problem for psychoanalysis, behavioural psychology, existential psychoanalysis, etc., to be practiced in regard to oneself. But there are particular difficulties in the case of oneself that are widely theoretically and empirically attested to.⁶² Without an account

of the significance but not privilege of perspectival in self-knowledge and action, we seem to have no accounting for akrasia. We make mistakes in interpreting others, of course, and we misunderstand motives and intentions of our colleagues and friends, but it is rarely willfully blind or “bad” in the same way as our efforts at self-knowledge can be. Sartre’s account of consciousness, self-consciousness, and the Ego, along with the irreducibility of the lived and the known, help us to understand how this complexity and bad faith may be pervasive in our lives (while also being liable to change).

Conclusion

Sartre’s account of the relations between self-consciousness and reflective consciousness develops an original analysis of the asymmetry between the first and third-person perspectives. He stresses the ontological significance and the irreducibility of our pre-reflective experience, a kind of experience that is implicitly indexed to a ‘here’ and ‘now’, but which nevertheless cannot be described as first-personal strictly speaking, insofar as it does not manifest one’s self and so does not constitute one’s perspective as one’s own. This lived-experience does not yet constitute ourselves *as selves*, and this is why it is best described as prepersonal rather than first-personal, in spite of the perspectival character of pre-reflective experience.⁶³ Consequently, the irreducibility of the specific kind of self-consciousness that unreflected experiences involves does not imply some kind of epistemic privilege to be granted to the first-person. Being perspectival but not first-personal, pre-reflective awareness is unable to provide us with any stable propositional knowledge of ourselves. Only reflection can bring this primary modality of awareness to a fully-

fledged form of self-knowledge, as it allows some distance and opens up a new perspective on oneself that is akin to the others' point of view and can be described as third-personal. However, this third-person perspective on oneself is just as ambiguous as the perspective that one experiences through pre-reflective consciousness: it deserves the title of "knowledge", as it makes oneself the object of a reflective act, but it misses the specific kind of intimate connection with oneself that pre-reflective experience allows. The description that Sartre provides of this asymmetry between first and third personal knowledge, in other words, is first and foremost a description of the limits of our self-knowledge, which exhibits the necessary failure of the one-sided conceptions of self-knowledge that ground it on one perspective or the other. The first-person perspective is just as insufficient as the third-person one, although their limitations are not the same, which is the reason why the irreducibility of these two perspectives must be maintained and is philosophically significant.

However, if the asymmetry between the first and the third-person perspectives reveals the limits of the forms of self-knowledge that they make possible, it does not rule out the very possibility of any kind of knowledge of oneself, which may be not about an object in form. While it might be disputed that this counts as knowledge, there is something more akin to an embodied "know-how" involved here, rather than knowledge-that, which is often very difficult to articulate as propositional knowledge. The peculiar access to ourselves that pre-reflective experience lets us enjoy might not be able to provide us with epistemic guarantees or certainty about ourselves, but it relies on a kind of immediate self-awareness that lays the ground for self-knowledge and makes it at least possible. If it does not provide a proper knowledge of ourselves, it is not non-knowledge either, to recall the enigmatic remarks from Wittgenstein that

we considered earlier. Our pre-reflective experience allows us to enjoy (and sometimes be saddened or haunted by) a peculiar access to ourselves, and it is this which makes practices of bad faith potentially liable to transformation. However, this access is not such that it could grant an *epistemic* privilege to the first-person or provide some grounds for first-person authority. Neither the first nor the third-person description of oneself is sufficient to fulfill the expectations and promises of self-knowledge, but their asymmetry opens the possibility of a *negative* form of self-knowledge, which comes to be accomplished within our failed attempts to know ourselves and through the negative experience of the insuperable limits of both perspectives.

References:

- Boghossian, P. Content and Self-knowledge. *Philosophical Topics*, 17(1), (1989). 5-26.
- Bortolotti, L. *An Introduction to the Philosophy of Science*. New York: Wiley. (2008).
- Brownstein, M. Rationalizing Flow: Agency in Skilled Unreflective Action. *Philosophical Studies*, 168(2), (2014). 545-568.
- Byrne, A. Introspection. *Philosophical Topics*, 33(1), (2005). 79- 104.
- . *Transparency and self-knowledge*. New York, NY: Oxford University Press. (2018).
- Cappelen, H. and J. Dever. *The Inessential Indexical*. Oxford: Oxford University Press. (2013).
- Carruthers, P. *The Opacity of Mind*. Oxford: Oxford University Press. (2011).
- Cassam, Q. *Self and World*. Oxford: Clarendon Press. (1997).
- Castañeda, H-N. “He”: A study in the logic of self-consciousness. *Ratio*, 8, (1966). 130- 57.
- . *The phenomeno-logic of the I: essays on self-consciousness*. Edited by T. Kapitan et J. G. Hart. Bloomington: Indiana University Press. (1999).

- Catalano, J. Successfully Lying to Oneself. *Philosophy and Phenomenological Research*, 50(4), (1990). 673-693.
- Cath, Y. The Ability Hypothesis and the New Knowledge-How. *Noûs*, 43(1), (2009). 137–56.
- Knowing How without Knowing That. In J. Bengson and M. A. Moffett (eds.) *Knowing How: Essays on Knowledge, Mind, and Action*, pp. 113–35. Oxford: Oxford University Press. (2012).
- De Caro, M. and Macarthur, D. (eds.) *Naturalism and Normativity*. Harvard University Press. (2011).
- Dennett, D. *Consciousness Explained*. New York: Basic Books. (1991).
- Dennett, D. Heterophenomenology Reconsidered. *Phenomenology and the Cognitive Sciences*, 6 (1-2), (2007).
- Dreyfus, H. & S. Kelly. Heavy-handed sleight of hand. *Phenomenology and Cognitive sciences*, 6 (1–2), (2007). 45–55.
- Garfield, J. *Engaging Buddhism*. Oxford: Oxford University Press. (2016).
- Gertler, B. *Self-knowledge*. New York: Routledge. (2011).
- Guillot, M. I me mine: On a confusion concerning the subjective character of experience. *Review of Philosophy and Psychology*, 1, (2017). 23- 53.
- Jackson, F. Epiphenomenal Qualia. *Philosophical Quarterly*, 32, (1982).127–136.
- Kapitan, T. Is subjectivity first-personal? In S. Miguens Travis, C. Morando Bravo, G. Preyer (eds.), *Prereflective Consciousness: Early Sartre in the Context of Contemporary Philosophy of Mind*. London: Routledge. (2015).
- Moran, R. *Authority and Estrangement*. Cambridge, MA: Harvard University Press. (2001).
- Nagel, T. What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), (1974). 435.
- Perry, J. The Problem of the Essential Indexical. *Noûs*, 13(1), (1979). 3- 21.
- Richmond, S. “Introduction”, *The Transcendence of the Ego*, Trans. A. Brown, London: Routledge, 2004.
- Reynolds, J. and P. Stokes. “Writing the First-person”. In G. D’oro and S. Overgaard (eds.), *Cambridge Companion to Philosophical Methodology*. Cambridge: Cambridge University Press. (2017).
- Ryle, G. *The Concept of Mind*. London: Routledge. (2009).

Sartre, J-P. *Transcendence of the Ego*. Trans. F. Williams and R. Kirkpatrick. New York: The Noonday Press. ([1936] 1957).

----- *The Transcendence of the Ego*, Trans. A. Brown, London: Routledge. ([1936] 2004)

----- *Being and Nothingness*. Trans. H. Barnes. London: Routledge and Kegan Paul. ([1943] 1958).

----- *La Transcendance de l'Ego et autres textes phénoménologique*, Paris: Vrin. ([1936] 2003).

Sellars, W. Empiricism and the Philosophy of Mind. In *Science, Perception and Reality*. Routledge: London. (1963).

Shoemaker. S. *The first-person perspective and other essays*. Cambridge: Cambridge University Press. (2010).

Thompson. E. Self-No-Self? Memory and reflexive awareness. In M. Siderits, E. Thompson, D. Zahavi. (eds.), *Self and No-Self*. Oxford: Oxford University Press. (2010).

Webb. S. *Connaissance de soi et réflexion pratique. Critique des réappropriations analytiques de Sartre*. Sesto San Giovanni: Mimesis. (2022).

Wider, K. A nothing about which something can be said: Sartre and Wittgenstein on the Self. In R. Aronson, A. van den Hoven (eds.), *Sartre Alive*. Detroit: Wayne State University Press. (1991).

Wittgenstein, L.W. *Philosophical Investigations*. London: Blackwell. (1950).

Zahavi, D. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, MA.: MIT Press. (2005).

----- Consciousness and (minimal) selfhood: Getting clearer on for-me-ness and mineness. In U. Kriegel (ed.), *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press. (2019).

¹Zahavi (2005; 2019).

² Such a position appears vulnerable to arguments concerning the reducibility of the first-person perspective to the third in the metaphysical register (e.g. Cappelen and Dever 2013), as well as other arguments that urge a similar move on epistemic grounds and for integration with the relevant cognitive or psychological sciences. And while there is a kind of mineness associated with the pre-reflective experience of bad faith for Sartre, he argues it is an impure form of self-awareness, as we will see.

³ We borrow this term from Alex Byrne though pursue it in very different ways to his work, which remains much more Rylean and inferentialist. For Byrne, self-knowledge must be said peculiar since the method by means of which we come to know our own mental states is a method that only applies to our mental states (not to others'), and that is not available to anyone else, Byrne (2005: 81; 2018: 8).

⁴ Dennett (2007: 252)

⁵ Notwithstanding the so-called "replication crisis". In short, many experiments are not, in fact, able to be replicated or reproduced by subsequent experimenters, especially in social psychology.

⁶ Sellars (1963, 173).

⁷ cf. Dennett (1991).

⁸ cf. Reynolds and Stokes (2017).

⁹ Dennett (1991: 76).

¹⁰ Ibid., 68.

¹¹ Dennett (2007).

¹² Ibid., 252. Dennett's teacher, Gilbert Ryle, made some similar points, famously declaring that: "the sorts of things I can find about myself are the same as the sorts of things that I can find out about others, and the method of finding them out are much the same" Ryle (2008).

¹³ e.g. De Caro and Macarthur (2011).

¹⁴ Jackson (1982).

¹⁵ cf. Cath (2009, 2012).

¹⁶ Dennett (1991).

¹⁷ Wittgenstein (1950: §304). The parallel with Wittgenstein on this point has been pointed out by Kathleen Wider (1991), Joseph Catalano (1990: 675), and more recently by Katherine Morris in various places.

¹⁸ These questions align with those that are posed to agentialist conceptions of self-knowledge, like those of Richard Moran who draws on Sartre (2002). If self-knowledge is some sort of commitment is it right to even call it knowledge, as Brie Gertler (2011) asks?

¹⁹ Boghossian (1989: 5).

²⁰ "Toute conscience positionnelle d'objet est nécessairement conscience non-positionnelle de soi." Sartre (2003, 98).

²¹ (Sartre, EN 19)

²² (Sartre, EN 18)

²³ cf. Garfield (2016).

²⁴ see Zahavi (2019); Thompson (2010).

²⁵ Carruthers (2011).

²⁶ Catalano (1990: 681).

²⁷ (Sartre TE 31)

²⁸ (Sartre TE 80-1)

²⁹ (Sartre TE 80-1)

³⁰ Zahavi (2005: 119).

³¹ Shoemaker (2010: 157).

³² (Sartre, TE 102, 49).

³³ (cf. also TE 56)

³⁴ Cappelen & Dever (2013: 1).

³⁵ Ibid., 2.

³⁶ Zahavi (2006: 17)

-
- ³⁷ Castañeda (1966; 1999); Perry (1979); Nagel (1974); Cassam (1997).
- ³⁸ Chalmers (1996: 197).
- ³⁹ Zahavi (2006: 27).
- ⁴⁰ For a disentanglement of the different notions used to express the subjective character of experience, such as ‘mineness’, ‘me-ness’ and ‘for-me-ness’, see Guillot (2017).
- ⁴¹ But cf. Thompson (2010), who thinks we can preserve the idea of subjectivity without the Ego.
- ⁴² (*translucide*, TE 41).
- ⁴³ (Sartre, TE 63-4).
- ⁴⁴ (Sartre, TE 63-4)
- ⁴⁵ (“premier degré”, TE 41)
- ⁴⁶ In some passages of *Transcendence of the Ego*, then in *Being and Nothingness*, Sartre draws on a distinction between pure and impure forms of reflection. A reviewer has suggested that pure reflection might form the basis for comprehending one’s own implicit and pre-reflective experiences more directly, and not in the bad faith of impure reflection. If so, it might then appear that Sartre endorses a privilege to first-personal awareness, contrary to our arguments. This material of Sartre’s is not especially clear and raises numerous difficulties (see for instance Sarah Richmond’s introduction to *The transcendence of the Ego*, Richmond 2004: xii-xiii); Sartre himself substantially modified his account of pure reflection between *Transcendence of the Ego* and *Being and Nothingness*, bringing up more obscurities as to the epistemic status of such reflection (Webb 2022: 79). While interpretations of Sartre can and will differ, we note that he still describes pure reflection as nothing “but a quasi-knowledge” (*quasi-connaissance*, BN 162), hence we maintain that it does not provide propositional knowledge of ourselves. As such, we think his account of pure reflection is compatible with the position we outline.
- ⁴⁷ (Sartre, TE 80-1)
- ⁴⁸ Dreyfus & Kelly (2007)
- ⁴⁹ Dennett (1991: 76)
- ⁵⁰ Sartre (1958: 90)
- ⁵¹ cf. Moran (2002)
- ⁵² Sartre gives it more attention in his abandoned *Notebooks for an Ethics*. Judgments differ as to the ultimate success of this account.
- ⁵³ Sartre (1958: 564); cf. Catalano (1990: 683).
- ⁵⁴ Catalano (1990: 683).
- ⁵⁵ *Ibid.*, 693.
- ⁵⁶ While this is available to all, coming to have knowledge of it may require engaging in something like existential psychoanalysis of it, especially in regard to self-knowledge. Elaborating that account, and Sartre’s discussions concerning the fundamental project, is something we cannot do here.
- ⁵⁷ We cannot here engage with intellectualist replies, but see Jason Stanley’s *Knowing-how* (2011).
- ⁵⁸ Brownstein (2014).
- ⁵⁹ cf. Sartre (1958: 346)
- ⁶⁰ (TE 86)
- ⁶¹ Bortolotti (2008: 102)
- ⁶² Bortolotti (2008).
- ⁶³ cf. Kapitan (2015).