Early Sartre on Freedom and Ethics

Peter Poellner

Abstract: This paper offers a revisionary interpretation of Sartre’s early views on human freedom. Sartre articulates a subtle account of a fundamental sense of human freedom as autonomy, in terms of human consciousness being both reasons-responsive and in a distinctive sense self-determining. The aspects of Sartre’s theory of human freedom that underpin his early ethics are shown to be based on his phenomenological analysis of consciousness as, in its fundamental mode of self-presence, not an object in the world (Section 1). Sartre has a multi-level theory of the reasons-sensitivity of consciousness. At one level, consciousness’s being alive to reasons is a matter of the affective perception of values and disvalues as features of phenomenal objects (Section 2). This part of his theory, a development of Scheler’s, is, however, situated within a broader phenomenological analysis resulting in the claim that the ultimate reasons acknowledged by consciousness neither are nor justifiably could be values adequately presentable as intentional objects. Consciousness’s ultimate reasons are, in this sense, not given by the world but by itself (Section 3). Section 4 reconstructs and assesses Sartre’s argument that consciousness cannot rationally have an ultimate end other than self-transparent (‘authentic’) freedom itself.

Jean-Paul Sartre has usually been regarded as the advocate of a libertarian, quasi-voluntarist ‘radical freedom’, a humanization of the kind of free will that Descartes attributed to God (Plantinga 1958/9; Føllesdal 1981; Taylor 1982). The present paper aims to contribute to a revision of this still influential image of the early Sartre’s philosophy. I shall argue that there are resources in it for a more subtle and attractive conception of human freedom, and I shall reconstruct Sartre’s argument, in the Notebooks for an Ethics and elsewhere, that a proper appreciation of this can ground, indeed in his view rationally mandates, an ethics of intersubjectivity that acknowledges the value of ‘humanity’ (Sartre 1999: 108) as the fundamental value.

Any interpretation of the early Sartre’s views needs to take a stance on how seriously to take Sartre’s description of his project, in the subtitle of Being and Nothingness, as ‘phenomenological ontology’. As Gardner (2009; 2011) notes, there is something hermeneutically unsatisfactory about approaches that simply deny the ambitions indicated by the term ‘ontology’ and have it that Sartre’s aspirations were entirely phenomenological in the sense intended by the early Husserl: descriptive of experience and its contents without incurring ontological or metaphysical commitments. On the other hand, we need not accept Gardner’s (2011: 67) strong claim that a metaphysical dimension of Sartre’s analyses is
presupposed by nearly all aspects of his reflections on freedom. Sartre, despite occasional vacillations (Sartre 1999: 81–2), holds that the practical significance of his philosophy, if it has any, must be based exclusively on the plausibility or otherwise of its phenomenological insights and on inferences from them: ‘even if God existed, that would make no difference from [existentialism’s] point of view’ (Sartre 1980: 56; cf. 1999: 108). Further evidence of Sartre’s belief in the independence of his practical philosophy from metaphysics can be found in the text of Being and Nothingness itself, and I shall return to it below. For now, I shall adopt the working hypothesis that the concept of freedom which is essential for Sartre’s conceptions of authenticity and ethics does not involve metaphysical commitments, and that it can illuminatingly be interpreted as an attempt to complete the endeavour, left uncompleted by Kant, of emancipating ethics from metaphysics.

The structure of the paper is as follows. Section 1 outlines two basic claims about human consciousness that function as premises in Sartre’s account of freedom. In Section 2, I argue that he conceives of the freedom of consciousness in terms of its being self-determining and responsive to reasons, and I reconstruct his account of worldly reasons—those reasons which appear to be located in the worldly objects of consciousness. Section 3 explicates what Sartre means by consciousness’s self-determination in terms of its ultimate reasons not being given to it by the world. Section 4 examines his justification for this claim and shows its essential link with a distinctive conception of ethics.

1. Consciousness

According to the interpretation proposed in what follows, Sartrean freedom is a necessary feature of finite, reasons-acknowledging (‘human’) consciousness.4 In the ‘Introduction’ to Being and Nothingness, Sartre advances a number of important claims about consciousness, of which two are central to his characterization of freedom (Sartre 2003: 6–12). (1) Human conscious experiences are generally intentional, including an awareness of phenomenal (‘intentional’) objects, all of which are external to the consciousness to which they appear. (2) All intentional consciousness is conscious not only of some object or objects, but also of itself: every intentional experience is consciously ‘present to’ itself, though not as an object. Both claims would require detailed justification, but for my purposes some clarificatory remarks will have to suffice.5

1.1

Sartre specifies as the essential characteristic of an intentional (phenomenal) object its ‘transcendence’ relative to any one experience of it. He thus adopts Husserl’s broad conception of intentional objects, which includes in its extension real, fictional and imagined particulars (including events), types, properties and
their instances, states of affairs, aspectual modes of presentation attended to as such, linguistic senses when thought about rather than simply understood, and experiential attitudes as thought about rather than simply ‘lived through’ (i.e. experienced). Husserl’s more technical exposition of what Sartre calls an intentional object’s transcendence may be summarized as follows:

X is an intentional object for a subject S at time t just in case (1) S at t is conscious of X through an intentional experience E, which may present X either directly or indirectly, e.g. via linguistic symbols; (2) S at t has the ability to become conscious of X again, without modifying it, in other experiences E₁, E₂, … Eₙ, with diverse basic attitudinal characters (such as perceiving, neutrally imagining, recollecting, willing); (3) S at t has the ability to become conscious of X as the same across these different experiences of it; (4) on the basis of (1–3), S at t has the ability to engage in conscious judgings or other, judgement-guided actions concerning or involving X. (cf. Husserl 1973: Section 13)⁶

For Husserl, if and only if something is an intentional object for me in this sense, I can be said to be representing it conceptually (ibid.). It follows that all conceptual representation and all action that is guided by conceptually structured reasons is necessarily phenomenally conscious. Nothing can be such a reason for me unless it, or something understood as presentifying it, displays some phenomenal character (some ‘what-it-is-like-ness’) to me. There is so far nothing in these foundational theses of Husserlian phenomenology that Sartre disagrees with. He also concurs with Husserl on a further claim concerning objecthood: in order for an item to be an intentional object, it needs to be ‘thematized’, that is, attended to (Sartre 2003: 292), though thematization is not sufficient for objectification.

Sartre’s thesis that all intentional objects are external to consciousness involves, first, the phenomenological claim that none of the object-presenting phenomenal properties given in an intentional experience are inner sensations—e.g., ‘colour sensations’, or Husserl’s hyletic data—but that they are all (re-)presentational, pertaining to the object as (re-)presented. Second, it includes a radicalization of Husserl’s denial of the Veil of Ideas theory of consciousness: consciousness is not directed at its objects via images if these are understood as inner objects numerically distinct from the objects intended. Phenomenological analysis shows that not only perception, but also imaginative experience is devoid of such inner objects (Sartre 2004a: 68–83). Third, even if there were such mental image-objects, they could not be constituents of consciousness, for Sartre defines consciousness as the experiential attitude towards some object-involving content—the cogitare/noesis as opposed to the cogitatum/noema.⁷ While there can be no such attitude without some content or other (hence talk of ‘consciousness’ is an abstraction; Sartre 2003: 27), it makes no sense to say that the content is a component of the attitude, contributing to its identity, since the attitude can evidently vary independently of the content. So when Sartre speaks of consciousness, he refers to what is expressed by the italicized terms in sentences like these:
I am imagining the president’s arrival.

I am perceiving that the racquet is broken.

I am willing that he should be helped.\(^8\)

I shall henceforth use ‘consciousness’ in this strict Sartrean sense to refer to occurrent experiential attitudes. The expression ‘for-itself’ will be used, as in Sartre, interchangeably with ‘the human reality’ to denote the complex mode of appearance to itself of the finite existent who is conscious in a reasons-sensitive way—the human being—as well as that existent itself.\(^9\)

1.2

But is not my imagining, perceiving or willing—if it is indeed conscious at all—itself an object of consciousness for me when I am engaged in it? Sartre rejects two mutually exclusive and seemingly jointly exhaustive views about the nature of occurrent experiential attitudes. On the one hand, they are not simply unconscious, since there is an evident experiential difference between, say, imagining, perceiving, or willing the same object-involving content. But, on the other hand, the experiential attitude also cannot be present as an intentional object to me while I am ‘living through’ it. To show this, Sartre initially appeals to Husserl’s point that reflection—the thematization of one’s experiencing—normally presents the experience reflected upon (and not just its content) as having been conscious prior to the inception of the reflective act (ibid.: 7, 9). If this is correct, and if that mode of presentation is veridical, then the experiential attitude prior to reflection, while conscious, cannot have been conscious as an intentional object, since intentional objecthood requires thematization. But this point does not get to the heart of Sartre’s view of the matter. Even if I were thematically conscious of my lived-through ‘consciousnesses’, they could not be objects for me at the time. Recall that it is an essential characteristic of an intentional object that it can be presented, without itself being changed, in numerically distinct intentional experiences instantiating diverse basic intentional attitudes to it: for instance, in order for a certain sound sequence to be an intentional object for me, it must be possible for me to represent and re-identify that very sound sequence as the same, without modifying it (the objective content of my representation) by recalling it, or imagining it, or when hearing it again. But this is in principle impossible for experiential attitudes as they are ‘lived through’: My auditory perceiving of the beginning of Beethoven’s Emperor concerto is necessarily modified when it is the object of other basic intentional attitudes, e.g., when I merely imagine or merely recollect my perceiving of those initial bars—for in doing so I am, necessarily, not actually perceiving them. This is part of Sartre’s point when he says that ‘the immediate consciousness which I have of perceiving does not permit me to judge or to will or to be ashamed of it. It does not know my perception, does not posit [i.e. objectify] it’ (ibid.: 9).
2. Freedom and Worldly Reasons

We are now in a position to approach Sartre’s most famous claim: humans are ‘radically’ free whenever they engage in reasons-responsive intentional action. Sartre emphasizes that the freedom he attributes to the human being is the ‘technical and philosophical’ concept of autonomy:

The formula ‘to be free’ does not mean ‘to obtain what one has wished’ but rather ‘by oneself to determine oneself to wish’ (in the broad sense of choosing). In other words success is not important to freedom. (ibid.: 505)

The autonomy in question is logically independent of whether an agent is negatively free in a physical, political or executive sense, that is, roughly, of whether or not the agent is prevented from attaining the ends of her all-things-considered desires or choices by obstacles external to her or by limitations of her executive capacities, such as physical or (some) mental frailties: an agent remains fully autonomous in the relevant sense even under the most severe constraints to her freedom in these latter ‘popular’ senses, as long as these constraints are compatible with reasons-acknowledging mental agency. In an initial rough statement, human autonomy as conceived by Sartre may be characterized as follows:

(Freedom) The for-itself is free whenever (F1) it makes conscious decisions to act (F2) in the reasons-sensitive pursuit of ends (ibid.: 475), which (F3) are not given to its consciousness at the time of decision from outside it, but are rather determined by that consciousness (ibid.: 465–6). In choosing its ends in this way, consciousness at the time of choice, and thus the for-itself, determines itself, and is therefore autonomous.

Each of the components of this outline calls for explication and defence. But first a further element needs to be mentioned. I cannot decide anything in pursuit of an end without a grasp of what I need to do in order to (attempt to) realize that end. And this implies that I need to understand what speaks in favour of taking one course of action rather than another. Sartre calls what I grasp here a motif, defined as a ‘rational justifying consideration’ (ibid.: 468). A motif is an (at least apparent) instrumental reason to attempt to perform a certain action, a presumed objective feature or state of affairs that makes the action good or suitable for the purpose of achieving a more ulterior end (ibid.: 469–70). Hence freedom requires a practical, not necessarily thematic, understanding of means–end relations. Instrumental reasons are often themselves ends of subordinate actions involved in the pursuit of more ulterior ends. Clovis’s motif for converting to Catholicism is securing the support of the powerful episcopate in his pursuit of the end of conquering Gaul. Such subordinate, instrumental ends depend for their intelligibility on the role they play in some more comprehensive project, and thus on some intrinsic end(s) that is (are) not instrumentally subordinate to others (ibid.: 468). While it is an objective matter whether some consideration is a genuine,
and not merely an apparent, instrumental reason (an appropriate *motif*), it is a separate question, to be considered shortly, whether intrinsic ends can be said to be rationally warranted.

Returning now to (F1) in the above initial statement of Sartre’s account of human freedom, the conscious ‘decision’ (ibid.: 475) that is claimed to be essential to free action is what he more frequently calls the choosing of an action. Unlike some contemporary philosophers of action, Sartre does not mean by ‘decision’ a response to uncertainty about what to do (e.g., Mele 2003: ch. 9). Rather, the term refers to the agent’s awareness of an effective commitment to an end, whether instrumental or intrinsic: a normally non-thematic consciousness of effective desire. The decision is experienced as the agent’s own (rather than as an alien force) precisely in so far as it is the definitive embracing of an end as *his* end, to be realized or pursued *now* (Sartre 2003: 471). This does not commit Sartre to the view that, in specifying the content of the agent’s awareness of his end, reference needs always to be made to the agent himself or to his intentions.10 Sartre’s position is closer to the Jamesian ideo-motor theory that effective desires are certain kinds of firm occurrent *convictions* about future occurrences or states of affairs.11 But whatever the correct full characterization of a decision (choosing) may turn out to be, for Sartre’s purposes it suffices to say that it is that element in an intentional action which distinguishes it for the agent at the time, on the one hand, from behaviourally type-identical happenings that are merely passive behaviours (being pushed) or reflexes, and, on the other hand, from other attitudes towards the same content such as wishes or mere entertainings. It is that aspect of the experience of action that grounds the agent’s normal non-observational and non-inferential ability to say, with respect to many of her behaviours, under some descriptions: ‘I did it’, in contrast to other behaviours which are truthfully reported by her as: ‘it happened to me’.

The main burden of Sartre’s theory of freedom rests on components (F2) and (F3) in the initial characterization: free action is action for the sake of intrinsic (i.e. non-instrumental) ends that are nevertheless fully determined by the for-itself. The standard interpretation of this claim in the literature has been that there is nothing on the noematic side of experience—nothing in its intentional content—that grounds or motivates the for-itself’s adoption of some intrinsic end at a time. What makes some intrinsic end suitable to figure as such is only the for-itself’s ‘choosing’ it, committing itself to it (e.g., Follesdal 1981: 397–8; Taylor 1982: 121). But, a familiar objection goes, a ‘choice’ which has no grounds in *what* is chosen, which cannot specify good-making features in the noematic content that are independent of the choice and motivate it, is not intelligible as a *choice* at all (Taylor: ibid.). This reading of Sartre’s thesis that ‘nothing outside consciousness’ can determine the for-itself’s choice of intrinsic ends is not forced upon us by the Sartrean texts. There are strong indications in them favouring a different interpretation.

Sartre does indeed assert that no object can causally determine consciousness (Sartre 2003: 11), that ‘nothing’ justifies me in adopting this or that value (ibid.: 62), that freedom is the ungrounded ground of values (ibid.), and that human
reality does not receive its values from outside (ibid.: 463). Now even if these claims are read, as they need to be, with a phenomenological prefix, they may seem problematic. Since what is ‘outside’ consciousness also includes objects as experienced, he seems to hold the view that nothing in the phenomenal world can motivate the for-itself to adopt any particular or specific intrinsic value or end. And this might be thought to imply that any noematic content whatever might in principle be chosen by a for-itself as an intrinsic end, and by virtue of this ‘choice’, become a value for it. But this interpretation—that Sartre’s for-itself resembles Descartes’ God in that at least its fundamental choices are not made on the basis of good-making features of what is chosen—comes under pressure from some of his explicit statements:

The Good, Unity, Identity are necessities [even] for God . . . We will not get out of this by making God some kind of absolute and unconditioned freedom, as Descartes does. For this is pure gratuitousness and man has to consider the world as the pure trace of a freedom he cannot comprehend. Which comes down to saying: that’s how it is . . . [God’s] activity itself, if it is not to be a pure caprice, a pure contingency, an arbitrary choice, must proceed from a being logically prior to it. (Sartre 1992: 517–18)

The challenge for Sartre, then, is to reconcile his acknowledgment, here and in other passages, that freedom cannot fundamentally consist in such gratuitous, capricious, or ‘arbitrary choice’ (cf. Sartre 2003: 474–5) with his equally emphatic claim that the grounds of the for-itself’s choices cannot lie outside of consciousness, if its freedom consists in not being determined by anything other than itself. Strewn throughout the Sartre’s early œuvre, alongside his denials of the idea that the for-itself’s ends are, in the final analysis, ‘given’ as objects, we find passages suggesting that all intrinsic ends are ‘values’, and that at least some values must manifest themselves as features of worldly objects if there is to be a world for a subject at all (ibid.: 60–2; Sartre 2004b: 17–20; 1999: 50, 88). Sartre also asserts the stronger claim that even the end of authentic action is a ‘good’ that is (partly) noematic, i.e. that qualifies the content of what is desired in such action, although it cannot be adequately captured by any intentional content (ibid.: 113; 1992: 555f). Is he simply being inconsistent?

The key to a resolution of the tension here lies in recognizing that Sartre’s account of practical normativity is based on, while importantly modifying, Scheler’s phenomenology of value. For the remainder of this section, I shall reconstruct the Schelerian elements in his account. Sartre claims that all intrinsic ends are or instantiate at least apparent values (Sartre 2003: 60–2, 459, 465). We cannot intelligibly specify an intrinsic end—that which an action aims to achieve and for the sake of which it is performed—except in terms of some apparent good, some feature of the end aimed at which makes it for the agent, not just desired, but ‘desirable’ (ibid.: 456). Values or goods are characteristics that involve ‘demands’ (ibid.: 62): ‘by their nature they “ought to be”’ (Sartre 1999: 88). Sincerely to take something to be a value is to acknowledge that it merits, at
least pro tanto, being favoured, pursued, realized or maintained. Therefore, if that for the sake of which I choose an action (its intrinsic end) is or instantiates a value, there is necessarily something that speaks in favour of the action: there is a justifying reason for the action, though not necessarily a conclusive or overriding reason. Only if my actions are done for the sake of what I take to be in some respect good or valuable are they properly intelligible to me.\textsuperscript{14} This is why Sartre says that personhood and ownership of action are inseparable from the idea of ‘the good’ (Sartre 1992: 558). Acknowledging that some putative value ‘demands’ or merits being pursued or realized implies, if the agent takes himself to be appropriately situated, and in the absence of other, overriding considerations, an inclination or motivation to pursue or realize it (Sartre 2003: 60–2).

Theoretically unprejudiced phenomenological analysis shows that the value or desirability of an intrinsic end in many, although not all, cases appears to us as a property of some actual or possible object that is such as to inherently merit or demand an exemplification of that property to be realized or maintained or appreciated. For example, the beauty of a painting appears to us as a property of the painting that merits appreciation. There are no conscious internal pleasures from which that property might conceivably be ‘projected’. Where the end is the removal of a disvalue, the latter appears as an objective property demanding action to remove it: ‘everything happens as if we lived in a world where the objects, apart from their qualities of heat, odour, shape, etc., had qualities such as the disgusting, attractive, charming, useful, etc., and as if these qualities were forces exerting certain effects on us’ (Sartre 2004b: 19). These apparent forces of objects are, phenomenologically, valuational demands which often entail normative demands (‘exigencies’) upon agents conscious of themselves as relevantly situated. The normative demand, or more broadly, the normative force of an actual or possible course of action is what I am conscious of when I am conscious of it, for example, as impermissible, acceptable, categorically required, etc. What I call a valuational demand is acknowledged by acknowledging that it is good (bad) that some possible item should (not) be actual, that it ought (not) to be actual—indeed independently of whether it is possible to act in any way relevant to its actualization. Values constitutively involve valuational demands (Sartre 2003: 62). For Sartre, as for Scheler, normative demands necessarily depend on these valuational demands—on the ‘ideal ought’ that is constitutive of value (Sartre 1999: 88; 1992: 555). But there clearly are many things which can be acknowledged as values by a subject although they do not entail any normative demands upon her, or even upon any finite agent. I can acknowledge the value of the freedom of unknown people who are no longer alive, and for the sake of whose freedom I cannot act. I can also intelligibly take as a value the fact that the structure of the physical universe is as it is.

Value, without which there could be no intelligible ends, therefore has a peculiar character: it is frequently encountered as a property of everyday phenomenal objects, but it is also essentially such that to-be-pursuedness or to-be-favouredness are internal to it. Acknowledging value, where it is encountered in objects, accordingly requires a mode of (re-)presentation with ‘mind-to-
world direction of fit’ that is capable of grasping these apparently objective phenomenal features and is, by virtue of its representational content, capable of providing reasons and of motivating action. Sartre follows Scheler in holding that many intentional emotions combine these seemingly incompatible characteristics: ‘My indignation has given to me the negative value “baseness”, my admiration has given the positive value “grandeur”’ (Sartre 2003: 62). ‘So we see that . . . these notorious “subjective” reactions, hatred, love, fear, sympathy . . . are merely ways of discovering the world. It is things themselves which suddenly reveal themselves to us as hateful, likeable, terrifying, lovable’ (Sartre 2002a: 383–4, also 2004a: 68–9).15 His idea is, then, that many emotions are (re-)presentational acts and that what is given in them are apparent value characteristics of phenomenal objects. A disvalue such as moral ‘baseness’ manifests itself in emotions such as indignation, and it is given through them as qualifying a particular actual or possible object—for example, an (objectified, represented) intention such as Paul’s expressed intention to betray his friend Pierre’s trust. Being given in this affective way to Anne who knows both, it is necessarily presented to her as meriting, not just causing, her negative affective response and as demanding certain actions: to warn Pierre, to avoid or reprimand Paul, and so forth. Emotions such as indignation have, literally, an intuitive character: they present a certain characteristic phenomenology which is representational—just like my perceptual experience of a red tulip represents an instance of the secondary property redness as a property of the tulip. But, for Sartre, they are also doxastic: in order not to be motivationally inert, their content needs to be taken to be as it is presented (Sartre 2002b: 49). His account of some emotions as affective perceptions of values entails that they have the condition of success pertinent to perceptions—empirical veridicality—and that their supplying of genuine reasons depends on this condition being satisfied.16 The view, shared by Sartre with his phenomenological predecessors, that perceptible properties of objects can be literally perceived also in merely imagined instantiations allows this account of affective value perception to be extended to imaginatively envisaged or anticipated objects (cf. Husserl 1982: Sections 4 and 70).

If values are in many cases manifested as noematic aspects of conscious emotions, it is easily intelligible how some emotions can ground choices which are not simply arbitrary—that is, how they can rationalize actions.17 A soldier who flees from fear of an enemy assault, and whose flight is a reasons-responsive intentional action—unlike in cases of ‘shell shock’—correctly apprehends through his fear a certain event (the enemy attack) as a disvalue, as a terrifying threat to his life, and this affective perception presents avoidance of that threat, making it disappear by running away, as something he should do—as a (decisive) motif.18 But motifs, as we saw, depend on non-instrumental ends. The fear, on Sartre’s construal, also reveals the soldier’s dominant intrinsic end at the time: preserving his life (Sartre 2003: 459, 465).19 This end, to be sure, is not thematically manifested in his fear, but is implicit in his terrified apprehension of the enemy assault as lethally threatening to him, as a specific kind of overriding disvalue, and of avoiding that threat in that particular way as a conclusive motif.20 Being
implicit, his end is available for explication: if the soldier is subsequently asked about his end at the time, and if he is not in bad faith and does not suffer from memory failure or other cognitive defects, he will concede that the fleeing was something he did, and that his end was indeed the preservation of his life. The soldier’s behaviour is therefore a reasons-sensitive intentional action, involving a decision, a choice in the light of non-reflectively apprehended reasons. His reasons are manifest, though not fully explicit, in his fear—his affective perception of a disvalue in the world. For Sartre, value, practical normativity and choice are therefore at this basic level non-reflective phenomena—he plausibly maintains that they are not generated by reflective distancing.

In so far as action is guided by unreflective affective perception that is alive to reasons, it satisfies (F2) in the initial characterization of freedom. Still, such reasons-responsiveness is clearly a matter of degree, while Sartre wants to insist that the for-itself is either ‘wholly free’ or not at all (ibid.: 464). This puzzle can only be properly addressed once the contours of Sartre’s account have been more fully exhibited (see note 34). What can be said at this stage is that, for Sartre, there is a certain level of reasons-awareness which, when present, is sufficient for full autonomy and that this level may be, and usually is, attained by unreflective affective consciousness.

3. The Self-Determination of Consciousness

So far, Sartre’s account corresponds closely to Scheler’s. In acting ‘from passion’ the for-itself typically makes choices, potentially on the basis of genuine reasons. But Sartre also claims—in (F3)—that a for-itself’s fundamental reasons cannot be found in objects. Value ‘can be revealed only to an active freedom which makes it exist as value by the sole fact of recognizing it as such’ (ibid.: 62) and nothing external to consciousness, and thus no object, can affect the for-itself in such a way as to determine its actions, for consciousness determines itself (ibid.: 464). The challenge is to make these assertions compatible with the Schelerian story above and with Sartre’s insistence that the commitment to some end or other is not, in the final analysis, ‘arbitrary’. I submit that what Sartre means by the determination of any end, qua value, by the for-itself, and the former’s independence of anything outside consciousness, is a complex thesis that can be articulated as four distinct claims (F3.i)–(F3.iv), of which only (F3.iii) and (F3.iv) are specific to Sartre.

(F3.i) Values are not objective in the strong sense of being either primary properties or Platonic entities (Sartre 1992: 555). Such objectivity, assuming it makes sense at all, would be incompatible with the demand-structure that is constitutive of value. A demand’s being depends on being experienced, and value is therefore existentially dependent on consciousness (Sartre 2003: 62). Nevertheless, as we have seen, values are not invariably ‘subjective’ either—many of them, and all of those that are primary in the epistemology of value (the order of knowledge), pertain
to the object-involving noematic correlates of consciousness without which there can be no intentional consciousness (Sartre 1992: 556).

(F3.ii) In his more considered formulations, Sartre does not say that nothing external can affect the for-itself—which, apart from being bad phenomenology, would make his position inconsistent (cf. Sartre 2004b: 18–19)—but, rather, that ‘events can affect you only if they are assumed by your own possibilities’ (Sartre 1999: 96). The ‘forces’ to which human reality in the first-personal perspective of agency is distinctively subject are, necessarily, not given to it as brute causal forces (Sartre 2003: 54). Qua intentional agent, the for-itself experiences itself as affected by apparent valuational and normative demands, which can only be effective if they are consciously acknowledged as such (ibid.: 55, 60). In being ‘moved’ by such a demand (by acting on it), an agent is not simply passively caused to behave in some way or other, but necessarily actively ensigans some possibility—an as yet absent, futural act or worldly state—as a value to be realized (ibid.: 148–51).21 What this envisaging, which typically is an implicit, unthematic presentification of the valued end, precisely consists in is a difficult phenomenological question requiring a detailed examination of imaging consciousness that is beyond the scope of this paper (see esp. Sartre 2004a: 68–83, 184–8).

(F3.iii) The ‘common-sense’ view would have it that some aspect of my facticity—say, an intense bodily pain—can exert an overriding, all-things-considered valuational and normative force, such that I am compelled by it to adopt a specific end (‘cessation of the pain’) and to try to act accordingly. Since my facticity pertains to the noematic (content) aspect of experience and is a particular present or past reality, being normatively compelled by such a reality ‘external’ to consciousness would be incompatible with consciousness determining itself, and hence with the for-itself’s autonomy. Sartre firmly denies that such a normative compulsion is possible. If the case in question is one of reflex-like non-action, such as wincing caused by a severe toothache, I am indeed compelled, but am not acting intentionally. But if, instead, we are dealing with a case of reasons-responsive intentional action—the mountain hiker sitting down by the wayside because his fatigue has become ‘intolerable’—it is behaviour guided by values, and no values can be imposed on consciousness as overriding by the for-itself’s facticity (Sartre 2003: 458).22 One reason for this is that, by Sartre’s lights, part of the valuational and normative force of a possible end depends upon its position within a holistic system of ends. While a for-itself may acknowledge a multitude of individual states of affairs as intrinsic ends, it is essentially conscious of these as partial ends. They depend for their full intelligibility to the agent himself upon an all-comprehensive, albeit normally not thematic end towards which he projects itself, namely some specific relation to the world as a whole. Less comprehensive ends such as ‘being free of pain’, ‘enjoying this musical performance’, ‘being loved’, may be intrinsic (non-instrumental) for me, but they are incompletely characterized. The specific way any such limited ends are entertained involves an implicit (‘horizonal’) under-
standing of their position in an overall, itself mostly implicit, ‘fundamental project’, a description of which would be an explication of what I take my life to be about at that time. A person whose ends during a particular stretch of his life would be correctly, albeit incompletely, described as hedonistic does not just understand individual actions of his as aiming at this sensory pleasure or the avoidance of that sensory displeasure, but also normally understands these particular ends as fitting into and partly realizing a fundamental orientation towards his life, within which other kinds of ends show up as ‘not worth the effort’. It is as part of this overall orientation that his particular pursuits and their goals make whatever sense they make to him. A person whose pursuits were apparently atomic and had no intelligible relation to such a comprehensive orientation to the world within which they ‘make sense’ would rightly be described as, to that extent, alienated from himself. But even such a person’s ends would not be genuinely atomic: an adequate characterization of them would still have to make reference to their experienced context within the person’s life as a whole, to wit, to their experientially lacking a certain kind of overall intelligibility in relation to it.

If Sartre’s analysis of the consciousness of ends as implicitly holistic in this manner is correct, then it should be clear, first, that no particular, limited aspect of my facticity, such as a pain located somewhere in my lived body, can by itself normatively compel me—make a conclusive, overriding normative demand on me—to adopt any particular overall orientation to the world or to my life. How could it do this, its valuational and normative demands, such as they are, being essentially confined to that particular, limited aspect of my facticity (‘this pain is bad, try to get rid of it’)? Second, it is plausible to say that the specific force of these valuational and normative demands will depend, in part although not entirely, on the place of the values involved within my ‘fundamental project’ (ibid.: 482–7). Sartre’s idea is this: it does not necessarily depend on my specific freely chosen ends whether some aspect of my facticity, or some objective worldly content I am aware of, instantiates a generic positive or negative value (Sartre 1999: 50–1). When I am aware of certain sensations produced by a dentist’s drill without anaesthetic, or when I am attentively contemplating a beautiful object under the relevant aspect, I will be aware of these as values or disvalues whatever my freely choosable ends are (or those ends consistent with the relevant attention, in the second example). But the fully determinate character of the (dis)values experienced, and consequently their precise normative force for me, will indeed be a function of their place within my system of ends. If my overall orientation—the way I implicitly understand my life, my relation to the world—is fundamentally ‘hedonistic’, the normative demands made by a severe bodily pain will be different from what they will be if that orientation is about mastery of a recalcitrant nature, or religiously ascetic (Sartre 2003: 476–9). For the committed ascetic, the pain may be experienced as demanding not to be removed, but to be endured, perhaps even to be welcomed. Note that this does not imply that the pain is not (correctly) experienced by the ascetic as pro tanto bad—to claim this would be to misunderstand what pain is. Consciousness is
not autonomous vis-à-vis the for-itself’s facticity or ‘nature’ in this way, and Sartre does not say that it is. Indeed, it is only because the pain is a given albeit not response-determining disvalue that ‘enduring’ it makes sense within the religious ascetic’s project (‘mortifying the flesh’). The upshot, then, of this part of Sartre’s argument is that no past or present reality external to consciousness can impose conclusive, all-things-considered normative demands on the for-itself.

(F3.iv) The fourth aspect of Sartre’s understanding of the for-itself’s ‘radical’ freedom in its choice of intrinsic ends is one of his most distinctive claims. Not only can no aspect of my facticity by itself present overriding ends/values to me, no actual or possible object can do so (Sartre 1999: 110–12; 2003: 117–19, 500–1). If we think of the world as a maximally comprehensive meaningful structure of objects that are, have been, or will be actual, we might say that, according to Sartre, neither the actual world nor any possible world, considered in this way, can supply values which I ought to acknowledge as overriding and I am always implicitly aware of this non-ultimacy of the value of ‘the world’. No possible envisaged state of the object-world is such that I do not justifiably, i.e., with good reason, experience it as a non-ultimate end, and as ‘demanding to be surpassed’. For Sartre, the for-itself is free of the world in precisely this sense, in that all values that the for-itself can encounter as fully objective in the world, and any worldly values that it can in principle adequately envisage as objective, can rightly be experienced and judged by it as non-ultimate.

The for-itself cannot find ‘an ultimate limit to his surpassings’ in any object (Sartre 2003: 393). Sartre’s point is both evaluative or normative and factual: not only does the for-itself have a reason to ‘surpass’ any worldly value, it implicitly always does thus surpass it. Clearly, one might be sympathetic to the first claim while rejecting the second. But for Sartre the ultimate end of human desire is always an ‘unrealizable’—something that cannot be a real or realizable object (Sartre 1992: 551–3; 1999: 199). The most frequent, inauthentic, manifestation of this is the for-itself’s aiming at completion or wholeness, implicitly taken by it as realizable—a determinable end which Sartre interprets ontologically as an impossible condition in which the for-itself would be both consciousness and entirely non-contingent, self-grounding object (Sartre 2003: 113–15). This is an incoherent project since, for one thing, qua consciousness it cannot fundamentally be an object to itself. We need not concern ourselves at this point with the plausibility of Sartre’s thesis about the ultimate value pursued by inauthentic consciousness.23 What matters for now is that even this alleged common structure of inauthentic fundamental projects manifests merely a motivated misapprehension of an evaluative orientation towards ‘unrealizables’ which is not only humanly inescapable but, when authentically understood and modified, is also justified (Sartre 1999: 110–11, 198; 1992: 556–7). But even when authentically modified it cannot find its sufficient justification in any object—whether worldly-particular, or Platonically universal, or affective-internal (Hume’s sentiments), or purely rational (Kant)—and is in this sense ‘groundless’ or ‘unjustifiable’ (see Section 4). Nevertheless that fundamental evaluative orientation is,
Sartre insists, not ‘capricious’ like the ‘arbitrary choice’ of Descartes’ God (Sartre 1992: 517–18). An explication of the for-itself’s ultimate end would have to use evaluative concepts suitable to make it intelligible as a possible value, that is, to rationalize it, even if necessarily incompletely. In the case of the ultimate ends of inauthentic agents, an appropriate intelligibility-preserving explication presumably would have to articulate their common structure through a description such as ‘the permanent overcoming of experienced lack, taken as really possible’ (cf. Sartre 2003: 114–15). It is clearly incumbent on Sartre to motivate this thesis of the for-itself’s being free of the world, as expressed in (F3.iv). But before turning to his attempt to do this, and granting for now that it is successful, I first want to consider a more immediate objection to his account of freedom as interpreted here.

The objection targets the very foundations of Sartre’s approach, which is resolutely first-personal and phenomenological. Sartre’s thought is that from the perspective of agency, which is made explicit by existential phenomenology, we experience ourselves as undetermined by unconscious subpersonal events, by past conscious states (whether objectifyingly recollected or unthematic), and by the phenomenal object-world, and that the perspective of agency is incapable for us. Hence we are ‘condemned’ to experience ourselves as free in the sense explicated so far. But, the objection goes, this does not license Sartre’s conclusion that we are condemned to be free and hence to be responsible. Phenomenology may be ontologically misleading. We have good reasons to believe (the objector continues) that phenomenal consciousness supervenes on physical events and that the latter are governed by deterministic or probabilistic laws such that antecedent conditions and laws obtaining in the supervenience base determine unique behavioural outcomes or at least probabilities, and it is therefore fundamentally erroneous to accept the phenomenology at face value. And if consciousness is in fact determined or constrained by laws and by its own psychological and (ultimately) physical past, then it is not responsible. There are two responses available to Sartre, and he deploys both of them. He can argue, first, that the ontology presupposed by the objection is simply mistaken, that the entities and properties recognized by physical science (such as unactualized forces, dispositions and propensities) exist only as correlates of consciousness, and that the phenomenology of consciousness is simultaneously fundamental ontology. Elsewhere, he takes a different tack. He suggests that even if determinism were metaphysically true, this would make no difference to the truth and relevance of his analysis:

[W]e wished to establish anguish in its existential structure as consciousness of freedom. Now from this point of view the existence of a psychological determinism would not invalidate the results of our description. Either anguish is actually an unknown ignorance of this determinism—and then anguish apprehends itself in fact as freedom—or else one may claim that anguish is consciousness of being ignorant of the real
causes of our acts . . . But in this case we should suddenly appear to ourselves [entirely] as things in the world. (ibid.: 57–8; first emphasis mine)

Since only the first disjunct is compatible with the phenomenology (cf. Section 1) and thus only it could possibly be true, the clear implication is that an ‘unknown ignorance’ of a putative determinism in the unreflective perspective of action, or in practical reflection about what to do, is sufficient for Sartre’s purposes. In other passages he concedes in a similar vein that from a third-personal, theoretical perspective a subject’s behaviour might well be predictable to a sufficiently knowledgeable observer with a comprehensive knowledge of the subject’s facticity (ibid.: 475; 1992: 31), which again implies that if determinism were true it would not invalidate the core of Sartre’s account of freedom. This account is therefore, in essence, compatibilist: the conditions on freedom and responsibility Sartre has uncovered are taken by him to be jointly sufficient in a practical respect and do not require a further negative condition such as the absence of deterministic causal laws in a presumed physical base of consciousness. If the for-itself decides to act in the light of an apprehension of genuine evaluative reasons which it adopts as its own and which in the last resort are not given to it by ‘the world’, and if it ipso facto does not experience itself as acting because it is compelled to do so either by its conscious past or by some alien, brute causal force, then it is free and ‘responsible’. This view is less counterintuitive than it may first appear if we recall that Sartre is concerned with freedom as autonomy and therefore with the conditions of what makes an action genuinely mine—with ownership—rather than with questions of desert. 24

We should not leave this topic without a clarification of Sartre’s assertion that no moment of consciousness has sufficient antecedent causes (Sartre 2003: 11–12; 2004b: 35). Again it is advisable to read this, in so far as it has relevance in a practical respect, as a phenomenological claim, although Sartre takes the phenomenology also to be ontologically revelatory. When I act autonomously, I determine myself in the light of evaluative reasons which I am conscious of now, albeit often only implicitly, and in the final analysis for the sake of values which are not given to me by the world but which are ‘projected’ by myself now (Sartre 2003: 51–2). 25 This does not mean that my conscious past is not present to me or does not propose possibilities to me. Often, like the recidivist gambler, I thematically recall that I used decisively to favour φ-ing (e.g., not-gambling). But this recollection cannot determine me, if I am responsive to reasons, because the mere fact that I used to favour or desire something is no reason. My thematically apprehended conscious past can only affect me if I now endorse, or acquiesce in, its evaluative content, or some other relevant evaluative content, such as ‘being faithful to my past resolution’ (ibid.: 520). But such present endorsement or active acquiescence is obviously different from merely recollecting, or contemplating, that I used to desire or favour such-and-such (see also Moran 2001: 138–48). In this sense, my present consciousness is independent of my objectified past consciousness and, phenomenologically, is not caus-
ally determined by it. The reasons-responsiveness of consciousness explains what Sartre takes to be one of his fundamental insights: the necessary absence of psychological determinism in the first-personal perspective of agency and practical deliberation.

4. Freedom and Ethics

What are Sartre’s reasons for his most striking claim about freedom (F3.iv), that no actual or possible object does or justifiably could present overriding, all-things-considered evaluative and normative demands to a for-itself? An answer might be expected to emerge from his reflections on a properly self-transparent, authentic for-itself, which are to be found mostly in the texts on ethical questions (including the War Diaries) written in close temporal proximity to Being and Nothingness. A core theme of these ethical writings is that once the constitutive inconsistency of a for-itself’s inauthentic fundamental project is ‘grasped in full clarity’ by it, that project cannot be rationally maintained and, unless there is a relapse into bad faith, it is ‘consequently nihilate[d]’ (Sartre 2003: 496). In order to ‘grasp in full clarity’ the inconsistency of its inauthentic fundamental project, a for-itself needs to recognize the incompatibility of the end of that project with the very structure of consciousness, and in order to do this it needs to attain a thematic awareness of the phenomenal character of consciousness, i.e. of experiencing, of which it normally has a mostly un thematic (‘pre-reflective’) awareness. More generally, in order for a for-itself to be able lucidly to adopt an orientation to the world whose ultimate end it knows not to be incompatible with its own fundamental character, it needs to become explicitly aware of that character. Such a thematizing of experiencing as it is ‘lived through’ is of course what Sartre, adopting Husserl’s terminology, calls reflection. Reflection is a necessary feature of explicit self-transparency, and the latter is what he, after Heidegger, terms ‘authenticity’.

In order for it to be possible to thematize consciousness in its fundamental non-object-like character, a special, ‘pure’ kind of reflection is needed which, unlike ordinary (‘impure’) reflection, does not objectify the experiential attitude reflected upon. Without going into the details of Sartre’s conception of pure reflection (see esp. Sartre 1992: 472–82), it suffices to say that is supposed to accomplish such a non-objectifying thematization of my experiential attitude in its correlation with some object(s) (Sartre 2004b: 11–12; 1992: 4–5, 473–4). In pure reflection my ends are explicitly revealed as ‘in question’ (ibid.: 480)—as dependent on my choice and open to revision in the light (say) of possible critical interrogation. But if a commitment or project survives pure reflection, then it has been, in Sartre’s technical sense, ‘assumed’ by me. In this case, the outcome of pure reflection can be described as ‘a willing of what I will’ (ibid.: 479), as an endorsement of my originally pre-reflective choice: it ‘introduce[s] into the internal relationality of the Person the relation of solidarity, which will subsequently be modified into solidarity with others’ (ibid.).
Surprising as it may seem to any reader of Being and Nothingness in isolation, solidarity with others turns out to be an essential element of full-fledged authenticity (ibid.: 9, 485). A reconstruction of Sartre’s elliptical argument to this effect might be expected to bring out his reasons for his so far unsubstantiated claim that no object can justifiably constitute an ultimate value for a for-itself. First, recall that according to Sartre we act on the basis of values which we either affectively perceive in the objective world or which are taken by us to qualify our projected ends, the former being partly dependent on the latter. While what we choose, at least in the basic cases, is something particular—to listen to this musical performance, or to relieve this friend’s pain—we act for the sake of these particular ends because of value characteristics we take them to have, and these are in principle universal in the sense of being instantiatable also by other particulars:

To choose between this or that is at the same time to affirm the value of that which is chosen; for we are unable ever [knowingly] to choose the worse. What we choose is always [what seems to us] the better; and nothing can be better for us unless it is better for all. (Sartre 1980: 29–30, my emphasis. cf. 1992: 555–7)

Sartre’s imprecisely stated point in this popular lecture has often been taken to be of Kantian inspiration, but its provenance is again Schelerian. Scheler argued that values are at the logically basic level not agent-relative, because basic values either are or can become without evaluative distortion shareable characteristics of noematic contents which phenomenologically are not relative to, or indexed to, the particular agent experiencing or appraising them (Scheler 1973: 12–19, 265–75). They are in principle multiply instantiatable features of phenomenal objects, such as the specific beauty of a landscape (as seen from a certain point of view with a certain attentional focus), or the specific baseness or generosity of an objectified intention (as disclosable to any appropriately attuned affective sensibility).28 Nothing can be valuable for me or any other particular agent unless it is in this sense valuable simpliciter and therefore valuable, although not necessarily with the same determinate valuational force, ‘for all’ who adequately grasp the relevant features. Note that Sartre, following Scheler, does not derive what he calls the essential universality of value from its intersubjective share-ability, but vice versa (Sartre 1992: 555–7). But for Sartre, the justifiable variety of projects is subject to a further constraint:

I declare that freedom, in respect of concrete circumstances, can have no other end and aim but itself; and when once a man has seen that values depend on himself, in that state of forsakenness he can will only one thing, and that is freedom as the foundation of all values . . . Obviously, freedom as the definition of a human being does not depend upon others, but as soon as there is commitment, I am obliged to will the freedom of others at the same time as mine. I cannot make freedom my aim unless I make that of others equally my aim. (Sartre 1980: 51–2; cf. 1992: 414).
The necessity in question is both logical and moral (ibid.): I must ‘will’ my freedom and that of others if I am not to be both practically inconsistent and unethical. Sartre’s core argument, condensed in these passages, for an ethics of freedom (AEF) seems to be that, since (i) all actual value depends existentially on the reasons-sensitive consciousness of myself and others, (ii) if I value anything (and I cannot not value anything) (iii) I cannot consistently not ‘will’ such consciousness in myself and others, hence (iv) I cannot consistently not will freedom universally; and therefore (v) I must take free subjectivity itself, in each of its instances, as the primary value for the sake of which I act—as my primary end. In fact, the formulation in the passage cited above does not capture Sartre’s actual conclusion quite correctly. That conclusion is rather (v’): I must take authentic subjectivity, in each of its instances, as my primary end. Recognition of this ‘obligation’ (ibid.) is essential to full authenticity.

Regarding another for-itself’s freedom as an intrinsic end involves an affirmative and non-appropriative attitude to the other’s freedom (unlike in inauthentic being-for-others) as well as a commitment to act for the sake of his freedom when it is threatened, for example by a threat to the other’s life, his being-as-object, since this exists in ‘indissoluble unity’ with his freedom (Sartre 2003: 311–12). It also involves a willingness to further or at least not to contravene others’ more important worldly ends in so far as these are based on genuine reasons (Sartre 1992: 500–1). And it necessarily includes a desire for others to overcome inauthenticity, for only thus can their relation to others’ freedom, including my own, surpass the objectifying-appropriative level described in Being and Nothingness. Such an affirmative relation to the other’s freedom is possible, contrary to what that text seemed to suggest (Sartre 2003: 430–2; 1992: 6), although I cannot but apprehend the other’s freedom as (at least in part) objectifying me. But ‘this is by no means a fall or a danger as such. It only becomes that when the other refuses to see also a freedom in me’ (Sartre 1992: 499–500). I am only confined in the ‘hell’ of alienation by dint of the other’s free subjectivity if I take the other to have an appropriative or entirely objectifying stance towards me and to be incapable of relinquishing that inauthentic stance (ibid.: 10, 20).

How persuasive is this argument? If it is successful, then Sartre will have vindicated his claim that no actual or possible object can justifiably constitute an ultimate value for consciousness. For if rationality requires that any for-itself should value above all free subjectivity in each of its instances, and if subjectivity (consciousness) is fundamentally and essentially not an object, then a for-itself cannot justifiably affirm anything other as its highest value than what is necessarily not an object.

The parts of Sartre’s argument that are likely to seem most problematic are propositions (iii) to (v’). Premise (iii) has it that, if I value anything for its own sake then I must also non-instrumentally value whatever is its constitutive condition of possibility, and this includes, by premise (i), consciousness to which value manifests or can manifest itself, that is, consciousness that is alive to reasons. Application of the Schelerian universality argument then commits me to value such consciousness not only in my own case, but wherever it can be found.
That argument in conjunction with premise (i) directly licenses the conclusion that I must value reasons-acknowledging consciousness not only in my own case but also in that of others. If anything is a value, then its constitutive condition of possibility is a value, at least in part because of its constitutive role, and this role is a universal in the sense of being indifferent to where or in whom it is instantiated.30

There may well appear to be, at least, a profound and endemic tension in Sartre’s position at this point. What I am rationally committed to valuing is any reasons-responsive consciousness in its constitutive role, that is, ‘lived’ consciousness in its necessarily non-object-like fundamental presence to itself. But if I thus value it on the basis of what I am aware of as one of its shareable characteristics (i.e. its constitutive role), I must ipso facto represent, that is, objectify it. This is why Sartre says that the good is in this case a noema for me, an object-involving intentional content (Sartre 1992: 556). The solution to this recurrent tension in Sartre’s texts clearly has to be that any authentic valuing of another consciousness—of the freedom of others—needs to recognize the essentially inadequate character of one’s objectifying representation of it, which cannot capture how that consciousness is for itself. The representation of another’s subjectivity which gives me my reason—i.e., my representation of its constitutive role—also needs to represent it as necessarily not adequately objectifiable. But while this certainly complicates the analysis of those representational contents, it need not evince a serious problem of principle. Consider the following analogy: we can be committed to removing the disvalue which we see in another’s observable bodily pain, partly on the basis of what we see in another’s observable bodily pain, partly on the basis of what we see, while recognizing—indeed because we recognize—that our third-personal representation of it does not capture its first-personal, felt, dimension. We can now appreciate that Sartre’s oft-repeated dictum that nothing ‘outside’ consciousness can constitute an (ultimate) end for the for-itself is true, by his own lights, in one sense but not in another. In affirming another’s freedom, an authentic agent evidently has an end beyond his own consciousness, but that end is not external to consciousness altogether in the sense in which an object is, nor is it alien to his consciousness in the way in which inauthentically experienced foreign subjectivity is (‘the look’). In affirming another’s freedom, a for-itself only acknowledges what is rationally mandated by its own character as freedom.

Further problems for AEF seem to be in store with premise (iv): an authentic for-itself has to value freedom universally. If this is taken to mean that I am rationally required not just to value reasons-acknowledging consciousness, but to value it equally in each instance, then it does not seem to be entailed by anything Sartre has given us grounds to accept so far. Recall that Sartre suggests in his analysis of inauthentic consciousness in Being and Nothingness that such a consciousness is insensitive to specifically ethical reasons (Sartre 2003: 434n). If we grant this point for now and assume, counterfactually, that we knew in a particular instance that such a consciousness was incapable of a ‘conversion’ to authenticity—a knowledge that, Sartre concedes, is not a priori impossible (ibid.: 475)—then the question whether the free subjectivity of that particular
consciousness would have to be valued by us equally to the freedom of another particular, but authentic, ethical consciousness does not obviously require an affirmative answer. In fact, Sartre himself places authentic consciousness evaluatively above inauthentic consciousness: ‘value of the person who assumes his life or authenticity. It’s the only absolute’ (Sartre 1999: 96). In order, then, for an appropriately modified version of (iv) to be acceptable, albeit still not rationally mandatory, Sartre requires at least a further, contingent, epistemic premise: that we do not know of any inauthentic for-itself that it is inherently incapable of a conversion to authenticity. It should be uncontroversial to accept at least such an epistemic reading of his dictum that ‘a radical conversion of my being-in-the-world ... an abrupt metamorphosis of my initial project ... is always possible’ (Sartre 2003: 486, my emphasis). Read in this epistemic way, and appropriately generalized, this claim seems hard to contest and (iv) emerges as appealing in its light.31

While the premises of AEF are therefore defensible, albeit with significant qualifications, they still seem to fall short of establishing Sartre’s conclusion (v’): a for-itself when authentic must not only value freedom universally, but it must value freedom—more precisely: authentic freedom—‘absolutely’ (Sartre 1999: 96) as a ‘supreme value’ (ibid.: 111). It must acknowledge that the value of free subjectivity is in each case greater than the aggregated value of anything that is not itself an authentic freedom, or capable of such, and that it is in this sense absolute.

One putative path to this conclusion, familiar from Christine Korsgaard’s Kant-inspired and in some respects structurally similar argument, relies on the idea that intrinsic value—in the sense of: value that something has non-relationally by or in itself—is possessed only by rational (‘free’) consciousness, from which it would follow trivially that nothing else can compete in intrinsic value with rational consciousness (Korsgaard 1996: 257–73). It should be clear from the discussion so far that Sartre would reject this line of argument. In many cases, including those which are primary in the epistemology of value, i.e. in the ‘order of knowledge’, value is a feature of phenomenal objects—of observed or envisaged actions, of bodies, of artworks, of natural landscapes, and so forth. While the value of an object is necessarily a correlate of consciousness and therefore dependent on it, it is not dependent on any one individual consciousness, nor is it internal to any or all of the consciousnesses to which it is given, nor can it with phenomenological plausibility be construed as wholly and consciously created or conferred by the consciousnesses to which it manifests itself, even if ‘consciousness’ is used in a broader than the Sartrean sense. It is no more thus conferred than the colours of the object as they are perceived (see also note 16). Moreover, the dependence is mutual: just as there can be no actual objective value without some consciousness to which it manifests itself, so there can be no free consciousness without value-manifesting objects (Sartre 2003: 60); free consciousness is itself relational also in this respect.

Sartre’s own attempts to address the question of the absolute value of authentic subjectivity can largely be found in the surviving War Diaries and in
the Notebooks, both unpolished and incomplete writings. I shall conclude by suggesting a direction which a reconstruction of these ideas might take. One line of thought in these texts starts from the tenet that ‘the source of all value, and the supreme value, is the substantiality or nature of the being which is its own foundation’ (Sartre 1999: 111; 2003: 112–13). This idea is by no means only a motivated error definitive of the inauthentic for-itself: ‘Human reality is moral because it wishes to be its own foundation’ (Sartre 1999: 110). ‘It’s not a question of its seeking any value other than substantiality—if it did, it would cease to be human consciousness’ (Sartre 1999: 112; cf. 2003: 115). In authentic existence, the project of self-founding is therefore not simply abandoned but rather ‘corrected’ and ‘purified’ (Sartre 1999: 112). The authentic for-itself seeks to found itself by reflectively ‘assuming’ (endorsing) both the contingency of its facticity as the condition of its freedom, and its freedom and its choices (Sartre 2003: 547; 1992: 506–7). Clearly this train of thought raises many questions, but arguably its most fundamental problem lies in its very premise that ‘the supreme value is . . . the being which is its own foundation’. This surprisingly traditional idea, with precedents going back to Aristotle, is certainly far from self-evident. We may rightly ask why an existent that has no foundation outside itself, or one that seeks to be self-founding, should by virtue of that characteristic alone be of greater value than an existent which owes its existence to another or one that does not aspire to be self-founding.

A second, in my view more promising approach, which in Sartre’s expositions is usually entangled with the one just sketched but is (contrary to his protestations) logically distinct from it, relies not on the ontological concept of self-founding, but on the phenomenological idea of completion, understood as the experiential absence of lack (Sartre 2003: 111–15). The for-itself’s constitutive end-directed, projective character implies that it is always aware of its current situation as lacking or deficient. Any conscious pursuit of a project necessarily entails such an (often unthematic) awareness of an undesirable lack, even if the lack in question is only the envisaged possible absence of a good currently present which needs to be actively safeguarded against future loss or destruction. The for-itself’s projective structure is tantamount to the ever-renewed endeavour to overcome or remove an undesirable conscious lack, a deficiency or incompleteness. Its ultimate implicit end is, therefore, to surpass the incompleteness which haunts it as the necessary correlate of its end-directed, projective character. What the for-itself ultimately desires, although it may deny this in bad faith, is a ‘coincidence with itself’, implying a removal of desire, understood as the awareness of lack-to-be-overcome. In its primary (i.e. necessarily temporarily prior) inauthentic manifestation, this ultimate end is implicitly taken as realizable and as involving the subject’s self-transformation into an in-itself-for-itself, a fully self-sufficient conscious being. Both in its authentic and its inauthentic modes, the for-itself is oriented towards an ultimate value that exceeds any possible fully object-like being. Due to Sartre’s argumentatively under-motivated ontological construal, noted above, of the desire for completeness or wholeness as a desire for necessary, ontologically self-sufficient being
(ibid.: 586–7), he concludes that the end of this desire is necessarily unrealizable (ibid.: 114–15). Nevertheless, in the context of the alternative, phenomenological interpretation presently under discussion, he also maintains that the conversion to authenticity does not involve the simple abandonment of that desire but its ‘purification’: ‘authenticity, I think, conduces to reserving [the unrealizables] a place around us as unrealizables. We must neither deny them nor vainly seek to realize them, but assume them as unrealizables’ (Sartre 1999: 199; my emphasis). ‘Subjectivity finds its meaning simply outside itself in this Good, which never is and which is perpetually to be realized’ (Sartre 1992: 556, also 551; 1999: 111). Staying within the strictly phenomenological dimension of Sartre’s analysis, we might say that what constitutes the absolute value of authentic subjectivity is its orientation towards an experiential wholeness or completeness, of oneself and others, which it grasps simultaneously as contingently never realizable. This orientation itself constitutes its distinctive worth. A fruitful further critical engagement with the early Sartre’s ethics would need to begin by addressing that thought.35

Peter Poellner
Department of Philosophy
University of Warwick
UK
peter.poellner@warwick.ac.uk

NOTES

1 Sartre’s works are cited from the translations specified in the References, which I have sometimes modified.

2 For a nuanced recent assessment of various criticisms of Sartre’s purported theory of freedom, see Webber 2009: ch. 5. Webber shows many of these criticisms to be misguided, but concedes that the theory of radical freedom remains a liability for Sartre’s philosophy, since he takes it to imply the absence of reasons at least at the level of our basic projects (2009: 66–7). I shall defend an alternative reading in what follows.

3 Sartre distinguishes ontology and metaphysics (2003: 639), but the distinction is irrelevant in the present context.

4 For another reading that stresses responsiveness to reasons as central to Sartre’s conception of freedom, see Thomas 2011. Thomas construes this idea in terms of openness ‘to the demands . . . of the world’ (ibid.: 168) which, however, is not an ‘openness to [independently existing] values’ (ibid.: 171). Both claims are suggestive but require clarification to permit assessment. See Sections 2 and 3.

5 See Poellner 2003 for more extensive discussion of (2).

6 Some clarifications of this definition are needed. What is a basic attitudinal characteristic of an intentional experience? I take it that such basic attitudes—modes of intentionally relating to a content—include at least perceiving, (non-committally) imagining, recollecting, and willing/choosing, but this list makes no claim to being exhaustive. In order for an item to be an intentional object for me, is it necessary that I should be able to adopt all the basic intentional attitudes to it which can intelligibly be taken towards it?
Perhaps not. In any case, an affirmative answer is not needed as a premise for the Sartrean arguments outlined below. But we should have justified doubts about whether someone had an objectifying representation of some perceptible item X if he claimed to be only capable of representing X when actually perceiving it (cf. McDowell 1994: 57–9), or if he declared that he can only imagine X but would not be able to recognize it when perceiving it.

Sartre normally uses the Husserlian term ‘noema’ for intentional content. The noema (strictly: the noematic sense) of an experience is the object as it appears in it—the object in its mode of presentation. For Sartre as for Husserl, the noematic sense in its fundamental mode of givenness—when it is merely grasped rather than itself thought about as a sense—is not an object ontologically distinct from the object of the experience.

I shall bypass the question of the sense and reference (if any) of the first-person pronoun in these sentences. See Longuennesse 2008 for discussion of Sartre’s view on this.

Sartre acknowledges the possibility of a non-human for-itself but shows no interest in this (2003: 306).

For a view of this kind, see Searle 1984: ch. 3.

See James 1983, also Sprigge 1970: 288–96. Sartre seems to advance a view of this sort in 1999: 39. I shall return below to the question of what else is required for an end to be embraced by an agent as his.

Sartre explicitly acknowledges his indebtedness to Scheler’s approach in 1999: 88.

Intrinsic ends need to be distinguished from intended outcomes. If I relieve someone else’s hunger for the sake of finding favour with God, or of being morally good, then the relief of the other’s hunger is an intended outcome but not an intrinsic end of my action. Actions can of course themselves be ends, when they are performed for their own sake. For Scheler’s arguments to show that the concept of ends cannot be used to elucidate the idea of value but rather requires analysis partly in terms of it, see Scheler 1973: 30–44.

Velleman 2000 rejects such a link between reasons for action and what is taken to be valuable in some respect by the agent. For criticisms of Velleman’s view to which Sartre would be sympathetic, see Alvarez 2010: 82–5.

Evidently not all emotions are ‘ways of discovering the world’. In Sketch for a Theory of the Emotions, Sartre at length analyses various emotions as pre-reflectively purposeful attempts to misrepresent the world in situations of difficulty, analogous to the sour-grapes response. But even in that text he recognizes another ‘main type’ of emotion which is not purposefully distorting in this way, but whose evaluative content is ‘motivated by the object itself’ (Sartre 2002b: 57). I agree with Richmond 2011: 153–5, that this second type of emotional experience plays a much more central role in Sartre’s thought than is acknowledged in the Sketch, not only in his account of being-for-others but in the foundations of his phenomenology of value.

However, an affective perception can be reasons-responsive without being veridical. The illusoriness of the content of an affective perception does not entail the latter’s insensitivity to reasons, just as the illusoriness of the content of a sense-perceptual experience does not entail this. How the evaluative content of emotions could be veridical is an issue I cannot explore here. It is discussed with respect to some classes of emotions in Poellner 2007: 235–48. For other recent defences of a broadly Schelerian-Sartrean construal of some emotions as value perceptions, see Johnston 2001, and Doering 2007. As Sartre emphasizes (contra, e.g., Korsgaard 2009: 112, 118, 122), there is no good reason to think that all human emotions prior to reflective control are regis-
terings of an object’s conduciveness or lack of such to human interests or needs describable without reference to non-instrumental values external to the agent (Sartre 2004b: 17–18).

17 To repeat, not all emotions rationalize. One type of circumstance where human affective life is not sufficiently alive to reasons is when the subject is in the grip of genuine (non-purposive) phobias or compulsive affects, or of ‘personality disorders’ such as severe depression. We are correspondingly less inclined to consider people responsible when we believe them to be subject to such conditions. I suggest that Sartre, consistently with the general structure of his theory, should concur with this intuition. Sartre does not discuss the bearing of such conditions on his theory of freedom in a systematic way, although he considers purposive, free analogues of, e.g., depression (‘melancholia’, Sartre 2002b: 43–4). The very fact that he takes such pains to rationalize certain pathologies, attributing a conscious, intelligible purposiveness to them, indicates that, if it turned out that some of them are really non-rationalizable, he would concur that they could in that case not be considered free.

18 To say that the soldier’s fear correctly apprehends a disvalue, and therefore a reason, in the world falls short of saying that that disvalue provides a conclusive justificatory reason for his action. The latter rather depends on the justification of his end and that justification ultimately cannot be perceptual. See Sections 3 and 4.

19 Can’t the soldier’s action be a case of *akrasia*? Sartre does not accept that there is genuinely acratic action in which an agent’s effective desires are not or insufficiently grounded in her valuations. The order of ends a person is committed to at a time is normally not revealed by his reflective beliefs about his commitments, nor by his reflective second-order desires, but by his choice of first-order ends and, partly dependent upon these, by his affectively perceptual responses to a worldly situation (Sartre 1992: 477–8). The purportedly acratic person typically acts against his beliefs about his own evaluative commitments, but his affective responses and the actions they motivate disclose his real (however fragile) commitments at the time of action as being different from the content of those beliefs.

20 Sartre maintains that the consciousness of ends, especially of higher-level ends, is mostly implicit or unthematic (Sartre 2003: 117–19, 483–5), which entails that its content is nonconceptual. For an account of how nonconceptual contents can contribute to a subject’s reasons, see Poellner 2003.

21 In principle, this orientation to the future is of course compatible with seeking to maintain a present condition in the future (see Webber 2009: 68, and Section 4).

22 The disjunction in the text between non-action and reasons-responsive action is not intended to be exhaustive. See the remarks on phobic and compulsive behaviour in note 17. For Sartre’s view on *akrasia*, see note 19.

23 Sartre’s ontological interpretation of the desire for completeness or wholeness that characterizes inauthentic consciousness is deeply problematic, since it implies that inauthentic consciousness is at the basic level not reasons-oriented: an incoherent end is not even a logically possible value, and hence cannot be a genuine reason. There are strong pressures internal to Sartre’s overall position for jettisoning that ontological interpretation (see note 34). Not least among these is its lack of phenomenological credentials. There is no phenomenological warrant for thinking that any adequate existential psychoanalysis of the fundamental project of an inauthentic for-itself would invariably uncover a desire for nothing less than ontological self-sufficiency.

24 For this concept of responsibility as attributability vs. liability to deserved sanctions or rewards, see also Watson 1996.
Note that the last two relative clauses have not yet been defended.

See also Webber 2009: 70–2. The for-itself cannot act on ostensible reasons that it explicitly recognizes to be inconsistent. A complex end that is recognized to be inconsistent cannot be aimed at while that awareness lasts, because it is acknowledged to be logically impossible and a fortiori not to be a possible value. The avoidance of inconsistency among an agent’s explicitly grasped ends is a constitutive feature of the reasons-sensitivity of consciousness. This is also why a for-itself cannot explicitly (reflectively) choose to be or to remain in bad faith (cf. Sartre 2003: 91).

For a detailed account of non-objectifying reflection to which Sartre would be sympathetic, see Zahavi 1999: 181–94.

Does this not entail that lived, non-objectified consciousness cannot be a value? No, for while objectifying an experiential attitude necessarily modifies it, it need not distort its value characteristics.

But the argument would of course still fall short of showing that the for-itself invariably does aim at ultimate ends beyond the object-world.

For a different defence of the requirement of valuing the freedom of others, see Gardner 2009: 195–7. Like the present interpretation, Gardner takes the phenomenology of Being and Nothingness and Sartre’s ethical claims in other roughly contemporaneous writings to be, for the most part, compatible (but see note 34).

For (iv) to be not just appealing but rationally compelling, the epistemic possibility asserted by the suggested additional premise is not sufficient, however. What would be needed would be at least the stronger assumption that any inauthentic for-itself’s conversion to authenticity is really possible, indeed probable.

Lack is not in the first instance grasped reflectively but as a lack in the world (Sartre 2003: 456, 499).

Therefore even the inauthentic for-itself is at least unthematicallly conscious of the impotence of any contingent object to supply what is lacked. The insufficiency of the presence or possession of any possible object as an appropriate ultimate end of the human reality’s fundamental desire for completion is not convincingly shown in Sartre’s account. Motivating this thesis phenomenologically rather than (as Sartre fragmentarily attempts) by way of contestable a priori ontological reasoning, seems to me an important task.

The desire for completion cannot be justified if it is, as Sartre wishes to insist, incoherent. His particular ontological construal of it, which is what renders it incoherent rather than merely contingently unsatisfiable, conflicts with his central thesis that even the inauthentic for-itself is wholly free, since it is to that extent not oriented towards even a possible value. If Sartre were to drop that phenomenologically undemonstrated ontological interpretation, this would enable him to say, consistently, that even the inauthentic (i.e. non-self-transparent) for-itself is fully autonomous if (1) it projects itself towards a good that is not given to it by the world, and if (2) that good is no less rationally grounded than the ends of authentic subjectivity (see note 31). Such a revision would also entail abandoning the claim that inauthentic subjectivity is necessarily insensitive to properly ethical reasons (Sartre 2003: 430–4). But if the structure of inauthentic ends is indeed at the fundamental level incoherent (as well as necessarily unethical), then it is impossible to see how this could be compatible with the inauthentic for-itself being ‘wholly free’ (ibid.: 464) in a non-Cartesian sense, a sense which Sartre is explicitly committed to (Sartre 1992: 517–18).

I wish to thank Fabienne Peter for very helpful comments on an earlier version of this essay.
REFERENCES


© 2012 John Wiley & Sons Ltd