It is central to the self-conception of the Western philosophical tradition that its most ancient, most orienting command is to know thyself. The centrality of this injunction is recognized by both participants and critics. So, for instance, Nietzsche often criticizes this tradition by criticizing this particular project. As he writes in The Gay Science, “the saying ‘Know thyself,’ addressed to human beings by a god, is almost malicious” (§335). In critiquing philosophy’s allegiance to the Delphic command, Nietzsche not only sardonically suggests that the pursuit of self-knowledge is doomed; he is also critiquing the total evaluative orientation and form of life within which that project and self-conception gets its bearings.

This suggests that one way to understand competing philosophical positions on self-knowledge is in terms of what Iris Murdoch calls “rival soul pictures,” by which she meant value-rich conceptions of personhood. Murdoch’s idea was that any philosophical moral psychology is always already an ethics, i.e., it is already taking a position on certain values and conceptions of the good life, bringing some readily into view and obscuring others. Thus, how one thinks about self-knowledge – its prospects and its values – will reflect how one thinks about much else about the human condition.

Much contemporary work on self-knowledge argues that we ought to see the topic not as a narrowly epistemological issue but as embedded in a nexus of related philosophical topics like responsibility, moral psychology, and ethics, such that one's philosophical position regarding the former will ramify and inform one's philosophical position regarding the latter. Self-knowledge emerges as an exercise of human agency, to be understood primarily as the way in which one takes responsibility for one's own mind, rather than the way one knows one's own mind as a kind of extant, inner realm of psychological facts. First-person psychological claims – I believe p, I want q, I intend to do y, I'm afraid of z – are authoritative because it is that person’s responsibility to make up her own mind one way or the other, not because she has special epistemic access to, again, an extant, inner realm of psychological facts. And – these philosophers continue – this presumption of first-person authority crucially informs and supports our interpersonal practices not only of holding each other responsible for judgments and actions, but of
believing each other, speaking with each other as rational peers, taking each other’s word seriously.

This conception of self-knowledge, then, is understood as an integral component of a broader philosophical anthropology, a total, value-rich vision of the human being. The presumption that we enjoy first-person authority with respect to our own minds and our commitment to this authority is a condition of being responsible, where this commitment ramifies throughout our personal and interpersonal lives insofar as our lives are organized around expectations of accountability and responsibility, both for ourselves and to others.

**Self-Knowledge Emerges as an Exercise of Human Agency, to Be Understood Primarily as the Way in Which One Takes Responsibility for One’s Own Mind**

Against this position are those who can be broadly conceived as sceptics about self-knowledge. Some version of such a position can be found in psychoanalysis, and while I think philosophers still have much to learn from Freud and post-Freudians, psychoanalysis has fallen into bad institutional favour (a student of mine reported that her psychology professor told her class to think of Freud “like your drunk uncle,” as in: you’re related to him but don’t have to listen to him). Study of the mind’s unconscious processing and products is now taken to be the proper purview of cognitive science and social psychology (see for example, the ubiquitously cited 1977 paper by Nisbett and Wilson, “Telling More Than We Can Know” and Wilson’s more recent *Strangers to Ourselves: Discovering the Adaptive Unconscious*). For these thinkers, and the philosophers who take their findings quite seriously, human beings enjoy a very limited range of privileged self-knowledge, if any, and, as Nomy Arpaly puts it, “murkiness is the rule, not the exception, in human life.”

From the perspective of those who argue for the validity of the first-person perspective, this latter position seems to advance an alien and alienating conception of personhood, what Victoria McGeer, following Akeel Bilgrami, describes as “a ‘psychiatric model’ of human behavior that replaces the structuring ideal of the responsible agent with the notion of a treatable patient.” The worry is that scepticism about self-knowledge, when understood as scepticism about our first-person authority – our ability to knowingly make up our own minds – forces us to adopt a third-person, objective, and objectifying perspective on our own minds, treating ourselves and others, not as self-conscious, self-determining agents but as complex organisms that can at best be managed. If I am a sceptic about self-knowledge, then it looks like I cannot take your word seriously, since I assume that you do not know your own mind, and I cannot take my own first-person experiences seriously, since I assume that I do not know my own mind. In such a scenario, we would rightly wonder, as McGeer does, how this general sceptical, managerial posture “affects [a person’s] long-term ability to understand her own experiences as manifestations of a stable and coherent persona.” That is, if I can trust neither my own conscious responses to the world nor the self-understanding that guides this, then I seem not only forced into a fundamentally alienated relationship with my own experiences, but might even lose the ability to so much as have such experiences if I cannot integrate them and believe them to be authoritative and mine.

So it can seem as though a conception of the human being as fundamentally, rather than just occasionally, opaque to itself would bring with it serious revisions in our basic form of life and our basic inhabitation of our own minds. But we might wonder if the dichotomy is as sharp as presented, where either we presume the authority and validity of the first-person perspective or we are entirely sceptical of it, where the former is imagined to be the position most consonant with pre-philosophical intuitions and values, and the latter entirely alien. Might there be resources for articulating a “rival soul picture” that emphasizes our opacity without being fully sceptical, austere, and alienating? Can we articulate a moral psychology that emphasizes both our self-opacity and our inescapable, essential self-consciousness?
One of the reasons that emphasizing self-opacity can seem to generate an alienating or inhumane or clinical picture of mind, one in sharp contrast to our own lived sense of ourselves, is that the culturally-dominant advocates of this position come from, and use the jargon of, cognitive science and social psychology, whose perspective and voice nervous philosophers seem to think we would then have to adopt regarding ourselves and others, if we were to take on their position. The language of scientific hypotheses or reports on clinical trials and empirical findings does not tend to be particularly intuitive or welcoming – human beings become just so much data – and it does seem convoluted and alienated to try to re-organize one’s own self-understanding around such scientific findings. This is why, again, it can seem tempting to conclude that emphasizing our opacity can only be alienating.

But we have other, more intimate and more humane resources for thinking about our own opacity, resources that do not just assert that the human being is opaque to herself, but work to convey, from the inside, the experience and significance of being a mind, or a person, that is at once self-knowing and self-opaque. Consider, for instance, the following passage from DH Lawrence:

It’s a queer thing is a man’s soul. It is the whole of him. Which means it is the unknown him, as well as the known [...] Here is my creed [...] This is what I believe: ’That I am I;’ ’That my soul is a dark forest;’ ’That my known self will never be more than a little clearing in the forest;’ ’That gods, strange gods, come forth from the forest into the clearing of my known self, and then go back;’ ’That I must have the courage to let them come and go.’

There are three things to note here. First, Lawrence is writing of his own first-personal experience of his own opacity. He is not reporting from his observations of others, let alone from psychological studies. Instead, Lawrence describes the ways in which opacity can “show up” in the course of experience, in the form of disruptions to the self as it knows and understands itself. By writing from inside of experience, the passage emphasizes both our self-consciousness and our self-opacity (“it is the unknown him, as well as the known”).

So it is not that either we inhabit this perspective credulously or we step outside and regard our own minds as if they were objects. Rather, there are ways of experiencing the opacity of the mind intimately, from the inside, where such disruptions can be disorienting or frightening or exhilarating, or all at once. Lawrence describes not knowledge that the human mind is opaque, but rather gives voice to his experience of this opacity.

The worry is that scepticism about self-knowledge forces us to adopt a third-personal, objective, and objectifying perspective on our own minds.

Second, Lawrence gives voice to this first-personal experience, and in turn prompts the reader’s sympathy with, or memory of, such experiences, by means of aesthetically rich, evocative, and metaphorical language. His language communicates the mysteriousness of having a mind that is not totally under one’s control and does not fall fully within the purview of consciousness. I might know, from reading psychological studies, that the human mind is opaque and so, by extension, is my mind, but this knowledge is theoretical and abstract. By contrast, Lawrence’s use of literary language does not facilitate theoretical knowledge of the kind produced by psychologists but something more like recognition or acknowledgment.

Finally, the passage suggests that Lawrence has adopted something like an orienting commitment to live in terms of this acknowledgment, i.e., that his mind is unknown to him, where this takes work, or as he puts it, “courage.” This is not the work or courage involved in taking up an objective perspective on one’s own mind, or regarding it as a kind of complex mechanism that you can trick into optimizing. It is instead the work and courage involved in assuming
responsibility for oneself, for one’s mind as one’s own, while also acknowledging its broader obscurity, opacity, and elusiveness. Moreover, it suggests that for Lawrence this is not a matter of being scientifically realistic about the mind or trying to live “in accordance with the facts” (to use Susan Wolf’s phrasing). It is not a case of a philosophical or humanist conception of personhood giving way to the hard and hard-to-bear deliverances of science. Rather, the passage suggests that Lawrence finds this “soul picture” compelling and attractive, mysterious, inspiring, even beautiful, if at the same time, occasionally unnerving and alien (such a picture can be found in Nietzsche, too).

***

As I described this stand-off in contemporary philosophy, it can seem as though retaining a commitment to first-person authority and a general presumption of validity about self-knowledge is the only way to retain a humane and recognizable picture of persons, against a destructive and value-deprived scientism, which proffers an alien and revisionary picture of mind, one seemingly quite remote from our ordinary sense of ourselves and others. What the passage from Lawrence reveals is that there can be more or less humane, more or less sympathetic, more or less value-rich articulations of opacity.

In fact, to take an example from philosophy, Murdoch – who was also a novelist – also advocated for just such a humane, sympathetic, value-rich, and opaque conception of personhood. Against the philosophical ideal of deliberative agency, according to which we should strive to become “perfectly self-aware and so perfectly detached and free,” Murdoch advocated
in *The Sovereignty of Good* for what she describes as “a darker, less fully conscious, less steadily rational image of the dynamics of the human personality,” as well as (in her essay “On Dryness”) for a renewed and deep appreciation of “the opacity of persons.” And in league with Lawrence, Murdoch herself argued for a specifically aesthetic dimension of, approach to, and appreciation of the mind.

**THERE ARE WAYS OF EXPERIENCING THE OPACITY OF THE MIND INTIMATELY, FROM THE INSIDE, WHERE SUCH DISRUPTIONS CAN BE DISORIENTING OR FRIGHTENING OR EXHILARATING, OR ALL AT ONCE**

Not only did she argue that philosophers should attend to an aesthetic dimension of the ethical life, specifically the role of vision – more or less creative, more or less adequate, more or less loving ways of seeing the world – but Murdoch’s own work exhibits an aesthetic, literary dimension, one which invites her reader not only to try to understand Murdoch’s darker, less fully conscious, less steadily rational soul picture, but sympathize with it and find it compelling. That is, the aesthetic dimension of Murdoch’s writing facilitates a non-alienated, non-clinical, sympathetic picture of opacity (and this despite her own puzzling insistence “that there is an ideal philosophical style which has a special unambiguous plainness and hardness about it, an austere unselfish candid style [...] a certain cold clear recognizable voice.” Her own philosophical style in fact speaks against this ideal, or at least shows it to be just one amongst others).

***

Lawrence and Murdoch give us two examples of how a thinker might present a conception of the human being as at once self-conscious and opaque to herself in a way that is not alienating or objectifying or clinical, a conception of the human being that is value-rich and that facilitates the appreciation of certain values (for instance, beauty, mysteriousness, curiosity, compassion, humility) while putting less emphasis on others (for instance, moral judgment, certain conceptions of responsibility). The literary dimension of their language can be understood as a strategy for articulating what it is like to be at once self-conscious and self-opaque, and so to do justice to our murky experience as persons.

By studying the human mind observationally, cognitive science and social psychology can offer powerful characterizations of how the human mind functions, and can indicate how its functioning diverges from its own self-conscious or reflective conception of its operations. But this perspective on the mind remains external to it; it does not reflect the mind’s own perspective but seeks instead to show that perspective to be deluded. What alternative, phenomenologically-rich forms of philosophy and literature have the power to present are characteristic ways in which the mind can register and experience its own opacity, as well as the values of those experiences. Lawrence, Murdoch, and Nietzsche each suggest in their own ways that the mind’s opacity can be conceived, not (always) as regrettable, and not (only) as a mere psychological fact, but as part of what Bernard Williams describes as “that worthwhile kind of life which human beings lack unless they feel more than they can say, and grasp more than they can explain.”

Francey Russell is Assistant Professor in Philosophy at Barnard College and Columbia University. Her research focuses on issues in moral psychology and ethics broadly construed, overlapping with topics in social philosophy and aesthetics (especially film). Website: franceyrussell.com