Describing Ourselves

Wittgenstein and Autobiographical Consciousness

GARRY L. HAGBERG

DESCRIBING OURSELVES

This page intentionally left blank

Describing Ourselves

Wittgenstein and Autobiographical Consciousness

GARRY L. HAGBERG

For Julia

This page intentionally left blank

What is good and evil is essentially the I, not the world. The I, the I is what is deeply mysterious!

Ludwig Wittgenstein
Notebooks, 1916

The total speech act in the total speech situation is the *only actual* phenomenon which, in the last resort, we are engaged in elucidating. Stating, describing, &c., are *just two* names among a very great many others for illocutionary acts; they have no unique position.

J. L. Austin
How to Do Things with Words

There is a picture of the mind which has become so ingrained in our philosophical tradition that it is almost impossible to escape its influence even when its worst faults are recognized and repudiated.

Donald Davidson
'Knowing One's Own Mind'

This page intentionally left blank

Acknowledgements

AT a time when the idea for this book had already been in the back of my mind for some years, a most welcome opportunity to return to St John's College, Cambridge, arose. I am grateful to that institution for once again having provided an ideal context for work on Wittgenstein, and it was there that I drafted the initial core of this study: the material examining that part of Wittgenstein's philosophy particularly concerned with self-referential and self-revelatory language. As if providentially from the point of view of the development of this book, a succession of invitations to write for various collections and to present material at conferences, colloquia, and philosophical and literary meetings followed. These projects and occasions allowed me—well, as any academically peripatetic, deadline-fearing author knows, this also carries in its connotative substrate more than a whisper of the words 'forced me'—to extend the material from that core manuscript into each of the subjects explored in the chapters now reunited in this volume. I am deeply grateful to all of the scholars named here for their sustained efforts in bringing about these events. (Readers less concerned with the process behind the product may safely turn to p. 1!)

I had wanted for some time to bring together a reconsideration of Wittgenstein's remarks on consciousness with particular cases of autobiographical writing, and Chapter 1 emerged as my response to John Gibson and Wolfgang Huemer's invitation to write for their edited volume *The Literary Wittgenstein* (London: Routledge, 2004). Parts of my chapter for them, 'Autobiographical Consciousness: Wittgenstein, Private Experience, and the "Inner Picture"', enjoyed (as did I, to say the least) the benefit of public presentation, in a visiting speaker series at the University of Erfurt, where Thomas Glaser provided an acute and helpful commentary.

The title of Kenneth Dauber and Walter Jost's edited collection *Ordinary Language Criticism: Literary Thinking after Cavell after Wittgenstein* (Evanston, Ill.: Northwestern University Press, 2003) accorded perfectly with a desire I had to reconsider a frequently misunderstood idiom of philosophical work in connection with autobiographical issues. My piece for them was 'The Self, Reflected: Wittgenstein, Cavell, and

the Autobiographical Situation', and it provided much of the content of Chapter 2.

Strong encouragement for the work leading to Chapter 3 was provided by Jean-Pierre Cometti, who invited a piece for a special issue of *Revue Internationale de Philosophie* (1 (2002), no. 219) entitled *Wittgenstein and the Philosophy of Mind*. My project for him, also presented in draft form at an unforgettable conference he organized entitled 'Wittgenstein, Language, and Perception' at the University of Aix-en-Provence, was published as 'The Self, Speaking: Wittgenstein, Introspective Utterances, and the Arts of Self-Representation'.

Seeing throughout the history of philosophy a number of misconstruals and simplifications—misleading pictures—the mind has made of its own workings, and seeing the force of Wittgenstein's undercuttings of those pictures, I wanted to reexamine some carefully selected remarks in connection with a particular dualistic picture of autobiographical self-investigation. Peter Lewis was kind enough to ask for a contribution to his collection *Wittgenstein, Philosophy, and the Arts* (Aldershot: Ashgate, 2004) and flexible enough quickly to accept and encourage my idea for a paper; this resulted in 'The Self, Thinking: Wittgenstein, Augustine, and the Autobiographical Situation' and has become the first two sections of Chapter 4. It was presented as the keynote address at the annual Building Bridges Conference (this year the bridge connected philosophy and literary studies) at the University of Illinois at Carbondale, where I was graciously invited by Christopher Nelson; another version was presented to the annual meeting of the Canadian Society for Aesthetics in Quebec City at the invitation of Bela Szabados and Alex Rueger. I must note that I am particularly indebted to Bela for his long-standing encouragement of (not only) this project and for his foundational article in the field on Wittgenstein and autobiography, as well as for his subsequent invitation to give another part of this book at a later CSA meeting in an extended session on Wittgenstein, this time in Halifax. The third section of Chapter 4, 'Wittgenstein Underground', was published as part of the symposium 'Dostoevsky Recontextualized', in *Philosophy and Literature*, 28/2 (October 2004). I remain indebted to Denis Dutton for his sustained encouragement of these symposia (and more generally for our stimulating and highly enjoyable ongoing co-editorship of that journal). This piece was presented in early-draft form to the Philosophy Research Seminar at Bard; I was on that occasion—as I so frequently am—grateful to and heartened by my students for their unstopping blend of stimulation, intellectual

engagement, encouragement, and steadfast refusal to assume very nearly anything as given.

When Michael Krausz told me he was editing a collection on the problem of single versus multiple interpretations and that he wanted an essay from me on the significance of Wittgenstein's philosophical writings for this project, I saw at a glance that singular versus multiple *self*-interpretation might be my subject, an idea he immediately fastened upon with his characteristic acuity and encouraging warmth. The result, providing here the basis of Chapter 5, was 'Wittgenstein and the Question of True Self-Interpretation' in his collection *Is There a Single Right Interpretation?* (University Park: Pennsylvania State University Press, 2002). Parts of this chapter were delivered, also at his invitation, at a session of the Greater Philadelphia Philosophy Consortium at Bryn Mawr College—another most helpful and thoroughly enjoyable occasion.

Another part of that originally drafted material concerned the perception of persons as investigated with irreducible complexity by Wittgenstein, especially where the question of the distinctive nature of person-perception links to issues of biographical and autobiographical understanding. So in receiving an invitation from Richard Raatzsch to write for a special issue of *Wittgenstein-Studien* (5 (2002)) entitled *Goethe and Wittgenstein: Seeing the World's Unity in its Variety* that he was co-editing with Bettina Kremberg and to present a paper at a thoroughly delightful conference on the subject at the University of Leipzig, I was given a chance to develop that material further. And then Catherine Osborne invited me to present still another version at her engrossing conference 'Wittgenstein, Literature, and Other Minds' at the University of East Anglia. 'The Mind Shown: Wittgenstein, Goethe, and the Question of Person-Perception' became the foundation for the first two sections of Chapter 6.

Toward the end of the preceding piece I could see that I would need to extend the discussion of person-perception—particularly where this turns recursively to self-perception—into aspect-perception and 'seeing-as'. It was thus my continued good fortune when William Day and Victor Krebs asked shortly thereafter if I might write for their collection *Seeing Wittgenstein Anew* (Cambridge: Cambridge University Press, 2008). I found this invitation especially attractive as it afforded the possibility of reviving and integrating some work on §xi of part II of *Philosophical Investigations* that I had done many years earlier while in residence at St John's College, Oxford; I thus finally have an opportunity

to thank them (if from across a considerable span of time), and particularly Dr P. M. S. Hacker, for having made that productive and stimulating time possible. I also benefited much from some characteristically insightful and helpful advice from Colin Lyas some years back on this topic. The long-delayed precipitate was published as 'In a New Light: Wittgenstein, Aspect Perception, and Retrospective Change in Self-Understanding'. As it was nearing completion, Jerrold Levinson invited me to give a paper to the University of Maryland's visiting speaker series. Presentation once again aided and abetted (as did Jerry), and the end point of this chain of events is, at last, the final section of Chapter 6.

I was also becoming increasingly aware that any study of this kind would need to consider Wittgenstein's too-little-discussed remarks on memory and the significance of these remarks for the clarification of issues pertaining to first-person narratives when David Rudrum invited me to write for his collection *Literature and Philosophy: A Guide to Contemporary Debates* (London: Palgrave, 2005). The project, as it unfolded, became 'Autobiographical Memory: Wittgenstein, Davidson, and the "Descent into Ourselves"', and now constitutes the first two sections of Chapter 8. David also kindly invited me to take part in a conference entitled 'Wittgenstein and Literature' that he organized at the University of London's School of Advanced Study, where some of the governing ideas for this book were aired.

Much of the content of the last section of Chapter 7 was part of another symposium in *Philosophy and Literature* (27/1 (April 2003)), entitled *Wittgenstein and Literary Aesthetics*. It appeared there as 'On Philosophy as Therapy: Wittgenstein, Cavell, and Autobiographical Writing', and versions of it were presented in San Francisco as part of a session, 'Wittgenstein and the Arts', at an annual meeting of the American Society for Aesthetics, and at the Seventh Annual Comparative Literature Conference at the University of South Carolina, which was devoted to the work of Stanley Cavell.

Sections of this study were also given at a number of other philosophical conferences, and I want to thank, if too briefly, the organizers and the participants of these for their generosity in making it possible to give various pieces of this book hearings in their formative stages. David Goldblatt extended a generous invitation to speak in the visiting philosopher series at Denison University, Ohio, where animated discussion with him and others proved most helpful. John MacKinnon invited me to give part of this book at a spirited conference entitled 'The Complementary of Human Perspectives' at the Institute of Humane

symposiasts, audience members, and late-night bon vivants (these, happily, are not exclusive categories), and I send my heartfelt thanks to them all. Every book—or so I imagine—is, also like persons, in a manner of speaking a palimpsest, and this one bears every kind of mark, ranging from direct, strong, plainly evident influences to the slightest under-layered traces of the encounters and experiences recounted above. And despite the passage of many years the sense of indebtedness to earlier teachers and advisers has not dimmed; it would thus, for me, be gratifying if traces (or more) from those earlier years were discernible in this book. Out of a longer list, I must mention, in connection with this project and its philosophical aspirations, Henry Alexander, Renford Bambrough, Frank Ebersole, and, going back to my earliest formative influences, John Wisdom. In more recent years I have invariably learned much from (if not invariably agreed with) the distinctive philosophico-critical writings of (and from some helpful and encouraging conversations in various contexts with) Stanley Cavell, Arthur Danto, Richard Eldridge, Martha Nussbaum, and Richard Wollheim (among others far too numerous to mention here, alas—although many do appear in the text or footnotes). The one solitary part of this entire undertaking is this: its remaining shortcomings, at all layers, are my doing alone.

Lionel Trilling entitled a collection of his *A Gathering of Fugitives*, which was his way of naming a book that brought together a number of independent pieces each written for a particular separate occasion or collection. Although this book does bring together the pieces as indicated in the foregoing, it is not that kind of affair. It is, rather, something of a reunion of accomplices, and I owe it to Peter Momtchiloff and Oxford University Press, to the Press's two extremely helpful anonymous readers (one later emerging from the darkness as none other than John Gibson, whom I particularly thank again for now having done even more for this project), and to Kate Walker and Laurien Berkeley (to whom I am especially indebted for discerning and sensitive copy-editing) that this reunion has been a particularly enjoyable, philosophically helpful, and productive one. Peter, John, and the still-anonymous reader will see herein just how much I owe to their acumen, good advice, and judgment.

Bard College is an institution to which I am greatly indebted, and it continues to be a remarkable place for work in aesthetics (in fact increasingly so). It is with a particularly deep gratitude that I want to thank James H. Ottaway, Jr., and Leon Botstein for creating a new endowed

This page intentionally left blank

Contents

Introduction: Confronting the Cartesian Legacy

The voluminous writings of Ludwig Wittgenstein contain some of the most profound reflections of our time on the nature of the human subject and self-understanding—the human condition, philosophically speaking. Yet the significance of his writings for the subject (in both senses) can far too easily remain veiled. One of my aspirations throughout this study has been to help clarify that significance, while at the same time assessing and exploring the multiform implications of those writings for our understanding particularly of autobiographical (and more generally, self-descriptive) writing and thereby of the nature of the self and self-knowledge. Any such attempt to unveil significance of this self-reflexive kind—that is, of a kind that should prove central to reconsidering a nested set of beliefs concerning the self, self-knowledge, and self-understanding that are foundational to moral psychology—requires our going beyond Wittgenstein's texts into actual autobiographical practices. For this reason this study contains fairly detailed discussions (which I would like to think of as one kind of philosophical criticism) of: philosophers writing as autobiographers (including Augustine and Iris Murdoch); a number of autobiographers whose writings, once seen in this context, are clearly philosophically significant; philosophers whose philosophical writings are themselves intrinsically autobiographically significant (including Schopenhauer, Kierkegaard, Stanley Cavell, and Donald Davidson); and literary figures whose writings cast distinctive light on the self and its descriptions (including Goethe and Dostoevsky, among others).

In a moment I will say a bit more about what is to follow throughout these chapters, but if I were to enumerate the fundamental aspirations of the undertaking, they would thus include these interlocking attempts:

1. to mine Wittgenstein's later writings (and then to extend the discussion well beyond those writings but along discernibly Wittgensteinian

lines) for an account of the self of a kind that stands in striking, indeed revolutionary, contrast to the initially intuitively plausible alternatives;

2. to assess the significance of some of Wittgenstein's later writings on language and mind for our understanding and clarification of particularly self-descriptive or autobiographical language;

3. to turn to autobiographical writing as a valuable and heretofore little-explored resource for the philosophy of literature (taking these writings, themselves the best examples we have of human selves exploring themselves, in the light of issues in the philosophy of language and mind);

4. to reconsider in a new light Wittgenstein's multifaceted critiques of Cartesianism (on Cartesianism, see below), seeing in (and again, beyond) them a powerful way of clarifying the problems of autobiographical consciousness;

5. to see the self, if not as an inner entity we can explore through dualistically construed introspection, then as it is manifest in action (of both the word and deed types), but in such a way that the eviscerating reduction of behaviorism in both letter (the easier part) and spirit (the harder part) is avoided and where first–third-person asymmetries are acknowledged (if in nontheoretical, irreducible form);

6. most broadly, to take in turn the issues of self-consciousness, mental privacy, first-person expressive speech, reflexive or self-directed thought, retrospective self-understanding, person-perception and the corollary issues of self-perception (itself an interestingly dangerous phrase), self-defining memory, to bring these into (I hope) mutually illuminating contact with each other, and to develop a Wittgenstein-inspired account (I am being *very* brief here: a better term than 'account' might be 'conceptual clarification') of each; and

7. to help show, over the book's course, some small part of the value of interweaving questions of subjectivity and selfhood with both autobiographical and autobiographically significant writings on the one hand and a therapeutic, nonscientistic conception of philosophical progress on the other.

I should say at the outset that, consistent with widespread philosophical practice, in this study I use the term 'Cartesian' to refer to a cluster of intertwined metaphysically dualistic views in the philosophy of mind

and language. Their precise articulations will follow chapter by chapter but, briefly and roughly stated, they include the views (*a*) that the self is most fundamentally a contingently embodied point of consciousness transparently knowable to itself via introspection, (*b*) that its contents are knowable immediately by contrast to all outward mediated knowledge (and that self-knowledge is thus non-evidential), (*c*) that first-person thought and experience is invariably private, thus presenting as a brute first fact of human existence an other-minds problem, and (*d*) that language is the contingent and *ex post facto* externalization of prior, private, pre-linguistic, and mentally internal content. It has in recent years been argued that, as it has been memorably put, Descartes's Big Mistake occurred in the mid-twentieth-century. That is, anachronistic readings have retroactively converted him into what we now call a Cartesian, when in truth he was no more a Cartesian in that sense than, say, Freud was a Freudian (in the terms of what that has come to mean since his original writings) or even than Marx was a Marxist. So for the purposes of this study, my use of 'Cartesian' will refer to that cluster of metaphysically dualistic views, and not necessarily (although I do think occasionally) to the views explicitly endorsed by that historical figure.[1] Of course, the grip, the culture-deep initial intuitive plausibility of those dualistic views, in any case very much pre-dates Descartes as much as they outlive him, so to show that these dualistic views, or some of them, were not *his* explicitly endorsed positions, however historically interesting, is not at all to show that the views and positions contemporary philosophy debates under that heading have therefore evaporated. In this respect Gordon Baker and Katherine Morris, in their book *Descartes' Dualism*, take a particularly helpful line: to come to see that these views are not ones advanced by Descartes can help to revivify a sense of how strange, alien, or prismatically distorting of human experience these philosophical pictures of selfhood in fact are, i.e. this is itself one way to change radically, therapeutically, our point of view, our way of seeing, these problems. For a helpful and historically informed survey of the broadly Cartesian position, see Charles Taylor, *Sources of the Self*.[2] On the anticipations of Cartesian views in the writings of others, see Taylor's opening remark on Descartes: 'Descartes is in

---

[1] See Gordon Baker and Katherine J. Morris, *Descartes' Dualism* (London: Routledge, 1996), and Desmond M. Clarke, *Descartes's Theory of Mind* (Oxford: Oxford University Press, 2003).

[2] (Cambridge, Mass.: Harvard University Press, 1989), esp. 'Inwardness'.

many ways profoundly Augustinian' (p. 143), where what is centrally intended is 'the emphasis on radical reflexivity' and 'the importance of the cogito'. The self of the Cogito is, of course, a self necessary for the coherence of the deception that Descartes's universal doubt posits (for deception to occur there has to be a deceived), and that self is knowable unto itself independent of any other (external) claim. So, while it may not capture all and only Descartes's explicitly articulated views, the term 'Cartesian' still—even with these welcome and salutary concerns regarding anachronism—seems not utterly wide of the mark either. In any case, it is in this study the cluster of views that go by a well-entrenched name that is the focus, not historical attribution.)

In the preface to John Updike's collection of critical writings *Hugging the Shore*,[3] he explains his title by suggesting that literary reviews, because they stay close to the texts they are criticizing and do not sail out into the open sea of fictional creation, 'hug the shore'. Part of the discussion here, in that sense, hugs the shore of Wittgenstein's texts and then of autobiographical or other philosophical or literary texts in turn, trying in each case to disclose what is particularly helpful in them for the achievement of a perspicuous and comprehensive view of first-person or self-revelatory speech, thought, and expression. But what we see along each of these shorelines is not, as I hope becomes increasingly clear as this study progresses, transparently evident upon simply looking at it. On the contrary, what we are enabled to see along one of the two shores, that is, either in Wittgenstein's philosophy on the one side or in autobiographical or literary texts on the other, is powerfully shaped by what we have just seen—or more accurately (for reasons that emerge as the book advances) by what we have *said* about what we have just seen—on the other. Then, of course, other parts of the book sail into open waters.

In Chapter 1, I initiate a philosophical project central to the entire book and that continues, with increasing specificity, throughout it: unearthing a number of powerful but nevertheless often undetected influences on our thinking of conceptual pictures, or simplifying theoretical templates, in particular the fundamental pictures of selfhood that encourage correlated models of self-knowledge and especially of autobiographical self-investigations. Freedom from such pictures promotes conceptual clarity, which itself is a result of an acceptance of, or an openness to, complexity and particularity. The chapter begins with a reconsideration of the Schopenhauerian elements in Wittgenstein's

[3] *Hugging the Shore: Essays and Criticism* (New York: Alfred A. Knopf, 1983).

early thinking about the self, followed by a close consideration of some of the remarks showing his struggle against, and ultimate freedom from, those early, theoretically neat, simplifying templates. Wittgenstein came to see what he called 'the inner picture' as a source of a great deal of philosophical difficulty and confusion, and in the second section of this chapter I look at his own analysis of the cognitive forces, or pressures of thought, that buttress the traditional Cartesian conception of selfhood. In the third section of this first chapter I turn to cases of autobiographical writing, showing something of the gulf that separates our picture-driven ways of theoretically construing autobiographical self-investigation from actual autobiographical practices. And this permits a glimpse of the great difference between real autobiographical privacy and the philosophical misconstruals of first-person privacy.

The second chapter begins with a reconsideration of the very idea of observing consciousness and the distinctive picture of introspection that this idea can easily enforce. 'Introspecting' is a word that carries the concept, indeed the word, of 'inspecting' within it, and the act of inspecting requires an object of inspection. With that conceptual linkage we are all too quickly bound up with notions of the self as viewer of inward objects, and consequentially with introspective language being *descriptive* (carrying, as we shall see, distinctively philosophical implications) language. But a close look at Wittgenstein's remarks pertaining to this subject breaks this linguistically induced spell, and the second section of this chapter turns to the picture of metaphysical isolation engendered by this line of thinking, along with the correlated conception of autobiographical truth as verified correspondence between inner object and outward description. The third section of this chapter turns to some contributions Stanley Cavell has made to our understanding of the pressures that would lead us, seemingly inexorably but only falsely so, into this line of thinking. Here, telling asymmetries between the first- and third-person cases emerge, along with a deployment of a distinction between the metaphysical voice and its ordinary counterpart of the kind we will have encountered in Chapter 1 in connection with autobiographical privacy. And here the fundamental idea of self-narrative comes to the fore, an idea that will be examined in ever-closer detail throughout subsequent chapters.

A conceptual undertow can swiftly and powerfully drag us back into a way of thinking of the self and its description deeply aligned with Cartesian or dualistic metaphysics, and it can do this in ways that are not entirely obvious on the surface. One less obvious way

of staying within the template of dualism has been to argue directly against the inner half of the inner–outer picture. Behaviorism is, as we see in the first section of Chapter 3, such a position, and it has on occasion proven difficult to distinguish Wittgenstein's position from behavioristic reductionism. But in this section we see why he is not what he called 'a behaviorist in disguise', why the first-person case cannot be assimilated to the third-, and why the language-games of our mental vocabulary do not permit reduction to the language-games of physical objects. We also see here why the perception of personally expressive gestures is not, against what the inner–outer template and the metaphysics of isolation would suggest, *inferential* (a subject to which we will return in greater detail in Chapter 6). Behavior is misconstrued as evidence in the vast majority of cases (where, that is, we are not looking for evidence, or signs, because of a particular context-specific suspicion), and seeing this, along with gaining a grasp of the noninferential character of our perception of emotional states, helps to free us from the tyranny of a dualistic self-concept. But then how do we characterize—if we characterize them *generally* at all—our first-person reports on what we call inner states? In the second section of this chapter we excavate and then scrutinize the presumption implicit in the preceding sentence, that is, that such language is itself rightly described as a matter of *reporting*. Wittgenstein shows that the matter is, instructively, not so simple or direct; the philosophical grammar of expression of states such as pain are not innocently construed on the model of inner object and outward designation. Our language of this kind, as it emerges under closer investigation, is not best characterized as descriptions, but rather—again if we want a kind of shorthand or generic category—as avowals. But then this makes us ask: If the matter is not successfully characterized in terms of descriptions (where this term imports metaphysical freight), how do we understand the acts of introspection upon which autobiographical or self-revelatory language would so evidently seem to depend (given that reductive behaviorism will by then, I hope, have been moved beyond the bounds of plausibility)? In the third section of this chapter we thus progress to a study of introspection of a kind neither engendered by nor supportive of dualism, or introspection that, in the manner of privacy as introduced in Chapter 1, is real, i.e. drawn from—or better, shown in—our practices. And it turns out that Kierkegaard's 'Diary of the Seducer' is of great value in this respect: duplicity is not dualistic, and an inner secret is not metaphysically hidden.

Yet thinking—or our image of thinking—seems to require some residual form of 'mentalism', some conception of the self as private interior consciousness, where thinking of the self's experience, intentions, hopes, fears, regrets, aspirations, and so forth just is autobiographical reflection. In Chapter 4 I consider both the influences on our thought that lead us to picture the act of thinking in a decidedly dualistic way and those remarks on thinking of Wittgenstein's that can powerfully reorient our thinking about thinking. This reflexive analysis, really a layered diagnosis, looks into the way the mind tends to imagine its own workings, and the word 'thinking' turns out to be better understood as a particular tool in our language than as the name of a unitary mental event. And then turning to cases here as well, in this chapter's second section we see in Augustine's magisterial self-investigation a range of practices that, taken together as the raw material for an overview of self-directed thought, show autobiographical language to be far more diverse—and more interesting—than the picture of self-revelatory language as outward one-to-one linguistic correspondences of inward thought would begin to suggest. Augustine's practice shows that the relation between what we call a thought and what the metaphysical voice might generically classify as a proposition is anything but direct and immediate. And here, in the third section, it is Dostoevsky's underground man who helps show a further expanded set of practices that we would without question regard as self-directed thinking, but where this self-investigation shows a self positioned in relation to his remembered past, to his present self, and to his present utterances not with a transparent immediacy but rather with a layered complexity. The issue of speaker's privacy—the distinctive relation to our own language that no one else does or could have—resurfaces here, and we see again, for deepened reasons, that we need to 'de-psychologize' our conception of first-person speech and writing in order to accommodate the interesting—and from the Cartesian point of view, very ill-behaved—facts of the case. Indeed our language, construed as merely contingent *ex post facto* expressions of prior determinate mental events, or really as an afterthought, can only further mystify the autobiographical processes that fall on a continuum between self-revelation and self-constitution (a matter also to be taken up in greater detail in subsequent chapters).

The contest between interpretive singularism (i.e. the view that there is only one correct interpretation) and multiplism (i.e. the view that there can be, in its weaker form, different, and in its stronger form, incompatible, interpretations) has concerned, primarily, the interpretation

of cultural objects and culturally emergent practices, e.g. works of art, societal conventions and rituals, and so forth. In Chapter 5 I attempt to extend the discussion of the contest into the area of self-interpretation. Here I find those (still too-little-discussed) passages of Wittgenstein pertaining to retrospective meaning, i.e. the questions concerning what we did and did not mean on a given occasion, particularly helpful in excavating the metaphysical presuppositions embedded within stand-ard ways of framing, or indeed picturing, the very problem of true self-interpretation. In order to understand the (sometimes misleading) motivations for favoring an interpretive singularism with regard to past first-person utterances, we need to examine the belief or intuition that there is a determinate mental event that constitutes meaning something and that this mental phenomenon of meaning itself requires the prior existence of an inner locus of consciousness, an inner Cartesian self that is the private sphere within which the act of meaning occurs. But Wittgenstein's investigation into this alleged phenomenon, which runs parallel to his better-known investigation into the very nature of think-ing itself (as examined in Chapter 4), shows that meaning something, on close examination, is found not to be at all what we expect when coming to the subject with certain philosophical expectations in mind. Here it emerges that the subject does not reduce to a single, uniform mental act, process, or state, and that various phenomena, not a single phenomenon, are (perhaps surprisingly—given their power to unsettle our picture-driven presumptions) relevant to the determination of retrospective meaning. In particular, the metaphysically misled notions of having meant something as (1) an easily remembered process or state, (2) a process that follows a course and upon which we can report, (3) a mental picture constituting the determinate thing that we mean, (4) an act of stipulation, (5) a focused directing of inward attention or the inner referent upon which we concentrate, (6) an act of inner 'pointing' modeled on outward ostension, and more generally (7) the ineliminable essence required for the words 'meaning something' to themselves mean, are all removed as candidates for the explanation of retrospective meaning. Yet it is often incontrovertibly true that on a given occasion we meant one thing and not another, and this blunt fact persistently argues against the adoption of a generalized (and given the foregoing considerations, de-psychologized) interpretive multiplism.

Thus, with Wittgenstein's observation that guessing at how a word (like 'meant') functions will not yield valuable philosophical results and that the necessary task is to 'look at its use and learn from that', along

with a brief look at the distinctively active or 'mind-making' (as opposed to passively 'mind-reporting') character of self-reflection that Richard Moran and others have helpfully brought into focus, I turn back to the actual case of Augustine catching himself in the acts of retrospective self-interpretation. And I try, indeed, to learn from that. Among the various lessons I draw from this example (in addition to the more general one that one kind of project in literary interpretation can itself constitute the work of philosophical investigation), I suggest that the removal of prismatic metaphysical expectations itself constitutes indispensable progress toward a perspicuous, post-Wittgensteinian, highly context-sensitive and pragmatically situated understanding of retrospective meaning in particular, and more generally the nature of the linguistic self that is the subject of interpretation in the first place. And at this point we will be much better positioned to fathom the competing pulls, sometimes toward self-interpretive singularism, and at other times toward self-interpretive multiplism.

We next turn to this question: Is there a distinctive nature of the perception of persons, of human beings, that is unlike any other mode, or perhaps category, of perception, and if so, what does this tell us about our consciousness of self? Wittgenstein describes this most distinctive kind of perception, memorably, as 'an attitude towards a soul', and in elucidating this concept we see that it will not settle either into the traditional categories of Cartesian or behaviorist models of the self, or into any directly antithetical position advanced against these polemically opposed but category-sharing pictures of selfhood. Wittgenstein's statement concerning the separateness of the language-games of the mental and the physical, and the correlated claim that if we try to characterize generally or theoretically the relationship between them we shall go wrong, proves helpful here—but it also invites the question asking exactly how then we are to understand this complicated relation between language-games. Employing examples in Chapter 6 from both Wittgenstein and, more extensively, Goethe in his writing on the perception of human qualities and mental states in artistic representation, we begin to see something of the value of a fidelity to the nuances of lived experience and the value of an awareness of the circumstantially situated and embodied-yet-irreducible character of the expression of, and the perception of, emotional or affective states. Indeed, Goethe's concern with what has been called in this connection 'the whole mind', i.e. the experientially highly variegated and conceptually nonuniform aspects of mental life within what Wittgenstein called the

stream of life, proves strikingly similar (once one knows where and how to look) to Wittgenstein's investigations into mentally revelatory actions. Goethe, like Wittgenstein, was wary of what he called 'ossified doctrines' that, once lodged into our conceptual substructure, exert a powerful but undetected influence on our subsequent thinking and, owing to their seeming naturalness, resist direct investigative scrutiny. An examination of Goethe's writings on various works of sculpture, painting, and drawing show that distinctive mode of person-perception in contexts within which what Goethe called 'the unity of mind and body' is evident, even if difficult to describe succinctly (or without the examples to do the work of showing what is difficult to say compactly). Goethe, like Wittgenstein, sees the expressive self not *through* or *behind* the body, but rather *in* the contextualized action of the person. Goethe's (and our) perception of Leonardo's qualities of mind in his work, and his similar thoroughly human perception of a thought-induced tremble in a drawn figure of Rembrandt's, shows that these perceptual phenomena are indeed instructively resistant to any simplified formulaic statement of the relation between the mental and the physical. Indeed the relations between these language-games are not accurately, or with a respectable fidelity to the nuances of experience, describable with the reductive concision traditional competing models or pictures of the self would, again seemingly naturally, suggest. Goethe—if with his own distinctive conceptual equipment—thinks deeply about the mind, and throughout his writings he shows a good deal about the self, both directly but also, like Wittgenstein, indirectly. In short, he offers material, of considerable philosophical significance, that shows how to comprehend, without lapsing into polemic-generated theory, the phrase 'an attitude towards a soul'.

In a manner particularly fitting for a philosophical novelist, Iris Murdoch, writing in her diary, often gave voice to philosophical questions concerning the nature of that very writing. In the third section of Chapter 6 I turn to a number of those entries, particularly those concerning what she called 'the unfrozen past'. In them she puzzles over the nature of our relations to our own pasts, and she claims, strikingly, that so long as one lives, one's relationship with one's past should keep shifting. Strengthening the moral dimension of this claim, she adds 're-thinking one's past is a constant responsibility'. Here I suggest that one way of articulating this self-investigative process can be found through a reconsideration of Wittgenstein's remarks on aspect-perception, or 'seeing-as', in this context. These remarks can

prove particularly helpful here, because one often encounters three fundamental positions put forward concerning our relations to our pasts: (1) the view that we project onto past events new content, and thus see in them what we, perhaps only unwittingly, put there; (2) the view that the past is, contra Murdoch, 'frozen', and that it simply was what it was, period (and thus autobiographical verisimilitude reverts to simple correspondence between prior event and later description); or (3) the view that we construct, in narrative, an ever-evolving view of the past as we go along in the stream of life's self-descriptions. But Wittgenstein's remarks on aspect-seeing, seen, indeed, in a new light within this context, show that we need not opt for straight projectivism, perceptualism, or constructivism. Continuing the theme from the previous two sections of this chapter, we come to see that our relation to the past is neither systematically bounded nor unbounded in any of these specific and uniform ways. Indeed, the process of coming to see some part of our past—our past actions, words, thoughts, reactions, hopes, fears, aspirations, or anything else in life that calls for retrospective reconsideration—in a new light or new aspects (or seeing newly emergent patterns of these) gives greater precision and clarity to what Murdoch was alluding to as 're-thinking'. And, as we shall see, closer attention to the vocabulary of sight, the language we use to capture many fine distinctions between categories and kinds of seeing, can prove helpful in coming to understand more fully why we tend to think of, or picture (employing ocular metaphors), the self's relations to its past as we do.

Wittgenstein also wrote a set of remarks—still, I believe, insufficiently examined in the light of autobiographical issues—on memory, and in the first section of Chapter 7 I turn to an examination of a picture of memory, and of remembering, derived from empiricism but influential to the present day; this picture engenders the idea of *objects* of consciousness, where the concept of remembering is elucidated as an inward-directed act of perceptual scrutiny. But this image of memory (and thus of the alleged unitary mental act of remembering) cannot accommodate the relational embeddedness of memories, nor can it (as we see in reference to a closely related discussion of Davidson's) account for the distinctions we make based on what we might call a memory's semantic link to the world, i.e. some are true and some are not. This takes us in turn to Wittgenstein's delicate unearthing of a false picture of recognition: what it is to recognize (a close cognate of memory) turns out to be a far more nuanced and interesting matter than anything

the empiricist picture of mental-object-scrutiny might suggest. In the second section of this chapter I try to capture something of the change to our way of seeing the acts of remembering—presumably central to the work of the autobiographer and his or her self-descriptive writing—that is made possible by a conceptual freedom from the empiricist picture and its progeny. The phrase 'looking into the past' turns out not to refer (as we far too easily take it to do) to an inward act of consulting inert and hermetically sealed visual or auditory images stored in a retrieval system by date and time, but rather to a rich set of experiences, of composite kinds, that take on and cast off relational properties, or networks of interconnections to other experiences both similar and different to them.

By the end of this study we will have seen a large number of fairly detailed investigations into a multiplicity of ways in which the very formulations of philosophical questions house misleading implications, and the ways in which Wittgenstein's work displays a heightened sensitivity to this easily neglected fact. In the final section of the last chapter, I will consider, in the light of all the foregoing, the conceptual similarities (and a few important dissimilarities) between Wittgenstein's ways of working through a problem—or giving his readers the raw material to do so—and self-investigative therapeutic work. Here it emerges that the very phrase 'Wittgenstein's method' itself houses misleading implications (concerning the false hope that *a* unitary method might be first articulated in full and then applied), and that similarities to therapeutic work can be helpful, but only within limits. Still, the earned freedom from the impulsions to speak in a metaphysical voice in accordance with the dictates of captivating pictures is, Wittgenstein suggests, the result of allowing a misled way of thinking (i.e. one that rejects particularity and complexity in the name of a theoretically simplifying schematic concision while still attempting descriptive fidelity to psychological life, to lived experience) to run its natural course. A distinctive kind of patience, a philosophical sensibility receptive to ever-finer contextualized nuance, is necessary for the achievement of the kind of clarity, or perspicuity, of which Wittgenstein writes. And this shows something of the distinctively *personal* nature of this kind, or idiom, of philosophical work. The seeing of newly emergent connections of the kind examined in Chapter 6, and the layered and intricate analyses of linguistically induced conceptual disquietudes, take time and care—as does the kind of autobiographical work (whether it manifests itself in a formal autobiography, memoir, diary, or not) considered throughout this study. Some have argued that Wittgenstein's 'method' is essentially (there is a gentle irony that appears

on both sides of the debate here) therapeutic, while others have argued that it is essentially more conventionally argumentative in nature. Once one has seen the ways in which this kind of conceptual work has, in a number of important respects as outlined in the final section, affinities with therapeutic work (and we see that in practice throughout this book), and also that it extends into sectors of the philosophical land-scape where argumentation is employed to considerable effect, it seems somewhat less necessary to enter into this debate. Both of these aspects are in evidence—along with the self-investigative, self-monitoring labor undertaken throughout Wittgenstein's later writings, a kind of work that vigilantly monitors the pressures on our thinking and then turns to diagnose, through therapeutic (or picture-freeing) analysis and argu-mentation, the conceptual sources of these pressures. And the aim of the philosophical undertaking that utilizes therapeutic, argumentative, and self-monitoring modes of analysis (along with multi-voiced self-directed dialogue) is, of course, conceptual clarity, perspicuity, and—in a philosophical sense—self-understanding.

Conceptual perspicuity, as we see at the close of the final section, is the kind of thing one achieves in a case-by-case, piecemeal manner. The fact that we cannot characterize generally a system called 'Wittgenstein's method' or generically define 'perspicuousness' itself (as we will see by returning to Cavell and to Cavell's Emerson at this point), along with the fact that one cannot at the end of investigations of this kind elucidate a succinctly expressed theory of autobiographical or self-descriptive language that stands at the end of a single overarching argumentative line, is not a limitation. It is, rather, a source of conceptual freedom, one that allows an ever more clarified way of seeing the contextually emergent significance of particularities. To put the matter in any more generalized way would falsify one distinctive mode of conceptual engagement that seems to some of us well worth keeping alive within the larger world of the contemporary philosophical and literary-critical scene.

Lastly, I should perhaps state explicitly that the following chapters comprise more of a set, or more of a mosaic, than they do a sequential progression along a single argumentative line (although some chapters and many sections of chapters are so related). Each chapter attempts to cover (or dig into) one part of the larger terrain (or one part of Wittgenstein's 'landscapes', about which I'll have more to say), so the book is more perhaps of a philosophical analogue to an archeological dig than to a sequence of mathematical reasoning. The task is to unearth, as carefully as possible in each ensuing chapter, many very

particular and sometimes quite subtle linguistic forces that influence, to a greater extent than we often realize, our thinking about the self and self-description. Then standing back and taking them all together offers a way of understanding autobiographical writing and self-descriptive language very unlike the deeply entrenched and seemingly more intuitive Cartesian model and its derivatives—a different way of seeing, of describing, ourselves.

It will be clear that a number of classic Wittgensteinian *topoi* are investigated here (e.g. the relations of thought to language and the Cartesian-versus-behaviorist dichotomy in modeling the self), but as they are examined within the less familiar context of the question of autobiographical language, I hope they themselves are seen in new ways. And there are also a number of Wittgensteinian subjects that have been little investigated (among them the *positive* conception of introspection; the notion of hidden understanding in relation to self-knowledge; the nature of retrospective meaning-determination; the remarks on memory; and the too often concealed powers on our thinking exerted by what Wittgenstein called philosophical pictures as seen in relation, generally, to linguistic meaning and, particularly, to the self's reflexively descriptive, expressive, and constitutive language). In both categories of topics, I hope that the context of the investigation allows new light to be cast (and new aspects to dawn) on a philosophical problem surely as old as the first moment of reflective self-awareness.

1

Autobiographical Consciousness

POSITIONED on the edge—or the precipice—of solipsism, it was Schopenhauer who famously asserted that the world is my representation. We know that Schopenhauer's philosophy exerted a strong influence on the early Wittgenstein, whose equally famous—and equally metaphysical—claim in his *Tractatus* that the world is all that is the case resoundingly announced his early entanglement with grand metaphysics (if in linguistic, rather than ontological, form).[1] Schopenhauer's claim makes the world a mental, or individualistically interior, representation that is, indeed, private to the mind of the individual whose representation it is, a representation that constitutes at once the contents and the boundaries of private consciousness. It is thus, to borrow Thomas Nagel's phrase, not only a claim concerning the necessity of entering that individual's consciousness (where this possibility is denied by the solipsist and debated by others) to know *what it is like* to be that individual; it is, for Schopenhauer, a far stronger claim.[2] The world is not a larger, realist place within which that individual consciousness is contingently situated, but rather the very *idea* of the world is unintelligible without *first* positing the existence of an individual consciousness that constructs it as, indeed, its *own* representation.

---

[1] Schopenhauer, *The Word as Will and Representation*, trans. E. F. J. Payne (New York: Dover, 1966), vol. i, §§1 and 10; vol. ii, ch. 1; Wittgenstein, *Tractatus Logico-Philosophicus*, trans. D. F. Pears and B. F. McGuinness (London: Routledge & Kegan Paul, 1961). See, in this connection, the entry 'Consciousness' in Hans-Johann Glock, *A Wittgenstein Dictionary* (Oxford: Blackwell, 1996), which I have found very helpful. The Schopenhauerian influence does not wane quickly: Glock notes that as late as Wittgenstein's lectures in 1930–3 (as recorded by G. E. Moore), he said, 'All that is real is the experience of the present moment' (p. 85).

[2] Thomas Nagel, 'What Is It Like To Be a Bat?', *Philosophical Review*, 83 (1974), 435–50.

1. SCHOPENHAUER, WITTGENSTEIN'S TRANSITION, AND THE EDGE OF SOLIPSISM

This powerful claim was not far from the articulated view of the young and early Wittgenstein in his *Notebooks, 1914–1916* and the *Tractatus*—with which he was, in his far deeper reflections in his rapidly maturing remarks in the *Blue Book* and then his fully mature remarks of *Philosophical Investigations*, to wage a kind of reverse-Oedipal battle with his former self (the first skirmishes with that 25- to 27-year-old self breaking out in his lectures of 1932). In the *Notebooks* entry of 11 June 1916 Wittgenstein, having already said that he knows that this world exists—and one might reasonably load a good deal of metaphysical freight into his word 'this', i.e. he means it in the proprietary sense of Schopenhauer's 'my'—observes that he is placed in it just as an eye is placed in its visual field. Rather like an analytical commentary on Schopenhauer's grand claim, this striking way of putting how it is we are positioned in this world is a precise analogy to the visual field understood as the content and the boundary of our experience. Giving the claim a chiseled precision, he adds the potent sentence 'That life is the world', i.e. the world, its substance and its extension, is given within the mental, or interior, experiential analogue to the visual field of the eye, that is the private, individual consciousness. Returning to the theme on 24 July 1916, he writes simply, 'The World and Life are one' (which became *Tractatus* 5.621), but then adds three sentences: the first, 'Physiological life is of course not "Life"', defies any physicalistic reductionism and emphasizes by the exclusion of the physiological the mental, the psychological interior. But then the second sentence—as though anticipating his much later undercutting of the entire behaviorism-versus-Cartesianism dichotomy—quickly rejects the polar opposite: 'And neither is psychological life.' Having evaded these twin reductive exclusions, he states the more encompassing, and more Schopenhauerian, claim: 'Life is the world.' And the Schopenhauerian character of this pronouncement is further brought out in his entry of 7 August: 'The I is not an object,' that is, to put it one way that seems consistent with at least the spirit if not the letter of his remarks in those years, the referent of the first-person pronoun is not one among many other particulars in the world that exist autonomously from their names, in this case the 'I', but rather is itself the necessary condition of

that world. That referent, like the referent of Schopenhauer's 'my' in 'my world', is, like the visual field of the eye, not *itself* encountered, not *itself* seen.[3]

Thus Wittgenstein adds four days later: 'I objectively confront every object. But not the I.' And then closing, in a drop of grammar, whatever small gap there might remain between the early Wittgensteinian and the Schopenhauerian senses of these claims, on 12 August 1916 writes: 'The I makes its appearance in philosophy through the world's being *my* world.'

This conception of the self as an interior consciousness whose boundary we do not perceive and whose nonencountered existence is the precondition for the world—for that consciousness, a world that is *mine* without remainder—takes a central place in the *Tractatus*. In 5.633 Wittgenstein encapsulates, and advances, the Schopenhauerian points above: 'Where *in* the world is a metaphysical subject to be found?', his italicized '*in*' now calling attention to the notion that the I, the self, is *not* an object like others in the world that we come across, identify, describe, confirm-as-existing, and so forth. He continues (anticipating his debate with an interlocutor who consistently voices philosophical positions showing the grip of philosophical pictures): 'You will say that this is exactly like the case of the eye and the visual field.' Saying just that, however, implicitly leaves open the possibility of encountering the eye itself as an object in, or—differently—the limit of, that visual field, and so Wittgenstein adds: 'But really you do *not* see the eye.' And then, advancing the argument in favor of this conception, makes the point that the role the eye plays in the visual experience is, as it were, *offstage*: there is nothing observable '*in the visual field* [that] allows you to infer that it is seen by an eye', the consciousness of selfhood, on this model or picture, functioning analogously as (to cast the matter in terms reminiscent of Kant) nonperceived precondition for experience. And he had prepared the way for these observations with 5.632: 'The subject does not belong to the world: rather, it is a limit of the world.' That was in turn prepared for by his famous remark in 5.62 concerning what he there identified as the element of truth in solipsism: 'The world is *my* world: this is manifest in the fact that the limits of *language*', to which he then adds a densely compressed articulation of the picture of exclusively inward, or private, language with which he will also do

---

[3] Wittgenstein, *Notebooks, 1914–1916*, ed. and trans. G. E. M. Anscombe (Oxford: Blackwell, 1961), entry for 4 Aug. 1916 (p. 80).

battle in *Philosophical Investigations* and change the course of modern philosophy as a result—'(of that language which alone I understand) mean the limits of *my* world.' Thus the Schopenhauerian metaphysics is transmuted from an ontological to a linguistic thesis, and it reaches its culmination in 5.641: 'The philosophical self is . . . the metaphysical subject, the limit of the world—not a part of it.' And, as a limit, the self is, indeed, the *only* limit: in 6.431 we find: 'So too at death the world does not alter, but comes to an end,' thus reaffirming the Schopenhauerian–Wittgensteinian claim that any intelligible talk of an existent world will be, *ipso facto*, of a mind-dependent one. And thus, in 6.4311 we get: 'Death is not an event in life: we do not live to experience death,' and a sentence later, linking this to the deep analogy of the philosophical self and the eye and its visual world: 'Our life has no end in just the way in which our visual field has no limits.'

All of this, to this point, casts light on the metaphysical picture that rests beneath these self-defining utterances; this conceptual model, or picture, is expressed with the very greatest linguistic density in 5.63: 'I am my world.' The philosophical intuitions concerning the nature of the self that are formed and fueled by the conceptual picture Wittgenstein has adumbrated in his early writings in fact account for a good deal of our attraction to autobiographical writing. To the extent that we all-too-naturally think of the self and its place in its world in a fashion consistent with Wittgenstein's early position, we then all-too-easily construe autobiographical writing as a special *kind* of writing: a kind that promises not only a glimpse of the world as seen through other eyes (which would be interesting or magnetic enough), but rather a glimpse—or indeed a sustained, long look—into *another* world, a world that is, in the foregoing metaphysical and consciousness-dependent sense, '*my*' world, i.e. the world of the autobiographer's. And as such, we thus think of autobiographical writing as a kind of literary antidote to the true element of solipsism to which Wittgenstein referred within the larger context of his Tractarian metaphysics, and we—if only in a sense that could never attain true or complete entry into the mind of the other but still holds out the promise of other-mind understanding—expect a view not merely of what it is like for another to live in our world, but rather the far more personally and philosophically compelling view *into another's world*. But every component of this picture, this way of intuitively modeling the conscious self in its autonomous world and then subsequently longing to cross the skeptical divide into the mind-world of the other, Wittgenstein battled

against in his mature philosophy. Bewitching forms of language, lodging conceptual confusion deep in the intuitive substrate, inculcate in us not metaphysical truths of self-consciousness, but rather misleading pictures that shape all of our subsequent thinking on the subject. And it would be bewitching linguistic forms that Wittgenstein came to identify as the enemy—the enemy of conceptual clarity—during, and increasingly strongly thereafter, his 1932 lectures. The following of this struggle against these earlier, and seemingly natural, ways of thinking—ways of seeing consciousness and selfhood—should prove of immediate relevance to an increasingly full and increasingly clear account of the position of the self investigating itself, i.e. of autobiographical consciousness.

Consider the striking difference in method, tone, and what one might call *way of seeing* the issue, in the *Blue Book* (dictated to his Cambridge pupils in 1933–4):

The difficulty which we express by saying 'I can't know what he sees when he (truthfully) says that he sees a blue patch' arises from the idea that 'knowing what he sees' means: 'seeing that which he also sees'; not, however, in the sense in which we do so when we both have the same object before our eyes: but in the sense in which the object seen would be an object, say, in his head, or in *him*. The idea is that the same object may be before his eyes and mine, but that I can't stick my head into his (or my mind into his, which comes to the same) so that the *real* and *immediate* object of his vision becomes the real and immediate object of my vision too. By 'I don't know what he sees' we really mean 'I don't know what he looks at,' where 'what he looks at' is hidden and he can't show it to me; it is *before his mind's eye*. Therefore, in order to get rid of this puzzle, examine the grammatical difference between the statements 'I don't know what he sees' and 'I don't know what he looks at,' as they are actually used in our language.[4]

Wittgenstein, now in a different voice from that of his former Tractarian self, is not stating a metaphysical limit of experience and showing something of the solipsism that cannot within the limits of our world, the limits of our language, be said ('what the solipsist *means* is right'[5]). He is, rather, asking if the very formulation of the problem (now demoted to a 'puzzle') can make *sense*, and the tribunal that judges that question will not be the Schopenhauerian metaphysics of selfhood,

[4] Wittgenstein, *The Blue and Brown Books* (Oxford: Blackwell, 1958), 61.
[5] See Glock, 'Solipsism', in *A Wittgenstein Dictionary*, for a succinct discussion and a helpful set of references throughout Wittgenstein's published and unpublished writings on this topic; the Schopenhauerian influences are helpfully brought out by Glock.

but ordinary linguistic usage. Looking to see how such phrases *are* used in our language will break the hold of the picture, which it does by calling into question the very sense of the various articulations of that picture. Now Wittgenstein is placing that picture-driven and conceptually bewitching language up against the standards of our *usage*, and he will go on to conclude that there indeed can be, and are, contexts of human discourse within which we intelligibly speak of not knowing what someone sees, but these will prove soberingly unlike the problem of other-minds skepticism and mind-enclosed solipsism seen here: they do not reduce to the problem of not being able to get access to the inner content—the putative *immediate* content—of his experience. And similarly, he observes in the subsequent discussion that, in regard to the concept 'person', we are at liberty to choose from multifarious, context-sensitive usages that, as he tellingly suggests, amount to choosing 'between many different kinds of analogy'.[6] Analogies for personhood, for the self, for consciousness, exert great power on our thinking, and to think of consciousness as a locked chamber, to think of the contents of that chamber as perception that will thus seem ineluctably private, to think of our experiencing of the world as hidden, but hidden inwardly, all conspire in favor of the Schopenhauerian–early-Wittgensteinian conception of the self and its world, and that way of picturing the positioning of consciousness in turn fuels the kind of fascination with autobiographical revelation mentioned above. The breaking of that spell is accomplished in language, and thus Wittgenstein, in the final pages of the *Blue Book*, turns to the instructive particularities of linguistic usage of the first-person pronoun that loosen the grip of those analogies, those pictures. He is showing, if not quite yet explicitly saying, that understanding the variegated grammar of the 'I' is necessary if we are to understand the nature of the consciousness that defines selfhood, in a way unlike that which his former self envisioned.

Both calling a troublemaking group of phrases to the court of usage, and picking up a thematic thread from his earlier philosophy—though now addressing it in a transformative manner—he writes:

What tempted me to say 'it is always I who see when anything is seen,' I could also have yielded to by saying: 'whenever anything is seen, it is *this* which is

---

[6] *The Blue Book*, 62. At this point Wittgenstein writes into his discussion an implicit justification of his own method. He observes that the word 'personality' has no one legitimate heir any more than does the word 'philosophy'.

seen,' accompanying the word 'this' by a gesture embracing my visual field (but not meaning by 'this' the particular objects which I happen to see at the moment). One might say, 'I am pointing at the visual field as such, not at anything in it.' And this only serves to bring out the senselessness of the former expression. (p. 64)

Wittgenstein then articulates directly and forcefully the way in which we are all-too-easily misled into thinking that the foregoing metaphysical utterances possess meaning just as do nonmetaphysical expressions, 'for we wrongly compare our case with one in which the other person can't understand what we say because he lacks a certain information' (p. 65), and then adds, as if writing an abbreviated recipe for the sustained labors of *Philosophical Investigations* to follow, '(This remark can only become clear if we understand the connection between grammar and sense and nonsense.)' The grammars of the self, of the 'I', of consciousness, do not behave, on inspection, at all like the way we expected under the influence of misleading analogies, bewitching language, and conceptual pictures, and indeed we will be enabled to see those grammars as exhibited in usage clearly, nonprismatically, *only* if we therapeutically free ourselves of their domination. (That therapeutic project we will examine below, in his remarks on consciousness in *Philosophical Investigations*.) And he is writing in self-defense, against an anticipated interlocutor who will insist that, wholly independently of any tribunal of usage, he knows he means something intelligible and profound by his Schopenhauerian utterances on self and world. Against this expected reply—and very plausibly a reply made by his own former, Tractarian self, in the early 1930s—he writes: 'The meaning of a phrase is not a mental accompaniment to the expression. Therefore the phrase "I think I mean something by it" or "I'm sure I mean something by it," which we so often hear in philosophical discussions to justify the use of an expression, is for us no justification at all. We ask: "*What* do you mean?," i.e., "How do you use this expression?" ' (p. 65). And that test, using the measure of intelligible usage, is one that the metaphysical utterances concerning the 'I' cannot pass, which the final part of the *Blue Book* sets out to demonstrate on the level of grammatical detail. The work he undertakes in those pages of the *Blue Book* is far too intricate to recount fully here, but a few passages may stand for the whole.

Exposing the influence of misleading analogies, Wittgenstein observes that when we use the word 'I' as a subject in a sentence, we can far too easily believe the illusion, created by the empirical fact that we do not use it because we recognize a given person by his bodily characteristics,

that we really use the word to refer to a bodiless something, an inner, metaphysically hidden ego that has its seat in our body, but is of a different ontological kind from it (p. 59). And *that* conception of selfhood is, of course, foundational to the Schopenhauerian metaphysic and the correlative conception of autobiographical revelation; it is fundamental to the entire conception of the reading of an autobiography as a philosophical event, i.e. looking—to the extent that we can do so using language—into the mind of another with the pre-later-Wittgensteinian metaphysics. And of the sense of metaphysical privacy endemic to that way of thinking, he writes: 'When I say "Only this is seen," I forget that a sentence may come ever so natural to us without having any use in our calculus of language' (pp. 65–6).

But the conception of the self—and thus the subject of autobiography—as an inner point of consciousness whose seat is, contingently, the body, is a picture that dies hard. It is, as we now know, preserved, even against our better judgment,[7] by the illusions of grammar. He writes: 'Now the idea that the real I lives in my body is connected with the peculiar grammar of the word "I," and the misunderstandings this grammar is liable to give rise to' (p. 66). At this stage of his development Wittgenstein draws a contrast (one to be made with much greater subtlety in *Philosophical Investigations* and his mature writings on the philosophy of psychology) between categories of cases where 'I' is used as an object and those where it is used as a subject. If I refer to my broken arm, the bump on my forehead, or the fact that I have grown six inches (p. 66), I am using the first-person pronoun in the 'I-as-object' sense; conversely, if I say that I think it will rain, or I see an elephant, or I hear a distant flute, or I have a toothache, I am using the first-person pronoun in the 'I-as-subject' sense. The distinguishing mark of the former category of usage is that such usages involve the recognition of a person, where, Wittgenstein pointedly adds, there is (however remote) a possibility of error. For example, in an automobile accident I may feel a pain in my arm, in disoriented confusion see my neighbor's broken arm, and mistakenly (probably very briefly) think it mine, or I could look into the rear-view mirror, see a bumped head, and take it (momentarily) as mine. There is, importantly, no such possibility of even the most fleeting error in the 'subjective' cases of toothache, hearing a flute, seeing an elephant.

---

[7] I discuss the persistence of this picture of selfhood and its continual revivification by particular ways of speaking in Ch. 6, Sect. 3, below.

That the pseudosentences formed by the words 'But are you sure it is you who sees the elephant?' and 'But are you sure it is you who has the toothache?' are nonsensical linguistic curiosities is self-evident—and that they convey a hint of metaphysical depth is telling. Such linguistic curiosities transgress the boundaries of sense; they are, Wittgenstein observes, not only bad moves, but 'no move of the game at all' (p. 67).[8] It is, indeed, impossible to 'moan with pain by mistake, having mistaken someone else for me' (p. 67). Thus, as he encapsulates the point here, to make a statement about one's pain using the first-person pronoun is no more a statement *about* that person than is the moaning of that person.[9] The conflation of these twin categories of cases—indeed very close to what Ryle famously termed a 'category mistake'[10]—yields strong and deeply misleading support for the conception of metaphysical privacy that is at the heart of the Schopenhauerian–early-Wittgensteinian way of thinking that so quickly generates the correlated picture of autobiographical consciousness. If we can be wrong about the arm or the bump, perhaps we can be wrong about the pain, thus driving a wedge between ourselves and our embodied experience. And this misplaced self-directed skepticism makes that embodiment seem contingent and, in a metaphysical sense, superfluous to who we truly are—which in turn makes a Cartesian interior seem to be precisely what we want autobiographical writing to report on,[11] i.e. the 'walled garden'[12] to

[8] I offer a fuller discussion of the language-game (and its significance for aesthetic understanding) in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), ch. 1.

[9] I discuss some of these nuances of self-description in Wittgenstein's philosophy, particularly in connection with the behaviorist picture of the self and the reasons Wittgenstein's position cannot be reduced to, or encapsulated as, behaviorism in the first sections of Ch. 3 below.

[10] Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949); see esp. 'The Origin of the Category-Mistake', 18–23.

[11] In Ch. 5 below I try to identify a number of the conceptual pressures pushing us into the very idea of *reporting* on the contents of the Cartesian interior.

[12] For insightful (and instructively discordant) discussions of the picture of privacy and its attendant problems of self-knowledge in recent Wittgenstein-inspired philosophy, see Crispin Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's 'Philosophical Investigations'* (Cambridge, Mass.: Harvard University Press, 2001), sect. 3: 'Privacy and Self-Knowledge', 215–374; and John McDowell, *Mind, Value, and Reality* (Cambridge, Mass.: Harvard University Press, 1998), sect. 3: 'Issues in Wittgenstein', 221–321. See esp. McDowell's 'Intentionality and Interiority in Wittgenstein', 297–321, where his intricate diagnosis of the grip of a conceptual picture (what he here usefully calls 'the framework') indeed shows how deeply such pictures are lodged in our language and how persistent they continue to be not only in contemporary Anglophone philosophy of language and mind but also in work that is allegedly Wittgensteinian in nature.

which we want, as readers, entry. Or, to take the grammatical conflation the other way (and thus cause different metaphysical trouble), we might all too easily think that because there exists no logical room for error in the 'subjective' cases, there similarly exists none in the objective, and we respond to any imagined counterexamples (of the automobile-accident kind above) by saying that we never make errors about what we *seem* to see or experience, thus driving a wedge between ourselves and the world we inhabit and with which we interact.[13] And this in turn, through an appeal to a variant of the sense-data picture of human experience, makes our experience seem incorrigible within the Cartesian interior, and thus when we turn to autobiography it will be—and here the two kinds of conflations, the subjective category of I-usages assimilated to the objective and vice versa, converge—the truthful presentation of that interior content that we will want, and we will correspondingly think of autobiographical truth as that interior content accurately represented externally. And if we, or some parts of us, insist that regardless of the disentangling of grammatical conflations that might preserve categorical clarity with regard to usages of the first-person pronoun, the word 'I' still must mean one determinate thing, Wittgenstein reminds us that the first-person pronoun is a tool, an instrument in our language, with a variety of context-sensitive employments: 'The word "I" does not mean the same as "L.W." even if I am L.W., nor does it mean the same as the expression "the person who is now speaking" ' (p. 67). The mistake to make at precisely this juncture would be to thus believe that 'L.W.' and 'I' in this case mean *different* things—that would be to cling to a fixed-referent conception of meaning-determination, where each word functions as, or like, a name. So he adds: 'But that doesn't mean: that "L.W." and "I" mean different things. All it means is that these words are different instruments in our language' (p. 67).

Such observations do not, of course, *replace* one philosophical picture with another: they, by contrast, loosen the grip of the picture and the way of seeing the problem that is enforced by that picture. This deeper philosophical process brings a kind of light, or a new way of seeing, not accessible when, more superficially and conventionally, the picture in question is merely supplanted by another. And such transformations of

[13] See Donald Davidson's remarks in 'Knowing One's Own Mind', in Quassim Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994); see esp. pp. 60–4 on the difficulties of overcoming a pernicious picture of the mind enforcing the belief that thoughts require (external-world-mediating) mental objects; we will return to these matters in Ch. 7, Sects 1 and 2, below.

our ways of seeing problems, invariably within particularized contexts of inquiry, just are the process and progress of the later philosophy in its full maturation. Emphasizing that the meaning of a term or an expression depends entirely on our use of it, Wittgenstein, in the closing passages of the *Blue Book*, writes—against a way of seeing codified in his early writings—'Let's not imagine the meaning as an occult connection the mind makes between a word and a thing, and that this connection *contains* the whole usage of a word as the seed might be said to contain the tree' (pp. 73–4). Transmuted to the understanding of autobiography, the meaning determined by the occult connection would constitute something of a holy grail of other-understanding: not only knowledge of the other's Cartesian interior, but, in a grander metaphysical Schopenhauerian manner, the world as represented in that other's inner-meaning-laden consciousness. But that conception, that picture, of the interior self and its meaning-determining occult processes, if only a myth born of misleading analogies and grammatical conflations, must be intricately removed in order to grasp rightly the nature of the self and its contents as revealed in our autobiographical practices. With that thought, we turn—having seen both Wittgenstein's starting point on these issues and his middle-period reaction against his own former way of thinking—to the intricate removal of misconceptions of consciousness of his late, mature work. Only this kind of removal will allow a clarified view of the richly human autobiographical endeavors in which we engage.

2. THE 'INNER PICTURE'

In *Philosophical Investigations* §416 Wittgenstein's imagined interlocutor suggests that, because we humans agree in saying that we see, hear, feel, and so forth, we must thus be our own witnesses that we have *consciousness*. Mindful of the snares of language, of misleading analogies, and of the necessity of context for making intelligible, sense-bearing moves within a language-game, Wittgenstein replies: 'But how strange this is! Whom do I really inform if I say "I have consciousness"?' And then turning the focus from communicating the fact of our consciousness to others to self-reporting, asks: 'What is the purpose of saying this to myself . . . ?' And he reminds us that there *are* cases, contexts, in which we might tell someone who witnesses our fainting spell, 'I am conscious again', but the implicit warning is that such cases are, on the level of grammatical appearance, similar to the interlocutor's philosophical

utterances, but in truth nothing like them at all. Pursuing this contrast between grammatically disguised nonsense and the intelligible, he begins §417 with the question 'Do I observe myself, then, and perceive that I am seeing or conscious?' He dismantles the apparent but deeply misleading sense of obviousness about this in layers. First, we find the concept of *observation* called into question. Wittgenstein does not fully undertake such an investigation here (he did describe his book as 'a machine to think with' and does characteristically leave a good deal of work to the reader), but if we were to consider extensively cases in which we readily and unproblematically speak of *observing* a thing or situation, we would see that the case of self-consciousness is *nothing* like that—indeed to the point where we would feel disoriented in trying to apply the term 'observation' to the situation at all. And the close reading of autobiographical self-investigation shows precisely this, as a number of cases considered below will suggest.

But then Wittgenstein, at the next layer, asks: 'Why not simply say "I perceive I am conscious"?' This would, however, constitute a kind of thinly disguised conceptual recidivism, since the concept of *perception*, on a similar investigation, would seem equally remote, equally detached from the real language-games of self-investigation. Thus, he asks, again in a kind of analytical shorthand: 'But what are the words "I perceive" for here?' This phrase, as an instrument in our language, would show *not* the true underlying nature of the ordinary case of consciousness, but rather, if applicable at all, only that our attention is disposed in a particular way. One might say (Wittgenstein does not, but a thorough investigation of the ground he is rapidly moving over would suggest it) that the phrase 'I perceive I am conscious' is a kind of pleonasm of self-reportage for the cleaner 'I am conscious'. But these statements do not capture the ordinary case of a person's consciousness either, for in what cases do we *report*[14] on our states, and in what contexts do we *say* 'I am conscious'? In the fainting case in which the *question* of consciousness has come to the fore, perhaps; in the stream of conscious life, never. This is a sentence grammatically similar to 'I am hungry', 'I am cold', 'I am delighted', and, by slight extension, 'We are not amused'. And the sense of these remarks lends illegitimate support to their grammatically similar self-report of consciousness, one that can deliver only a false promise of intelligibility.

[14] I offer a discussion of Wittgenstein's particularly relevant passages in Ch. 2, Sect. 1, and Ch. 3, Sect. 2, below.

But how then should we characterize the evident fact that most of us go through life conscious most of the time (and *that*—the particularities and nuances of that experience that is for each of us distinctive to us—is again how we think of the subject matter of autobiographical writing)? In §418 Wittgenstein asks: 'Is my having consciousness a fact of experience?' And he assembles some stray thoughts that seem to endorse, indeed to necessitate, such a general claim: don't we say that persons have consciousness while trees and stones do not? And (in §419) if we speak of a tribe with a chief, must not that chief have *consciousness*? 'Surely', he continues the voicing of this grammatically misled line of thinking, 'we can't have a chief without consciousness!' There are contexts in which we speak of consciousness as an attribute or as a state of a person, including ourselves, but there is a deep logical difference (disguised by grammatical similarity) between saying that green or hot are properties some things possess and others don't and that consciousness is a property some things have and others don't, and so it is a property *of that kind*. Some chairs are green, others are not, so we can ask how many chairs in a room are green, how many not. It would transgress conceptual–logical limits to inquire in turn, however, into the nature of greenness itself: the disorienting generality of this question (a question of a metaphysical kind logically *very* different from either a physicist's question about the wave-behavior of green color, or the neurophysiologist's question concerning the ocular system that allows us to differentiate green from blue) is a symptom of our having lost our link to legitimate moves in the language-game. Similarly, various kinds of things come into, and pass out of, existence in contexts of all kinds; such particularized contexts are, again, *very* remote from the generalized question concerning, not when the self-reflective awareness of a kind evinced in Petrarch's sonnets or the hyperromanticized, nervously overwrought self-image epitomized by Goethe's Werther came into existence, but rather—although grammatically similar but logically different in the extreme—what the nature of existence, or indeed of Being, *itself* [15] might be. If some chairs are green, and some exist, is then existence not a property? If some entities are conscious, and others not, is not consciousness a property? And here again, the word 'as', versus 'is', condenses a cloud of philosophy into a drop of grammatical difference.

---

[15] For an exceptionally lucid, compact review of what one might call the etiology of this kind of intellectual quandary, see P. M. S. Hacker, *Wittgenstein: On Human Nature* (New York: Routledge, 1999), 5–14.

   Wittgenstein asks if we cannot imagine that the people around us
are automata, that they lack consciousness, and that, for example,
the liveliness of a group of children is the result only of automatism
(§420). He predicts that we will, in one reaction to this suggestion,
find the words, the sentences, expressing this 'automata view' of persons
becoming quite meaningless when they are, as it were, held up against the
reality of the persons. Or as a second reaction to the suggestion, we might
momentarily provoke an uncanny feeling in ourselves. He then speaks of
this in terms of 'seeing a living human being as an automaton . . .', and
therein lies the grammatical cue (which relates directly to his important
remarks on aspect-perception and the phenomena of 'seeing-as' in part II,
§xi[16]) that calls for more elucidation than Wittgenstein gives it in this
passage of his 'machine'. That we can try to see persons as automata
(and if slightly successful, experience the sense of the uncanny) suggests,
erroneously, that if we aren't seeing them as automata, then we must,
in the ordinary case, be seeing them as something else—which in this
context would obviously be characterized as beings-with-consciousness.
This is a grammatically parallel construction that leads us astray precisely
because, in the ordinary case, we don't see persons *as* anything; we
see that it *is* a person, or group of children, etc., before us. And
those we see, not on an additive model (which the misapplication
of the seeing-as construction would here strongly suggest) where the
perception of the person is analyzable into its constituent parts with the
separable property of consciousness being added to an isolated body or
a humanoid mechanism, but in terms of the irreducible, unanalyzable
'attitude towards a soul', an *Einstellung zur Seele*, of which Wittgenstein
writes.[17] That this distinctive attitude is *basic* to human beings, i.e.
fundamental to *any* understanding of who, what, and how we are, is
what gives the strength to the ordinary nature of our person-perception
sufficient to make the first reaction Wittgenstein articulates common,
where the words and sentences expressing the automata-view quickly
seem meaningless. This, to express it in a drop of grammatical difference,
is the triumph of 'is' over 'as'. The picturing of the consciousness of

---

   [16] I take this matter up directly in Ch. 7, Sect. 3, below.
   [17] For a more detailed discussion of this irreducible attitude, see my *Meaning and
Interpretation*, 'Against Reductionism', 129–38. I think one can see a good deal of directly
relevant philosophical detail (on person-perception in relation to the irreducible attitude)
in Mozart and Da Ponte's *Don Giovanni*: I argue the case in 'Leporello's Question:
Don Giovanni as a Tragedy of the Unexamined Life', *Philosophy and Literature*, 29/1
(Apr. 2005), 180–99.

the person as ontologically distinct from, and only contingently *related* to, what we will then call *embodiment* is fueled, in part, by the notion that we see persons as entities-with-the-property-of-consciousness, and it is not then long until we are, in keeping with this dualistic picture of selfhood, asking how we can—or, solipsistically, *if* we can—gain access to that inner realm, and we then look to autobiography in that light, with those philosophical motivations and their correlated expectations. In short, this is yet another juncture of thought at which a misconstrual of the self, as a kind of grammatical optical illusion, makes us inquire into the nature of consciousness *itself*, i.e. in isolation from the human practices, engagements, and interactions that assure the intelligibility of the concept of consciousness in the first place, and this in turn causes us to miscast deeply the nature of autobiography and, particularly, autobiographical truth.

In *Philosophical Investigations* §426 Wittgenstein contrasts the chiseled precision and seeming clarity of the pictures, the conceptual models, we employ in philosophical thinking about sense and meaning with the actual diverse uses we make of the concepts under scrutiny, saying that the latter invariably seem, by contrast with the former, 'something muddied'. He likens this to our conceptual tendencies in thinking about set theory, where 'the form of expression we use seems to have been designed for a god', i.e. the way we tend to speak of such issues *in the abstract* sounds like what he now, in his mature philosophy, understands to be the falsifying neatness—indeed, a false rigor—of the way of thinking given articulation in the *Tractatus* (where the true rigor is now to acknowledge, comprehend, and painstakingly earn an overview of the particularities of the concept in question). That imagined godlike view in set theory, or in the philosophy of language in determining sense and meaning, is perfectly parallel to the case of picture-driven autobiographical thinking: that godlike perspective—precisely the position we desire to occupy in understanding the mind of another and which, on the model in question, we hope autobiography might provide or approximate—is possessed, he writes, only by one 'who knows what we cannot know'. And then adding a phrase that renders the connection we are now considering between a conception of sense and meaning and a conception of autobiographical content explicit, Wittgenstein writes: 'he sees the whole of each of those infinite series and he sees into human consciousness'.

But Wittgenstein as quickly places his mature thought against that reiteration of his old way of thinking, saying that those icy pictures

where everything would be fixed '*unambiguously*', high and remote from any genuine use in human contexts, are like pontifical vestments we might don but nevertheless can do nothing with, since we lack what would give such vestments meaning and purpose. In the actual use of expressions, he adds: 'we make detours, we go by side-roads'. Autobiographical writing is similarly a process of taking detours and following up side roads; the autobiographer is, rightly if more messily, understood not in the position of the god with regard to himself; nor do we, as readers, ascend to that Olympian position.[18] But those pictures, again, die hard; we might well say to ourselves in such contexts of person-interpretation (or any biographical project, broadly construed) that 'while I was speaking to him I did not know what was going on in his head' (§427). But Wittgenstein adds, diffusing the misleading power of the picture and thus showing the great importance a seemingly small grammatical restructuring can assume, that, in saying this, 'we only mean what elsewhere we should mean by saying: we should like to know what he is thinking'. And thinking is an activity that we can *make* seem occult and ineluctably private in accordance with the dualistic picture of the self that underlies the presently disputed way of construing autobiography—but it *need* not seem so.[19]

  Thus §428 begins with the interlocutor's sentence 'This queer thing, thought', and that misled remark only heightens the sense of mystery of first-person content and the sense that such thought is the private maker of the hidden Schopenhauerian world of the other. But the mature Wittgenstein of *Philosophical Investigations* counterposes: '—but it does not strike us as queer when we are thinking. Thought does not strike us as mysterious while we are thinking, but only when we say, as it were retrospectively: "How was that possible?"' In accordance with the picture, the genuine understanding of another person's thinking, another person's thought, would be a metaphysical impossibility; in accordance with a far less neat reality, such understandings, some-times characterized as genuine in contexts where that word marks an

---

[18] The reality of those autobiographical processes is far closer, in fact, to what pragmatism might suggest. Very briefly, we would find an individual or group of individuals working within a problem-field of self-or-other understanding—or working within particularized contexts within which the criteria relevant to the determination of the particular issue will emerge as relevant in highly context-sensitive ways that no generalized *theory* of autobiographical writing or person-interpretation could possibly accommodate.

[19] I attempt to show this much more fully in Ch. 4, Sects 2 and 3, below.

important contrast, will come in a thousand different forms (it is perhaps literature that best provides the vast catalogue of cases of other- and self-understanding of precisely the kind Wittgenstein repeatedly suggested we assemble)—reminders of what we *actually*, contra the picture, say and do—in order to change our way of seeing, to loosen the grip of the falsely unifying picture. Returning to a theme from the *Blue Book*, where he reconsidered the question of a word as a sign, the 'life' of that sign, and the dichotomy this terminology insinuates, in §432 Wittgenstein, having said that every sign *by itself* seems dead (having said what he did in the *Blue Book*, of course he is wary of this sign-versus-life distinction, although he does not record this wariness here), he states that the life comes from its *use*. No godlike speaker, as sole owner of inward meanings, breathes life into them. And if that is true, then we, as readers of autobiography, ought to be given our freedom from the false prison[20] of two deep misconstruals, in which we attempt to break out of ourselves and gain entry into the meaning-determining, or linguistically 'life-breathing', mind of another. The position—of us, and of the autobiographer—is *wholly* different. It is the *use* in our language of self-descriptive, or self-investigative, terms that gives them their 'life'; free of a misleading picture of autobiographical consciousness, we can begin to see the varied examples of autobiographical writing for what they are: instructive reminders, against the deceptions of generalizing theory, of what we do.

The sense that the *content* of lived experience, whatever else one says (or whatever Wittgenstein has said), is, as a brute fact of life, metaphysically private is one that seems to want to survive Wittgenstein's reflections to this point, and there is good reason to think that this sense, however incompatible with all the ground Wittgenstein has covered and however clearly picture-driven in all the ways heretofore considered, was felt by Wittgenstein himself. For after finishing part I of *Philosophical Investigations* in 1945, he turned (in 1946) to problems exclusively in the philosophy of psychology, to which he devoted the following three years almost without interruption. And it is not long before the question of the content of experience comes up. In *Remarks on the Philosophy of Psychology*,[21] volume i, §109, Wittgenstein asks: 'Where do we get the

---

[20] I borrow this telling phrase from David Pears, *The False Prison: A Study of the Development of Wittgenstein's Philosophy*, 2 vols (Oxford: Clarendon Press, 1987).

[21] 2 vols (Oxford: Blackwell, 1980), i, ed. G. E. M. Anscombe and G. H. von Wright, ii, ed. G. H. von Wright and Heikki Nyman.

concept of the "content" of an experience from?' This idea, of course, co-conspires with those guilty conceptions and pictures, and in giving us a dualistic conceptual model for the content of experience, at the same time gives us a conceptual model for the content of autobiography. At this stage of his reflections the answer to his question is clear, and he wastes no time in answering it. He writes—giving voice to his early way of thinking even now, only a few years before his last writings in *On Certainty*[22]—'Well, the content of an experience is the private object, the sense-datum, the "object" that I grasp immediately with the mental eye, ear, etc.' That private object[23] would be the inner object upon which the private diarist introspects,[24] and (for our present concerns) it would be the inner representation of world-constituting Schopenhauerian significance, knowable only to the autobiographer, and upon which that self-describing, self's-world-defining author introspects in order to capture the inner content for which the autobiographical writing serves as external descriptive-narrative counterpart. Thus, Wittgenstein adds, encapsulating in a phrase and placing the weight of all the previous work in *Philosophical Investigations* and the still earlier *Blue Book* against this entire way of construing our experience and its language, the words 'the inner picture'. Reminding himself, and us, once again of the gulf that separates this picture from the particularities of our practices, he then asks: '—But where does one find one needs this concept?'

The notion of the private content of experience, its attendant conceptual confusions, and the misleading analogies that give rise to the notion in the first place (i.e. the concept of the inspection of an object as it shapes our parallel notion of introspection and inner object; the concept of a description of an outward object generating the deeply misled notion of expressive utterances,[25] or avowals,[26] as inward-directed descriptions; and so forth) submerge and resurface recurrently throughout

[22] Ed. G. E. M. Anscombe and G. H. von Wright (Oxford: Blackwell, 1969).

[23] In *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958), §§243–363 (but not only here; these trace back to the lectures on private experience given in 1934–6). I offer a sketch of the significance this holds for an understanding of artistic creativity in *Art as Language: Wittgenstein, Meaning, and Aesthetic Theory* (Ithaca, NY: Cornell University Press, 1995), ch. 6: 'The Silence of Aesthetic Solipsism', 118–35.

[24] I take this matter up directly, in connection with Dostoevsky, in Ch. 4, Sect. 3, below.

[25] Ch. 7, Sect. 2, offers a fuller examination of the problems generated by the grammatical similarities between outward and inward objects.

[26] This is the primary subject of Ch. 3, Sect. 2, below.

the remarks of this period. But it comes to the surface explicitly as late as §896, indicating that this is indeed, as a problem that first appeared in his earliest writings, never very far from Wittgenstein's concerns. He begins by announcing the topic, and then proceeds to articulate the impulse one feels to speak in ways dictated by the underlying picture:

The *content* of experience. One would like to say 'I see red *thus*,' 'I hear the note that you strike *thus*,' 'I feel pleasure *thus*,' 'I feel sorrow *thus*,' or even '*This* is what one feels when one is sad, *this*, when one is glad,' etc. One would like to people a world, analogous to the physical one, with these *thus*es and *this*es. But this makes sense only where there is a picture of *what* is *experienced*, to which one can point as one makes these statements.

'One would like', meaning to feel impelled in the grip of a picture, to people such a world with these '*thuses* and *thises*', and that is precisely what we do undertake if we regard autobiographical, or any kind of self-descriptive (dangerously put) or self-investigative writing, as the narrative externalization of that content. The *use* of such sentences is, again, very much unlike that suggested by the picture, and real introspection,[27] or self-reflection, as it *actually* occurs in contexts of human inquiry is the kind of thing we can see for what it is, without dualistic-prismatic distortions, only when free of the picture that would falsely give the '*thus*es and *this*es' a sense. And if we *cannot* point inwardly to the inner 'what', how then should we proceed? Not by attempting to capture the essence of consciousness in the act of Cartesian introspection — where the content of experience is knowable only unto itself, and, in grander terms, where the world is made of those representations — but rather by *turning to cases*.

3. REAL PRIVACY (AND HIDDEN CONTENT)

It was Plato who first wrote of the tripartite self, introducing near the beginning of our tradition a conceptual model of the self that, by its very structure, secured the possibility of both inner harmony, and if that is possible, inner disharmony. (He himself established this inner divisibility with an example of a person who, on seeing a fallen soldier in a ditch, feels arising in himself a morbid curiosity to inspect the corpse more closely but at the same time a higher inner voice

---

[27] This is the subject of Ch. 3, Sect. 3.

instructing him to rise above this low impulse; he gives in, saying to himself 'Go ahead, damn you, and feast your eyes on this banquet for sordid appetites!'[28]) Rousseau, in his *Confessions*,[29] describes himself in accordance with the conceptual model of an internally divided self, and one would initially think of this as a case that quite clearly suggests that the human phenomenon in which we are interested here is hermetically sealed within the Cartesian interior—we are after all speaking of *inner* disharmony. Rousseau writes:

Thus there began to form in me . . . a heart at once proud and affectionate, and a character effeminate and inflexible, which by always wavering between weakness and courage, between self-indulgence and virtue, has throughout my life set me in conflict with myself, to such an effect that abstinence and enjoyment, pleasure and prudence have alike eluded me.[30]

It is all too easy to bring to this everything Wittgenstein encapsulated as 'the inner picture'. But then where indeed do we have a *need* for this concept? In understanding the sentences Rousseau employed to describe this characterological double tendency, we need to *contextualize* Rousseau's remarks (as he himself does in the pages from which the passage is taken), considering cases of his deeds, his words, and combinations thereof in which one can see sometimes pride, sometimes a sentimental affection that seems incompatible with that proud heart, or inflexibility and courage manifest in one strand of his engagements with life, and weakness and self-indulgence manifest through another. And in understanding the result he describes, i.e. a person who is as a consequence at once enduringly anhedonic and chronically imprudent, we need, not access to a ghostly realm, but access to those multiple engagements in life, and a grasp of the way those many and diverse of human action and interaction do plausibly divide into each of the two strands. Moreover, it is a grasp of such cases, such particulars, that will give us an understanding of the person sufficient to see that this self-description is an exaggeration made in the interest of dramatic flair as well as neat conformity to, indeed, a philosophical picture.[31] Importantly, it is a grasp of those lived particulars that allows us to

---

[28] Plato, *Republic*, trans. R. W. Sterling and W. C. Scott (New York: W. W. Norton, 1985), 134. [29] Trans. J. M. Cohen (New York: Penguin, 1953).

[30] Ibid. 23.

[31] This is hardly the only time Rousseau wrote in correspondence to philosophical-conceptual pictures of selfhood (thus showing, if it needs to be shown, that cases of autobiographical writing are not *ipso facto* antidotes to misleading philosophical pictures). He wrote, to take one example, 'the first, the greatest, the strongest, the most

judge (or if 'judge' is too harsh a word, develop a sense of the veracity or plausibility of) the self-description. We do *not* make this judgment by holding the sentence of self-description up against the inward psychic state to determine its correspondence to inner states of affairs. Where, as Wittgenstein inquires, *do* we need this picture?

When Jill Ker Conway describes Frederick Douglass as 'the most articulate chronicler of the male American runaway slave's experience',[32] the articulateness to which she refers is not Douglass's ability to describe an inner state in outward terms as the picture would suggest. And she puts the word 'inner' to work in the way that we, in ordinary language, understand it: she writes that Douglass 'begins the account of his life by showing the reader an inner world of emotional and physical suffering, and grief at betrayal, which is charged with romantic passion'.[33] The depiction of the inner world is accomplished by showing him being torn from the security of his grandmother's cabin as a youth, teaching himself to read against all odds, his rejection of and disgust with the slaveholders, a life-defining fight with a particularly brutal slave driver, and the growing anger, disgust, and unrelenting moral horror that led to that violent confrontation. It is formative experience, followed by a long strand of increasingly interconnected thought, speech, and action, that makes what we call his inner struggle intelligible, and the interconnections of that strand are strengthened in

inextinguishable of all my needs was entirely one of heart. It was the need for intimate companionship, for a companionship as intimate as possible, which was the chief reason I needed a woman rather than a man . . . This singular need was such that the most intimate physical union could not fulfill it; only two souls in the same body would have sufficed. Failing that I always felt a void' (*Confessions*, 386, quoted and helpfully discussed in Jill Ker Conway, *When Memory Speaks: Reflections on Autobiography* (New York: Alfred A. Knopf, 1998), 23). This enduring, emotionally sensed void is one consequence of picturing the self in terms of a metaphysical dualism that engenders the problems of other-minds skepticism and, in its most virulent form, solipsistic privacy. The conceptually therapeutic methods articulated by Wittgenstein in his later philosophy are intended to dislodge such conceptual pictures or heuristic models by effecting a change in one's way of seeing, not only the problem before us, but also the conceptual substructure, the framework upon which that problem is built. I discuss this therapeutic conception of philosophical work directly in Ch. 7, Sect. 3, below.

[32] Conway, *When Memory Speaks*, 23. As will be clear in what follows, I am deeply indebted to Conway's discerning work of assembling highly instructive cases (although she does not mention Wittgenstein, her collection of cases is reminiscent of Wittgenstein's notion of 'assembling reminders for a particular purpose'). Admittedly, the purposes to which I put the cases differ greatly from hers, but the indebtedness remains (indeed, I would never have encountered many of the more obscure cases she presents and discusses), as will be clear in the following.   [33] Ibid. 23.

the retrospective writing.[34] These are events, with an increasing sense of moral teleology, in a person's life that we would only call 'external' or 'outward' when thinking of them in relation to the inner picture. Thus, it would be wrong, if in a sense a corrective (somewhat indeed like a ladder to be climbed and then kicked away), to say that we understand the inner through a grasp of the external, the outer. That way of putting the matter would lead us (i.e. grammatically mislead us) to see what we know of the person's life as *evidence* for the inner life—but the concept 'inner', as Wittgenstein is showing, does not *work* like that. We understand the coincidence, deeply meaningful for Douglass, that during the speaking tours of his later years he found himself speaking in the courthouse just beside the very jail in which he had been incarcerated for his failed first break for freedom, not in terms of ontologically hidden deep inwardly knowable meanings, but in terms of the particularities of his experience and the context, or in Wittgenstein's sense, the form of life, within which those particularities assume significance.

Conway, noting an outburst of memoir-writing among women of the Progressive Era, turns to Jane Addams, and Conway observes that the 'extensive use of conditional tenses and the passive voice'[35] gives rise to questions concerning what we might call the transparent sincerity of the memoir. Conway herself, suggesting that Addams, by writing in this way, conceals the power she exerted in bringing about the events of her own life and thus presents a self that conforms more to the romantic image of the female than it does to Addams's life, reaches the exaggerated conclusion—too general to be convincing—that 'We can be sure that whenever women autobiographers are hiding behind the passive voice and the conditional tense, they are depicting events in which they acted forthrightly upon a preconceived, rational plan.'[36] Conway is surely right to say that the emergent patterns of such grammatical details in any autobiography are worth scrutiny, and that at least in some cases, we thus arrive at a question concerning that autobiographical writing as describing or as *making* the character of that life; she goes wrong in asserting that we will invariably know the conclusion to reach in the face of such grammatical patterns. This is not at all to endorse a generic skepticism in such matters—we *can* know, but not in any way that does

---

[34] I look into the ways such interconnecting strands can be strengthened by retrospective writing in Ch. 5, below.

[35] Conway, *When Memory Speaks*, 49.   [36] Ibid. 49–50.

not arise from a nuanced, detailed grasp of the particulars of the life in question.[37] And in cases where we find autobiographical dissimulation of any one of countless kinds, we will judge the truth of what is written in relation to what we can and do sensibly call the inner life, but where that inner life is assessed in the manner of the Douglass case above, and not in terms of Wittgenstein's inner picture. The intentions of a person, against which we would judge the dissimulation, are not ontologically kept secrets.[38]

But there are, of course, secrets. Ellen Glasgow (insightfully discussed by Conway), writer of an autobiography depicting her life in the aristocratic South in the late nineteenth- and early twentieth-century America, describes the moment of falling in love in a way that reminds us that Wittgenstein's 'attitude towards a soul' runs along a continuum, from barely noticing the presence of a person (but still, most significantly, *barely*; i.e. the attitude is ineliminably present as constitutive of person-perception) to being acutely aware of one person's presence in a large crowded room—where one's attention is, seemingly beyond volition, magnetized to a single human focal point despite where one is visually looking. Glasgow writes, 'I felt my gaze drawn back to him by some invisible thread of self-consciousness,'[39] and she adds the remarkable passage later, 'What I knew, through some vivid perception, was that the awareness was not on my side alone, that he was following my words and my gestures, that a circle of attraction divided us from those around us . . .'. Her articulation of this vivid perception provides an equally vivid description of the *Einstellung* in its heightened form, and she goes on to tell the tale of her seven-year secret affair with the man in question (who was married). Looking back on that part of her life and the way she lived it, she wrote: 'Only on the surface of things have I ever trod the beaten path. So long as I could keep from hurting anyone else, I

---

[37] It is noteworthy in connection with *some* autobiographical writing serving as antidotes to simplifying metaphysical pictures of selfhood, that Margaret Sanger (in *Margaret Sanger: An Autobiography* (New York: W. W. Norton, 1938), 86–7, quoted in Conway, *When Memory Speaks*, 53), describing her life's work, argues for just such a particularized understanding—and one that implicitly combats, not only the Cartesian picture, but also a kind of reductive-behaviorist picture of selfhood as well. She writes, 'A woman in childbirth was not merely a woman in childbirth. My expanded outlook included a view of her background, her potentialities as a human being, the kind of children she was bearing and what was going to happen to them.'

[38] See the foundational monograph by G. E. M. Anscombe, *Intention* (Oxford: Blackwell, 1957) for a broadly Wittgensteinian treatment of the subject.

[39] Conway, *When Memory Speaks*, 55.

have lived, as completely as it were possible, the life of my choice. I have been free.'[40] But secretly free: as Conway describes it, hers was 'a private world behind the mask of conformity'.[41] There is indeed a secret here, and it is perfectly sensible to speak of the mask of conformity behind which is lived the real life. But when we see this kind of language at *work*, we see that the mask is not one that conforms in any way to the philosophical picture of the inner versus the outer; it is not a hidden truth *in that sense*.

Continuing his writings in late 1948 and early 1949 (that appeared as *Remarks on the Philosophy of Psychology*), Wittgenstein wrote near the last of those manuscripts some remarks of central relevance to the gaining of a non-picture-driven, clarified overview of the hidden, or the secretive, against its misconceptualizations. In them (published as *Last Writings on the Philosophy of Psychology*, volume i, §974), he writes: 'Nothing is hidden here; and if I were to assume that there is something hidden the knowledge of this hidden thing would be of no interest. But I can hide my thoughts from someone by hiding my diary. And in this case I'm hiding something that might interest him.'[42] Glasgow's secret, what one might term the ontological nature of her hiding, should be placed on a continuum of cases next to the hiding of a diary and not at all next to the imagined metaphysical 'hiding' (where indeed the inner is modeled on the outer and the grammar of hiding objects generates the picture of hiding inner, private objects) within a Cartesian interior. The life of her secret affair is *lived* by a human being, not by a hermetically sealed point of inward consciousness, as is what we will intelligibly call her public life that masks what we similarly understand—*in* the larger stream of life—as her private life. And if the analogy-influenced direction of our thoughts turns back once again to the notion that the mental, whatever else one may say about the real-versus-the-metaphysical understanding of the hidden, is *always*, despite these usages, still hidden in a special way, Wittgenstein writes in §976: 'What I say to myself silently he doesn't know: but again this isn't a matter of a "mental process," although there may be a physical process taking place here which might do instead of words spoken out loud if the other did know it. So also a physical

---

[40] *When Memory Speaks*, 55.   [41] Ibid.
[42] Ludwig Wittgenstein, *Last Writings on the Philosophy of Psychology*, 2 vols, ed. G. H. von Wright and Heikki Nyman, trans. C. G. Luckhardt and Maximilian A. E. Aue (Oxford: Blackwell, 1982). I will return to the notion of hidden content and ways in which it can be misleadingly put to work in retrospective explanations of actions in Ch. 5, below.

process here might be called ''hidden''.' If we do make sense of the word 'hidden' here, it does not, against the inner picture, refer exclusively to the mental—just as Glasgow lived both a public and a private life, but, if we contextually render intelligible the words 'public' and 'private', we grasp them in a way that speaks against the grammatically misled way of construing the meanings of those words.

Siegfried Sassoon's war memoir offers a description of events that yields a sense of what it means to find oneself in a position where one can only guess at another's private thoughts. He writes that his company had been issued orders to expand a newly captured German trench, but that just as they arrived he was handed new overriding orders to go back. He writes (with an honesty concerning the unclarity of motive that itself unsettles the privacy picture): 'Just as we got there a second runner overtook us to say that my bombers were to go back again. I sent them back. I cannot say why I went on myself; but I did, and Kindle [his Lance Corporal] stayed with me.'[43] He continues to describe a scene of increasing confusion in battle, where he comes upon the body of a just-killed German officer, and he writes of 'an impulse' that made him 'lift him up from the miserable ditch',[44] propping the fallen enemy officer against a small embankment and wiping mud from the officer's eyes and mouth with his coat-sleeve. One could do worse than to characterize this as a momentary triumph of 'an attitude towards a soul' over the competing conception of a dehumanized enemy, and what Wittgenstein describes as 'discerning the humanity in a person' Sassoon illustrates very well. He writes, 'Hoisting him a little higher, I thought what a gentle face he had . . . ', and he goes on to give voice to what he describes as his 'dim'—but as is clear to any reader, rapidly strengthening—'sense of the futility which had put an end to this good-looking youth'. (Here again, his *speculation* on that dim sense—his words are 'Perhaps I had some dim sense . . . '—itself repudiates any residual notion to which we might at this relatively late point cling of inward retrospective transparency.) Sassoon shortly reports that Kindle, for whom he felt a special responsibility owing to both his youth and his staying with Sassoon while the others retreated (no doubt only exacerbated because he himself did not feel sure he knew why he stayed when ordered to retreat), was killed, immediately provoking in him an extremely dangerous rage-fueled attack on the German sniper who found Kindle in his sights. Later, he returns to headquarters, passing

---

[43] Conway, *When Memory Speaks*, 77.   [44] Ibid.

the bodies of Kindle, the German, and very many others en route, and he was overcome by the dismayed sense of utter pointlessness of losing thousands of lives over a few hundred yards of bombed-out land in rural France. Here we will not say that we can only guess at his private thoughts: he has expressed them forcefully—and the only one we would *call* private is the last sense of dismay at irrational and wholesale slaughter, where privacy is brought into the language-game in order to mark the contrast between what he could and could not say to his superiors at headquarters. But of what follows, we may well say this: back in London, he and his fellow survivors, many of whom were now amputees, blinded by shrapnel, or shell-shocked (i.e. clearly men who would never see battle again), were, incredibly, required to attend lectures on the nature of trench warfare given by a young staff officer who had never been in the field. *There* we say, subsequent to the distinction between the private and the public having been put into play to mark the particular contrast that it does, that we can only imagine his private thoughts at that moment.

Wittgenstein, continuing his remarks, writes in §977: ' "What I think silently to myself is hidden from him" can only mean that he cannot guess it, for this or that reason; but it does not mean that he cannot perceive it because it is in my soul.' It is likely that the young lecturing staff officer could not guess much of Sassoon's private thoughts, for the particular reason that he knew nothing of the experience that preceded Sassoon's arrival in that room along with the more general reason that he had not seen battle himself. Knowing more of Sassoon's experience, we can do somewhat better at guessing; his thoughts, or rather thoughts of the kind he might have had, are less hidden to us, and we say we can only guess at what he might have been thinking, not as a marker of any metaphysically enforced epistemic limit, but because this is a way of expressing our comprehension of the experiential preconditions of outraged indignation. Such a sentence is not used to identify an object that is unperceivable because it is hidden within a soul.

Glasgow, as we saw, lived both a public and a private life, with the private, mostly, hidden; Sassoon had both public and private thoughts, where for reasons of military authority he had—in that context—to keep his private thoughts private. And indeed, some of them would have likely been beyond the understanding of, for example, the inexperienced lecturer. But these forms of intelligible privacy and publicity are not

for metaphysical reasons beyond the lecturer's reach—it is not that *kind* of question of 'reach'. One might have said, looking at Glasgow's life, 'I wonder what's going on behind that public persona', and one might have said, of Sassoon in those lectures, 'I wonder what's going on behind that face'. Wittgenstein wrote, as his next remark late in those inquiries of 1949: 'You look at a face and say "I wonder what's going on behind that face?"—But you don't have to say that. The external does not have to be seen as a façade behind which the mental powers are at work.' And in a variant of this passage, even better for our purposes, he wrote: '—But you don't have to think that way. And if someone talks to me quite obviously holding nothing back then I'm not even tempted to think that way.' We do not read 'external' facial or bodily 'signs' as evidence for inner events; these sentences do not function in that way. And we ought not to think that because in some suitably particularized circumstances we can and do ask what is going on behind the public display that thus all person-perception should be modeled on guessing, or collecting outward evidence for, the hidden interior. The human experience of sensing that someone is holding something back can put such sentences to work; such a sense is hardly, as Wittgenstein is observing, the key to the universal nature of all human interaction. We think in accordance with those phrases, those wonderings, *when*, and only when, we have occasion to do so.

In another war memoir, by Peter Ryan,[45] we are presented with a story of an 18-year-old warrant officer in New Guinea, recently occupied by the Japanese, whose troop movements he was sent (with hopelessly inadequate provisions) to monitor and report on. After extensive and truly extraordinary, unrelenting, life-threatening difficulty behind Japanese lines, he was finally ordered back to the base, but just short of safety his group was attacked in what had been until recently fairly safe ground. He submerged himself under the slime of a disgusting swamp with only his nostrils protruding as he felt in his ears the pressure caused by the boots of the Japanese soldiers as they squeezed through the adjacent mud. Later, having developed malaria, he was evacuated, and while walking down the jungle airstrip to the evacuation plane he passed a Japanese soldier, very obviously mortally wounded, under severe interrogation by Australian intelligence officers. Ryan wrote: 'As

[45] Ibid. 79.

I looked at his face, wasted with fever and suffering, I suddenly felt more akin to him than to the Australians who would not let him die in peace. His eyes, wonderfully large and soft, met mine. In that brief second I hoped he could read the message in my face.' It is not only that Ryan discerned the humanity of the dying soldier and, here too *in extremis*, again showed the insuppressible presence of the *Einstellung*. It is more that in this fleeting moment of reflection upon a human face he wondered if what was going on in *him* was readable in his face. Such questions, in such contexts, do have meaning, and they can prove definitive of a person's subsequent character and allegiances; Ryan concludes his memoir with the rejection of his nationality and a newly cast and deeply felt sympathy for the sufferings of all soldiers in war. To understand this, and to comprehend the significance of his own question concerning the fundamental human solidarity the dying enemy may or may not have been able to read in his face, we need nothing of the 'hidden'—whatever that might be—in the sense to which Wittgenstein is objecting, nothing private *in that sense*. We do, however, need something private in the intelligible, contextualized sense that we have seen in Glasgow and Sassoon. Ryan wrote that, well beyond his initial naive enthusiasm for what he expected to be the high adventure of war, just before looking at the dying enemy soldier in that way he had come to recognize 'how useless your whole mission was, how futile and purposeless your death would have been, and, above all, when your sober but aching eye discerns that nobody whose business it might have been took the least trouble to see that you got at least a reasonable chance of living'.[46] Where we wonder what is going on behind a face, or where we wonder if someone can see what is going on behind ours, what we need is at least a sense of, and ideally an intimate knowledge of, the *lived* life that stands behind the utterance, the gesture, or the facial expression in question. *That* is the true substance of biography and, self-reflectively, autobiography. And, as Wittgenstein is arguing, none of that is hidden in the metaphysical sense. Were it not for misleading grammatical analogies to hidden objects and the bewitchments of language that generate what Wittgenstein so compactly labeled 'the inner picture', we would not, as he wrote in the variant to his remark 978 of March 1948, even be tempted to think that way.

---

[46] *When Memory Speaks*, 82.

It is now time to increase the level of magnification somewhat and to have a closer look at a number of these linguistic forces that, working in concert, lead us to conceive of selfhood in a manner consistent with a deeply entrenched philosophical tradition and—as we shall see in later chapters—instructively incompatible with our actual autobiographical practices.

2

The Self, Reflected

It is widely understood, first of all, that Wittgenstein's multifarious writings on language hold deep significance for the philosophy of mind, to an extent that, because of his writings as well as what has followed in their wake, many are now reluctant even to go so far as to draw a distinction between the philosophies of mind and of language. To achieve a fuller comprehension of language, to gain an overview of what Wittgenstein called in the Preface to *Philosophical Investigations*[1] the 'landscapes,' the sketching of which required his traversing 'a wide field of thought,' precisely is to gain an overview of many various and interrelated facets of the mind in its variegated deployments. Second, it is widely understood as well that Wittgenstein was fundamentally opposed to what he revealed to be a particularly pernicious conceptual model, or picture of linguistic meaning, the dualistic picture separating the inner from the outer. Third, it is equally widely understood that Wittgenstein's opposition to—or rather undercutting of—the inner–outer distinction is directed with as much force to its employment in the philosophy of mind—in our conceptual modeling of the human subject—as it is to its employment in the philosophy of language; thus this third point serves as one of the many supporting reasons for the first.

But there is, in this line of thought, a curious and at least seemingly strange inability to progress to a fourth point. One would like to be able to add to this list that the general overview of language, along with a developed appreciation of the significance of this work for our understanding of the mind, has brought us to a point of clarity concerning Wittgenstein's conception of the self. But this subject seems veiled; there is an air of mystery surrounding the conception of the self that lingers after we work through Wittgenstein's related yet separate discussions. Indeed, if there should be a widespread understanding here,

---

[1] Ludwig Wittgenstein, *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958).

or, given what we have to work with from the writings Wittgenstein has left us on this topic, then is the topic—indisputably among the most significant and, one imagines, deep questions not only of the philosophy of mind but rather of all philosophy—permanently veiled? Has Wittgenstein simply failed to provide an account anywhere near as full as his accounts of the related but different issues? Why, given the profundity of his work elsewhere in the philosophy of mind, would he leave the matter where he did?

The answer to this question requires our looking, in this chapter, somewhat more exactingly into not only where he left the matter, but where he started it, as well as into those surrounding issues closely related to the question of the nature of the self. These would include his remarks on philosophical seclusion and his intertwined critique of introspection; his stout rejection—and here perhaps not always so widely understood a rejection—of behaviorism; his equally forceful rejection of Cartesianism; his employment of the concept of avowal (supplanting our conception of the verbal *description* of an inner state with the verbal *expression* of an inner state—where the phrase 'inner state' is thus differently understood); his penetrating remarks on consciousness; and his positive or nondualistic conception, unfortunately more intimated than argued and explicated, of genuine introspection. And all of these, taken together, will lead to a position from which the epistemological value of those forms of literature in a sense closest to human beings, i.e. autobiography, and full literary depictions of mental life—in short, the literature of the self or, as it is now called, life-writing—can perhaps be perspicuously understood.

1. OBSERVING CONSCIOUSNESS

It cannot be plausibly argued that Wittgenstein simply did not care about the philosophical problem concerning the nature of the human subject. First, the problem of selfhood was certainly implicit, if not explicit, in the philosophical positions to which his work responded or reacted: it is clearly at the center of rationalism, most obviously, as we have already seen, in the Cartesian conception of the thinking thing, the inner point of consciousness that constitutes the self; it is in empiricism in the various forms of Locke's personal-identity problem, Berkeley's idealism, and notably Hume's 'bundle theory' of the self; it is in Kant's transcendental unity of apperception; it is in Schopenhauer's conception

of the Will; it is in James's introspectionist self in his psychology; it is in
Russell's work on the analysis of mind; and—perhaps most significantly
for Wittgenstein—it is in Frege's revised employment of the traditional
ontological distinction between the mental and the physical in his 'first
realm' of ideas which are private to the mind of the thinker and the
'second realm' of outward material objects. Second, the fact that this
concern is in truth deeply important to Wittgenstein and was so from an
early stage is clear from the memorable entry from 1916 in the *Notebooks*
that has already served as this book's opening epigraph: 'The I, the I
is what is deeply mysterious!'[2] But of course it is within the context of
his working on the *Tractatus Logico-Philosophicus* that the question fully
comes to the surface, and it is treated in a way suggesting that this may be
one of the areas in which Wittgenstein's thought remains fairly consistent
throughout his life, at any rate in terms of the conclusion—even if the
manner of reaching that conclusion differs in method and substance
considerably. He writes, fairly amazingly, at 5.631: 'there is no such
thing as the subject that thinks or entertains ideas'. But it is important
to realize that he has already said at 4.003—and here too is another line
of continuity in his thought from the earliest to the latest—'most of
the propositions and questions of philosophy arise from our failure to
understand the logic of our language'. There is indeed something—in
fact a number of different things[3]—that leads us to misconstrue the
logic of our language of the human subject, the self, the I; it is thus, even
at this very early stage of his work, a distinctively *philosophical* myth he
is opposing. As we shall see, the foregoing claim is perhaps not really
so amazing; it is not, for one thing, self-refuting, in that it is written

---

[2] *Notebooks, 1914–1916*, ed. and trans. G. E. M. Anscombe (Oxford: Blackwell, 1961), entry for 5 Aug. 1916. It is significant to the context of this entry that Wittgenstein was working at the time (1916–17) on the mystical themes in the *Tractatus Logico-Philosophicus* (trans. D. F. Pears and B. F. McGuinness (London: Routledge & Kegan Paul, 1961)).

[3] And indeed very many of these things are linguistic in nature, thus reconfirming the close, in fact indissoluble, link between the philosophies of language and of mind in Wittgenstein's work. One unambiguous symptom that language has misled us, that is, that an unwitting presumption *about* language has generated illusory conceptual difficulty, is the intuitively deep belief that, first given the (erroneous) presumption that all words get their meaning through direct reference (a belief that Wittgenstein subjects to the closest scrutiny in many different contexts of inquiry, throughout his conceptual 'landscape'), the first-person pronoun must also thus get its meaning through direct reference to, if not an outward, then an inward, substantive thing. Our understanding of the self, the 'I', is thus related to language in a way that certainly does not reduce to what is taken in some quarters to be mere semantics. Linguistic usage—and our close investigation of it—is metaphysically (and not merely semantically) significant.

by a thinking, idea-entertaining subject and communicated to another thinking subject, the reader, who is implicitly expected to entertain its central idea. There is, in short, not a kind of Cogito argument readily available against this claim. Wittgenstein is not self-contradictorily writing that a thinking writer is not now writing but rather that a particular—and particularly pernicious—philosophical picture of the self, the subject, is what is false, what does not exist. He later writes, in the *Blue Book*, that it is only through the tricks that language plays on us, it is only through the bewitchment of our intelligence by means of language, as he would go on to express the matter in *Philosophical Investigations*, that we come to believe in the myth that the first-person pronoun refers to 'something bodiless, which, however, has its seat in our bodies'. 'In fact,' he says, '*this* seems to be the real ego, the one of which it was said, "*Cogito, ergo sum*." '[4] One way to identify more accurately the mythical picture of the self to which he is opposed is to examine a number of his remarks on introspection and the picture,[5] the conceptual model, of inward psychological seclusion that it enforces.

It is apparent from virtually any inquiry into anything whatsoever that one needs an *object* upon which to focus one's attention, if we are to employ the concept of *inspecting*. But to go this far—which is seemingly not very far at all—is already to allow language to lead us astray, to establish a fundamental analogy—in truth a misleading analogy—between the mental and the physical, or, as Wittgenstein warns us, to lay the foundation for conceptually modeling the mental on the physical. To inspect, we think, we need an object; to introspect, we consequently think, we need an object of another kind. And that object, if we are puzzled about, or indeed mystified by, the nature of the self, will be separated from the physical or the material, and the sense of mystery will be heightened as the separation proceeds. In *Philosophical Investigations* §412 Wittgenstein opens by referring to the 'feeling of an unbridgeable gulf between consciousness and brain process' but then just as quickly establishes the other pole, the pole of ordinary experience: 'how does it come about that this does not come into the considerations of our ordinary life?' Back to the philosophical

---

[4] *The Blue and Brown Books* (Oxford: Blackwell, 1958), 69.
[5] For some of Wittgenstein's uses of the term 'picture', meaning, roughly, a conceptual model of which we may be unaware, or if we are aware of it, one that we erroneously take as unproblematically given (often when this is the result of misleading grammatical appearance) and that shapes or determines our subsequent thought on any philosophical matter pertaining to the picture, see *Philosophical Investigations* §§422–7 and II, p. 223.

pole, he goes on by identifying a feeling that accompanies the thought: 'This idea of a difference in kind is accompanied by a slight giddiness,' and now moving again, but this time to a new position, one in which, grounded in ordinary experience, we are aware of philosophical artifice: '—which occurs when we are performing a piece of logical sleight of hand.' 'When', he now asks, 'does this feeling occur in the present case?' The answer to this question, along with his subsequent reaction to the answer, reveals the conceptual model of introspection that he finds unacceptable. 'It is when I, for example, turn my attention in a particular way onto my own consciousness, and, astonished, say to myself: Tʜɪs is supposed to be produced by a process in the brain!—as it were clutching my forehead—.' And this depiction of introspection is just as quickly met with a return back to the other pole: 'But what can it mean to speak of "turning my attention on to my own consciousness?" This is surely the queerest thing there could be!'

It is thus not a Humean point Wittgenstein is here making; he is not asserting that when we introspect, when we focus our inward gaze on the self, we find not a self but one of a larger number of particular experiences of sensation or reflection, impression or idea, so that we can never really arrive at anything but the bundle of impressions and ideas. He is rather pointing to the very *strangeness* of this way of describing an attempted mental act (and this phrase 'attempted mental act', too, as is naturally the case, is in turn a strange way of describing the strangeness). But it is the very idea of an 'inward gaze' that is fundamentally problematic, and Wittgenstein's following words can be read as an implicit criticism of that familiar philosophical notion; in what follows he does *not* use the word 'gaze' metaphorically: 'It was a particular act of gazing that I called doing this. I stared fixedly in front of me—but *not* at any particular point or object. My eyes were wide open, the brows not contracted (as they mostly are when I am interested in a particular object). No such interest preceded this gazing.' And his glance, he adds, 'was vacant', and underscoring the word '*like*' precisely because the glance only resembled, it did not equal, as there was no object upon which visual scrutiny rested, adds: 'or again *like* that of someone admiring the illumination of the sky and drinking in the light'. Wittgenstein closes this section by observing that the proposition he found paradoxical in a philosophical voice ('Tʜɪs is produced by a brain process!') in another, ordinary voice has nothing at all paradoxical about it; it might have been said in a neurophysiological experiment in which the effect of light is produced in a subject by stimulating a part

of the brain. He adds two important final thoughts in this section, one concerning the social context of the utterance, the other concerning the psychological state of the imagined speaker, neither of which involves any inner occult process for its comprehension: 'But I did not utter the sentence in the surroundings in which it would have had an everyday and unparadoxical sense. And my attention was not such as would have accorded with making an experiment. (If it had been, my look would have been intent, not vacant.)' We thus have the grounds upon which Wittgenstein says, opening §413, 'Here we have a case of introspection,' and he further suggests that it is from a similar case that William James derived his conception of the self.

Seeing a direct connection here to the philosophy of language, he notes that 'James's introspection shewed, not the meaning of the word "self" (so far as it means something like "person," "human being," "he himself," "I myself"), nor any analysis of such a thing.' This is a telling remark in relation to our fundamental questions concerning Wittgenstein's conception of the self and the conceptions of the self to which he is opposed. First, it clearly indicates an anti-referential position with regard to the self; introspection does *not* show the meaning of the word 'self', supported by the twin claims that, (1) more specifically, the introspective act only *pretends* to focus on an object that would serve as an inward referent of the word 'self', and (2) more generally, we should free ourselves from the illusion (also generated by the misleading analogies language readily offers us, specifically that all words function like names, which goes back to the opening[6] of *Philosophical Investigations*) that words get their meanings exclusively from the objects to which they refer. Second, it moves, again, between the poles of the philosophical voice and the ordinary voice, and, speaking in the latter, Wittgenstein reminds us, if only in brief, of the contextualized employments of terms or phrases such as 'persons', 'human beings', and 'he himself'. The grammatically misled philosophical conception of the introspective self is utterly remote from our *uses* of such phrases; again, Wittgenstein is implicitly relying on the sense of strangeness[7] as an indicator of

---

[6] For two interwoven sets of remarks on the opening passages of *Philosophical Investigations* (written from different points in time) that gradually unfold what is at stake in those passages, see Stanley Cavell, 'Notes and Afterthoughts on the Opening of Wittgenstein's *Investigations*', in *Philosophical Passages*: *Wittgenstein, Emerson, Austin, Derrida* (Oxford: Blackwell, 1995).

[7] The instructive and enlightening employment of this criterion of strangeness (along with a highly nuanced sensitivity to it) was given one of its best expressions in the work

conceptual confusion or, in this particular case, of the presence of a philosophical myth. Third, Wittgenstein adds, significantly if briefly, that James's introspective project did not provide an *analysis* of the concept of the self; it does not in truth go behind or beneath the allegedly superficial appearance of language in anything like the style of his own earlier and Russell's atomistic works.[8] The stance taken here is thus anti-analytical (in this restricted sense of atomistic analysis only; it certainly is not against conceptual *clarification*) as well as anti-referential. Both of these, Wittgenstein is indicating, would lead us astray—the referential leading us into inward reification, the analytical leading us to look *beneath* precisely what it is we need to look at *directly*, that is, the language-games in which self-terms operate.[9]

Indeed, what James's introspective project shows is rather 'the state of a philosopher's attention when he says the word "self" to himself and starts to analyze its meaning'. Wittgenstein adds parenthetically that a good deal could be learned from this, which, given its clear relation to the belief with which Wittgenstein is concerned throughout his late philosophy, namely, that the meaning of a word is its referent, seems true enough. But at present it is important to see one of the things that Wittgenstein is *not* saying in this passage. He is not arguing that introspection, as a method for achieving self-knowledge, is fallible and therefore of dubious epistemological value. This is in fact the position James himself develops,[10] and thus on this point James differs substantially from the Cartesian position, which holds that introspection, because of its

---

of J. L. Austin; see in this connection the extraordinarily close study of the philosophical voice's claims and pronouncements on perception in his *Sense and Sensibilia*, ed. G. J. Warnock (Oxford: Oxford University Press, 1962).

[8] For an excellent overview of the position of Wittgenstein's late philosophy in relation to both earlier and later analytical work in the philosophy of language, see P. M. S. Hacker, *Wittgenstein's Place in Twentieth-Century Analytical Philosophy* (Oxford: Blackwell, 1996).

[9] There are two particularly helpful sets of metaphors on the matter of the distinction between (1) satisfying the philosophical impulse to attempt to look *beneath* the cultural phenomena in question versus (2) remembering the value of looking *directly at* those phenomena; see Hans-Johann Glock's distinction between 'logical geology' and 'logical geography' in *A Wittgenstein Dictionary* (Oxford: Blackwell, 1996), 278–83, and Cavell's distinction between the 'horizontal' and 'vertical' conceptions of the difficult Wittgensteinian phrase 'forms of life' in his 'Epilogue: The *Investigations'* Everyday Aesthetics of Itself', in *The Cavell Reader*, ed. Stephen Mulhall (Oxford: Blackwell, 1996).

[10] Centrally, in *The Principles of Psychology*, ed. Frederick Burkhardt, 3 vols (Cambridge, Mass.: Harvard University Press, 1981), although James's full development of his concept of the mind runs throughout his many writings.

unmediated nature, is infallible. Wittgenstein's point is *very* different; it is, again, not only that the idea of introspection as construed within this philosophical voice is mythological and that the self upon which we place our introspective gaze is a part of conceptual mythology as well, but also that we do not in truth *perceive* or *observe* our own consciousness.[11] One way briefly to recall to mind this element of Wittgenstein's philosophy as it pertains to the question of the self is to try to find an intelligible context in which we can give a ready answer to the question 'How do we know we are in pain?' This linguistic exercise suggests that *knowledge* is an ill-suited concept for this kind of context, because there is not an epistemic divide between there being a pain and our knowing of it. There is such a divide, by contrast, with external objects; there can be a book on my desk without my knowing that it is there. Thus, here too the grammar of the external world leads us astray when the internal world is modeled upon it. An insufficient intricacy in these matters, as Wittgenstein repeatedly shows, is an ever-present danger: to speak of the 'internal world' in contrast to the external is to incline our subsequent thinking in turn toward the idea of an inner mental world populated by mental objects, or, as Ryle put it, a private stage,[12] and this is precisely the picture of the mind Wittgenstein is combating in his multi-front war with grammatically induced misconceptions. The fundamental point, possessing the greatest power to undercut the very idea of Cartesian or Jamesian introspection, is thus again, contrary to our preconceptions that exert their power only so long as we remain under the influence of misleading analogies, that we do not perceive or observe consciousness.[13]

---

[11] For closely related remarks outside of *Philosophical Investigations*, see 'Wittgenstein's Notes for Lectures on "Private Experience" and "Sense Data" ', ed. Rush Rhees, *Philosophical Review*, 77 (1968), 275–320; see esp. 278–80. See also 'The Language of Sense Data and Private Experience: Notes Taken by R. Rhees of Wittgenstein's Lectures, 1936', *Philosophical Investigations*, 7 (1984), 1–45, 101–40; see esp. 111–12.

[12] See Gilbert Ryle, *The Concept of Mind* (New York: Barnes and Noble, 1949); in connection with present issues see esp. his classic chapter 'Self-Knowledge'.

[13] Indeed one might think, given that we speak of consciousness, that there must be a kind of 'second-order' consciousness from which we observe first-order consciousness (in order—and here again we are misled by the presumption concerning meaning-as-naming—to use the word 'consciousness' meaningfully, i.e. to attach the name to the thing). Against this picture-induced confusion, compare Wittgenstein's remark about philosophy itself in *Philosophical Investigations* §121: 'One might think: if philosophy speaks of the use of the word "philosophy" there must be a second-order philosophy. But it is not so: it is, rather, like the case of orthography, which deals with the word "orthography" among others without then being second-order.'

In *Philosophical Investigations* §416 Wittgenstein writes, again in the interlocutor's philosophical voice: 'Human beings agree in saying that they see, hear, feel, and so on (even though some are blind and some are deaf). So they are their own witnesses that they have consciousness.' The intuitively plausible introspectionist idea being put forward is clear enough, and Wittgenstein again responds to it immediately with a return to the ordinary voice, again employing the criterion of strangeness: '——But how strange this is! Whom do I really inform, if I say "I have consciousness"? What is the purpose of saying this to myself and how can another person understand me?' And, once again, returning philosophical language to its ordinary employment, Wittgenstein observes that we *do* say such things in medical contexts, contexts of repeated fainting ('I am conscious again'), and so forth. The challenge, indeed the assault, contained within this remark is powerful and profound. Can we get so far as even to *understand* the claim that would seem absolutely necessary to the articulation of the Cartesian conception of the self? Or are we relying, at the most fundamental level in the articulation of this model of the self, on a form of expression, on a phrase allegedly self-defining (in both senses, in that it would deliver its own first-person meaning intrinsically as well as give definition to the word 'self') that in truth delivers sense only if we take it away from its philosophical and remove it back to its ordinary voice—in which case it means something (as in the medical contexts) but (and here is the potency of the observation) not what it alleges or pretends to in its philosophical position. Disguised nonsense was the particular variety of nonsense that interested Wittgenstein the most in his later works, and one of the reasons for this is perhaps that the disguises can present appearances not only of seemingly obvious irrefutability ('I know I am conscious because I observe it, I perceive it') but also appearances of capturing the essence of the mysterious 'I'.

The argument gains clarity in the next section. In §417 Wittgenstein asks: 'Do I observe myself, then, and perceive that I am seeing or conscious?' This question is answered, indirectly, with a question asking why we should talk about *observation* at all; indeed, why not simply say that we *perceive* we are conscious, since the act of observation seems clearly otiose if we already perceive our consciousness? But again—and Wittgenstein is moving through intermediate steps from the philosophical to the ordinary voice—we can ask, in an equally direct manner, why should we say *perceive*? Why not just say 'I am conscious', since the act of *perceiving* now seems equally otiose? There

is, I believe, at this point a missing step that needs to be inserted, or made explicit if it is implicit, in order to show the force of Wittgenstein's position; it is also at the point of this missing step that the argument becomes significantly more intricate. If indeed we remove *observation* from the claim of first-person consciousness, then the very concept of introspection is losing its content: what would it mean to say that introspection is the inviolable source of knowledge of the self if the very act of in(or inward)spection, that is, the experiential substance of the word 'introspection', is absent?

We might then say, moving to a position closer to the Cartesian than to the Jamesian, that we still know it because we *perceive* it immediately, without the mediating mental act of in(tro)spection, but this retort in the philosophical voice has a two-pronged answer, here only one of which Wittgenstein has provided explicitly. The first prong, as we have seen, is to remove the mental act of perceiving just as we did observing, thus severing consciousness from the perception of it. The second prong is to make clear that this strands the philosophical voice, leaving it isolated from the *knowledge* of consciousness, from independently assessing the fact of consciousness (as, for example, we might imagine an animal being conscious without simultaneously 'knowing' of its consciousness). And without the capacity to assess independently the fact of consciousness, without the autonomy of the knowing subject who declares himself conscious, the content of the 'I' seems irremediably unclear in the first-person claim of knowledge of consciousness.

Another way to put this is to say that the second 'I' in the judgment 'I perceive I am conscious' is not isolable from the consciousness, and thus to note that it possesses this property seems at the very least otiose, if not unintelligible. And the initial phrase 'I perceive', where the object of perception is the consciousness of the 'I', seems, to put it another way, reflexively claustrophobic: the 'I' can never get away from the consciousness it (allegedly independently) attributes to itself or judges itself to possess. But to simplify, if, having removed observation and perception, we say now only 'I am conscious', we might well understand this—if only in the ordinary voice—but then it does not in any clear (or other) way rest on a foundation of introspection; indeed, it appears impossible to wedge introspection into this context. Wittgenstein's next explicit step, again with growing intricacy, is to remark '—But don't the words "I perceive" here shew that I am attending to my consciousness?—which is ordinarily not the case.' If this is true, he then observes, the sentence 'I perceive I am conscious' does not perform the

job of saying that we are conscious, but rather the job—a very different one, and one that again brings the philosophical back to the ordinary voice—of saying that our attention is disposed in a particular way. And again he closes this section by asking his question that continually reaffirms the necessity of context for intelligibility: 'In what situations do we say it?'

2. THE PICTURE OF METAPHYSICAL SECLUSION

Despite the qualms one may have, after following Wittgenstein's discussion, about the very possibility of getting so far as to comprehend the concept of introspection, much less to determine its precise character both as a mental act and as a fundamental source of knowledge of the self, one may still feel assured in positing the ultimate seclusion of each individual mind, each private self.[14] Wittgenstein also argues against this picture of the self, or more precisely, against this conception of the circumstances in which the 'I' finds, indeed, itself. In *Philosophical Investigations*, part II, §xi, we find the errant interlocutor, or the philosophical voice, again speaking a language bewitched by the tricks of grammar and misleading analogies. He makes this assertion: 'A man's thinking goes on within his consciousness in a seclusion in comparison with which any physical seclusion is an exhibition to public view' (p. 222).' This gives clear and forceful articulation to the view that many find philosophically obvious as the first brute fact of life.

Wittgenstein's initial step in unsettling this conception, this picture of the self's most fundamental predicament, is to respond in the interrogatory: Would people who were somehow able to 'read' the 'silent internal discourse' of other people—and here he adds, significantly (in that his remark suggests a *physical* rather than ghostly or immaterial mode of access, or mind-reading, to that inward discourse), that they accomplished this remarkable feat by closely observing the larynx—be inclined to employ this picture of complete seclusion, the picture of ultimately inviolable isolation that in turn motivates and nourishes other-minds skepticism, and at the logical extreme, solipsism? His

[14]   For a discussion and particularly helpful set of references throughout Wittgenstein's writings on this topic (and from which my section title is borrowed with emendation), see Garth Hallett, *A Companion to Wittgenstein's 'Philosophical Investigations'* (Ithaca, NY: Cornell University Press, 1977), ch. 38.

implication is clearly that they would not. Yet, for a number of interrelated, and as it were conspiring, reasons—here including (1) a misconception of language in dualistic terms of physical, outward signs and internal, mental meanings (thus lodging the permanent possibility of skepticism within language itself), along with (2) a corresponding misconception of the possibility, indeed the naturalness, of a private language that only the speaker can understand (because the inner referents of this language are private experiences[15])—we are, or can easily be, strongly inclined to say that they (and we, in that imagined condition of inward-discourse readability) indeed *would* embrace the picture of the metaphysically secluded self. It seems that the inner world is, in this distinct, ontological sense (in that the self is not and could not be the *kind* of thing open to public view), *hidden* in an even more extreme way than we saw in Chapter 1. Wittgenstein immediately takes this use of 'hidden' back to the ordinary voice, beyond the grasp of this particular manifestation of skepticism, by simply observing: 'If I were to talk to myself out loud in a language not understood by those present my thoughts would be hidden from them.' This bracing observation performs two services simultaneously, in that it contextualizes the concept 'hidden' much in the way we saw it done in connection with autobiographical writing in Chapter 1, reminding us where we readily understand it—where we use it—and thus, here again, how remote this usage is from its philosophical guise. And it reminds us, but here in philosophy rather than in literature, that the comprehension of 'hidden' in this sentence does not necessitate—and in this case in fact excludes—any metaphysical conception of an inner, private, ghostly realm wherein thoughts are hidden.

Wittgenstein continues to develop his argument in this vein, observing, in the imaginary case of a person who always guessed right what we were saying to ourselves in our thoughts, that the criterion for his guessing right is that we are truthful and confess that he has guessed right. And, severing the picture of secluded inner thought from the understanding of 'right', he observes that the criteria for truth of such a confession of thought are not coequal with the criteria for the description of a process; moreover, the importance of the true confession of thought 'does not reside in its being a correct and certain report of a process'. Thus, against the predictions cast by the picture of inner seclusion, when we look to the actual details of lived experience, we find—to

---

[15] I return to this matter in Ch.4, Sect. 3, below.

greatly abbreviate[16]—that what it means to understand a person is not
equivalent to what it means to understand the descriptions of the inner
processes of that person's private mind. And again, what will strike us as
important about a confession of inward thought is not that the report
of such an inner process is certain, but rather—and here too we see an
extreme abbreviation—that it is given 'in the special consequences' of
the confession, whose truth is guaranteed not by its correspondence to
an inner process but by 'the special criteria of *truthfulness*'.[17]

There is a further example Wittgenstein gives in this discussion
that serves to reorient us away from the conception of psychological
interiority. He mentions a game of guessing thoughts with variants: A
speaks a language B does not, and B is supposed to guess what A meant;
or A writes down a sentence unseen by B and B has to guess its sense;
or A is putting a jigsaw puzzle together and, although B cannot see A,
periodically guesses A's thoughts, saying things like 'Now where is this
bit?', 'The sky is always the hardest part', and so forth (p. 223). In each
case, what is hidden is not *metaphysically* hidden; the meaning of the

---

[16] One might characterize one of the contrasts between philosophy and literature in
precisely these terms; the latter does not in philosophy's sense abbreviate, and thus itself
constitutes both a corrective and an invaluable resource for philosophical understanding.
The diagnosis of the *impulse* to abbreviate in this sense is still another matter, about
which Wittgenstein has said a good deal in his various remarks on philosophical method
and, broadly speaking, his wisely cautious analogies between his later philosophical style
and psychoanalysis.

[17] Even the phrase 'the special criteria of truthfulness' is too abbreviated and thus
misleading, suggesting both uniformity among them and that they travel, as a coherent
and invariant set, from context to context. Such criteria are best *shown* in literature, and
perhaps best examined in the philosophical criticism of literature. (To show them *in
detail* is the only way to render them visible.) In this connection it is perhaps worth
noting that the many episodes in Wittgenstein's life as they pertain to truthfulness (as
discussed in Ray Monk's very fine biography or, for example, in Fania Pascal's memoir)
display philosophical significance. See Ray Monk, *Ludwig Wittgenstein: The Duty of
Genius* (New York: Free Press, 1990), esp. the chapter 'Confessions', and particularly
the observation that 'it is no coincidence that Wittgenstein wrote the set of remarks with
which he remained most satisfied at a time when he was most ruthlessly honest about
himself' (p. 367); and Fania Pascal, 'Wittgenstein: A Personal Memoir', in Rush Rhees
(ed.), *Recollections of Wittgenstein* (Oxford: Oxford University Press, 1984). Or to put
the matter in Gilbert Ryle's terms, there will not be one single overriding criterion of
autobiographical truthfulness that follows from 'Privileged Access' to 'facts of a special
[inner] status'. He writes: 'The fact that retrospection is autobiographical does not imply
that it gives us a Privileged Access to facts of a special status. But of course it does give
us a mass of data contributory to our appreciations of our own conduct and qualities
of mind. A diary is not a chronicle of ghostly episodes, but it is a valuable source of
information about the diarist's character, wits and career' (*Concept of Mind*, 167). The
criterion is not, in short, fidelity to *ghostly* episodes.

unknown language, the hidden words—these contextualize the word 'hidden' in ways that fail, instructively, to correspond to the picture of the self's complete seclusion. But does the puzzle case really do this, exactly? Here we may insist, under the influence of the picture, that the *thoughts* are themselves hidden. But Wittgenstein's point operates at a more subtle level: if B is right on occasion, his rightness does not depend on a correspondence between his sentences and the silent or out-loud utterances of A; indeed, Wittgenstein adds, 'but *I* need not be talking to myself either out loud or silently at the time'. What this discussion points to is: (1) the language in which we express our convictions concerning first-person privacy may be language that we in truth do not understand; (2) the picture of seclusion is motivated in part by, ironically, a misunderstanding of what it actually is to understand people; and (3) the picture of seclusion is nourished, again, not only by a large-scale misconception of linguistic meaning but more narrowly by an erroneous conception of the meaning-determinants, i.e. inner referents, of emotion-terms.

Each of these topics will resurface with differing inflections in what follows, but, to complete the reconsideration of this part of Wittgenstein's discussion as it pertains to the nature of selfhood, he now claims explicitly what he implied earlier, that thoughts are no more hidden than 'unperceived physical proceedings'. He suggests a new analogy for hiddenness. In reply to the interlocutor's philosophical voice restating with renewed emphasis (despite all these various conceptual desiderata) that 'what is *internal* is hidden from us', Wittgenstein says that the future is hidden from us, but that an astronomer is not thinking this way when he predicts an eclipse of the sun. He reminds us that when we see a person writhing in pain with clear cause we do not think that nevertheless his feelings are in truth hidden from us. He reminds us that we can in circumstantial fact be as '*certain* of someone else's sensations as of any fact', then adding, against the philosophical impulse to assimilate diversity into a simple uniformity, that this does not make reports on people's moods, mathematical calculations, and self-descriptions of the 'I am *x* years old' kind 'similar instruments', but rather, the certainty in each case is determined by the language-game in which it operates. There is an *appearance* of a psychological difference between the claims 'He is much depressed', '25 times 25 equals 625', and 'I am 60 years old', but the difference is rather logical, which at this point in his philosophy means that the language-games are different. We are not less certain that a person is in pain than that twice 2 equals

4; our inclination to believe this is yet another symptom of the widening influence of the picture of the metaphysically secluded self.[18]

The fairly uniform clothing of language can make very diverse language-games appear alike, and in response to the philosophical voice insisting 'While you can have complete certainty about someone else's state of mind, still it is always merely subjective, not objective, certainty,' Wittgenstein flatly replies: 'These two words betoken a difference between language-games' (p. 225). Again, a logical difference is misconstrued as a psychological one; the very word 'subjective' shows its danger[19] in calling us back to the philosophical voice, and the influence our own language holds over us (in making us want to say what fits the metaphysical picture) is approximating a condition of autoventriloquism.[20] But we should, Wittgenstein suggests, at this stage of these considerations show strong resistance if not outright rejection. If the doubt concerning our capacity to know the mind of another because of its inviolable privacy reenters through this (or any similarly metaphysically freighted) word, we should—having remembered the facts of our form of life[21] that are otherwise open to view—reject *artificial* doubt: The interlocutor says, 'But, if you are *certain*, isn't it that you are shutting your eyes in face of doubt?', and the reply comes, '—They are shut.'[22]

---

[18] The inclination to believe so is not *exclusively* a symptom of this picture of the self: Plato's divided line separating the sensory from the intellective, Descartes's fundamental epistemological dichotomy between sense and pure reason, and Kant's distinction between the a priori and a posteriori, all strengthen the plausibility of the picture of the metaphysically secluded self at issue here.

[19] The danger of the very word 'subjective' can be seen as one manifestation of a kind of linguistic danger well examined in the work of Wittgenstein's student Maurice Drury in his *The Danger of Words* (London: Routledge, 1976), repr. in *The Danger of Words and Writings on Wittgenstein* (Brighton: Thoemmes Press, 1996).

[20] For a full explanation of the ventriloquial model, see David Goldblatt, *Art and Ventriloquism* (London: Routledge, 2004); I offer an examination of the relation between the ventriloquial model and pictures of word-meaning in my introduction to this extraordinarily thought-provoking volume.

[21] Although this phrase is used only six times in all of Wittgenstein's published writings, it has generated a great deal of interpretive secondary writings. See *Philosophical Investigations* §§7, 19, and 23, and II, p. 226, for what are perhaps the central sources; see also the entry 'Form of Life (*Lebensform*)' in Glock, *Wittgenstein Dictionary*, for a helpful brief review of the matter, and Stanley Cavell, *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism* (Chicago: University of Chicago Press, 1990).

[22] The rejection of artificial doubt (and the correlated belief that doubt is or can be volitional) constitutes one of the affinities relating Wittgenstein's later work to American pragmatism. See e.g. C. S. Peirce, 'Some Consequences of Four Incapacities', repr. in

As one progresses through Wittgenstein's examinations of the concept of introspection and the picture of first-person seclusion, it becomes increasingly clear that (1) the developed idea of the subject as an interior point upon which introspection can focus, (2) the very conception of introspection itself, and (3) the forms of speech that a distinctively philosophical voice generate along with that voice's fundamental concept of the metaphysically hidden, are all expressions of a single self-concept, and this concept—against the initial appearances when coming to the subject laden with the presuppositions and conceptual preoccupations of traditional philosophy—not only is not given in experience but is a self-concept that is in truth incompatible with what is so given. And it is precisely the argumentative strategy of moving ever back and forth between the two opposed poles, the philosophical versus the ordinary voice, that gradually erodes, and perhaps ultimately breaks down, the false sense of givenness, the illusory sense of the experientially given obviousness of the metaphysical predicament of self-isolation. Among writers in the Wittgensteinian tradition it is perhaps Stanley Cavell who has given the most sustained philosophical attention to the complex inter-relations between the competing pulls of the metaphysical versus the ordinary voice on the one hand and the sense of metaphysical isolation on the other.

3. CAVELL AND THE STAGE OF SPEECH

Given these complexly conjoined concerns, it is not surprising that we thus find in the writings of Cavell not only a rich awareness of the significance of Wittgenstein's philosophy for questions of self-understanding, but also an Austinian sensitivity to the multiform distinctions between the utterances of the philosophical and of the ordinary voice. And it is, again, precisely in these writings, ranging across a number of years and volumes, that we also find a deeply sustained investigation into the logic of the very thinking that would—if unanalyzed, if unchecked—lead us, step by step, into the belief in the experientially given obviousness of the self's most fundamental predicament of metaphysical isolation as just described. With our central concern of autobiographical self-description kept in focus, let us

J. Stuhr (ed.), *Classical American Philosophy* (New York: Oxford University Press, 1987), 32–3.

glance at a few of the episodes of this sustained investigation, drawn from various stages along that investigation's way.

In his early essay 'Knowing and Acknowledging',[23] Cavell situates what is, for him, the problem of privacy and private experience. Of the much-discussed issue of the knowing of the pain of another (the Wittgensteinian example that has grown into the modern discussion of the classic problem of other minds), Cavell insists, against strong protestations, that the very phrase 'being unable to feel another's sensations' is inherently confused, is in urgent need of conceptual elucidation and clarification:[24] 'But there *is* someone who knows, there is a position which is totally different from mine in the matter of knowing whether he is in pain.' Insisting here on the fundamental metaphysical asymmetry between the third- and first-person cases, he adds, 'different not only in being better (as if certain factors in my position were increased in accuracy or range) but in being decisive, making the best position I can be in seem secondhand', and then, underscoring the metaphysical asymmetry, finishes the thought: 'namely, *his* position'.[25] And Cavell at this early stage finds the claim of authority of the other's mind—the distinctively privileged position of the feeler of pain—itself so deeply plausible from one's own first-person perspective that he writes: 'phenomenologically, as a datum, it seems to me undeniable'. After this claim—or rather, after this philosophical picture of the self in its most fundamental metaphysical position as it is given in experience, he adds, 'I think everyone recognizes the experience which goes with it, that it is some terrible or fortunate fact, at once contingent and necessary, that *I* am not in that position,' and, emphasizing that this thought propels other-minds skepticism, further adds that 'the skeptic merely comes to concentrate upon it' (p. 259).

It is notable that Cavell, again at this relatively early point, did not find John Cook's analysis (itself a model of ordinary-language philosophical method as initially inspired by Wittgenstein's philosophy) sufficient, or sufficiently conceptually satisfying for him to (as Wittgenstein puts it) stop doing philosophy on this topic, that is, to achieve a condition of

[23] In Cavell, *Must We Mean What We Say?* (Cambridge: Cambridge University Press, 1969).

[24] Ibid. 259. Cavell at this juncture is writing in response to John Cook's deservedly classic essay 'Wittgenstein on Privacy', *Philosophical Review*, 74 (1965), 281–314, repr. in G. Pitcher (ed.), *Wittgenstein: The Philosophical Investigations* (New York: Doubleday, 1966).

[25] Cavell, 'Knowing and Acknowledging', 259; this and following brief quotations.

conceptual equipoise. Against Cook's compelling linguistic analysis, in which he argues that the difference between the first- and third-person utterances regarding sensations like pain cannot be intelligibly, or if minimally intelligibly then only wholly misleadingly, captured in the language of *circumstance* (as Cook says, like the circumstance of 'being unable to see my neighbor's crocuses'), and that the difference lies in the language-game (and hence that identifying the circumstance as 'being unable to feel another's sensations' is pernicious and linguistically misled confusion), Cavell persisted. 'Why is "being unable to feel another's sensation" not a circumstance?' While acknowledging that the force of Cook's point derives from the notion that first-third-person asymmetry is not a circumstance because it cannot be imagined within the bounds of coherence to be other than it is, that the asymmetry does 'not describe an *inability* of ours', Cavell still asks: 'But why can't a general fact of nature be thought of, accurately, as a circumstance, a permanent circumstance?' And then giving a somewhat startling sense to his notion of a permanent circumstance, indeed a sense that I think transgresses the bounds of Austinian ordinary language,[26] adds, 'The circumstance is, I feel like saying'—this latter phrase does distance Cavell from the claim and betokens an awareness of the philosophical voice's impulsion to speak, an awareness of the possibly 'autoventriloquistic' aspect of this utterance, as discussed above—'*him*'.[27] One sees here that, again, Cavell is not stopped from doing philosophy on this fundamental topic at the level of language; despite the linguistic problems, he finds the more important residue at the level of experience.[28] He writes, 'The problem . . . may be that the formulation "inability to feel" tries but

[26] For the reason that—to encapsulate a lengthier linguistic investigation into the range of uses of the word 'circumstance'—a *person* is not, as Cook rightly sees, a *circumstance*; the person may be *in* a circumstance, and probably only temporarily so. The ability to use the word intelligibly in description of the person's condition does not outlive the context within which we say this of the person, after which we of course speak of the circumstance (however we might describe it, which itself is internally related to the intelligibility of the particular use of the word 'circumstance') in the past tense.

[27] Cavell, 'Knowing and Acknowledging', 260.

[28] My purpose here is to capture Cavell's position at this point in his fluid thinking; I in fact find the distinction between the level of language and the level of experience problematic: the problem is that the distinction is both too general and too quick. Following Austin, we could find cases in which we speak of the level of language apart from, or in contradistinction to, the level of experience (and vice versa), but I would not, on Austinian–Cavellian grounds, expect this distinction as made in a particular case to prove generalizable; an investigation of our various uses of the term 'level' would, I expect, show that the contextually circumscribed meanings are not in truth transferable out of context.

fails to capture my experience of separation from others,' adding a still stronger doubt concerning what he here finds to be the limited reach of linguistic analysis with the words: 'This does not make it inherently confused, but, one might say, much too weak—as though words are in themselves too weak to record this fact' (p. 260).

Given the sizable and conceptually rich body of work to follow on countless aspects of this topic, one might speculate that Cavell here defended the skeptic (against whom Cook's Wittgensteinian linguistic analysis was arguing) in this battle in order to go to war with him later. Cavell sees—or finds given in experience as an undeniable phenomenological datum—what he calls 'the truth to which he [the skeptic] is responding', despite, or beneath (as we might say, having introduced talk of the differing levels of language and experience), the very real difficulty of coherently articulating this metaphysical asymmetry. Cavell's position at this stage is thus expressed in this remarkable way:

I take the philosophical problem of privacy, therefore, not to be one of finding (or denying) a 'sense' of 'same' in which two persons can (or cannot) have the same experience, but one of learning why it is that something which from one point of view looks like a common occurrence (that we frequently have the same experience—say looking together at a view of the mountains, or diving into the same cold lake, or hearing a car horn struck; and that we frequently do not have the same experiences—say at a meeting, or learning the results of an election, or hearing your child cry) from another point of view looks impossible, almost inexpressible (that I have your experience, that I *be* you). What is it I cannot do? Since I have suggested that this question is a real one (i.e., that the sense of "cannot" here is real), and since nevertheless I have suggested that the question has no answer (or the ground that the words 'cannot have his feeling' are 'too weak' for the experience they wish to convey), I would need, in accounting for these facts, to provide a characterization of this sense of incapacity *and* provide the reason for our insistence upon putting it into words. I find that, at the start of this experience, I do not want to give voice to it (or do not see what voice to give it) but only to point (to others, or rather to the fact, of the being, of others) and to gesture towards my self. Only what is there to point to or gesture towards, since everything I know you know? It shows; everything in our world shows it. But I'm filled with this feeling—of our separateness, let us say—and I want you to have it too. So I give voice to it. And then my powerlessness presents itself as ignorance—a metaphysical finitude as an intellectual lack. (Reverse Faust, I take the bargain of supernatural ignorance.)[29]

---

[29] Cavell, 'Knowing and Acknowledging', 262–3.

Within our language-games, within the bounds of sense, we know we can have the same experiences: we say so. Yet, although I can walk a mile in your shoes and vice versa, I am not you, nor will I be, and you are not me, nor will you be. A philosophically sensitized Austinian ear will not want to attempt to give voice to the metaphysical truth thought to lurk beneath this empirical fact, yet one feels an impulsion to do just that—a philosophical impulsion which manifests itself in pointing first at the other, then to the self. Then the futility dawns on us, and we are left with the (seemingly pre-verbal) feeling of human separateness and the very human desire for company—for accompanied solitude, for being alone together—that then refuels the drive of the voice. And thus, with an awareness of the limits of philosophical language, we nevertheless speak—but in ways that seem either to fall far short of the true nature of this metaphysical human separateness or to transgress the boundaries of the coherently sayable.

This description of the self's state of affairs, I would suggest, defines Cavell's position toward the beginning of his investigation; it acknowledges the human need to voice the circumstances (if we can use that word) in which the self finds itself; it simultaneously acknowledges the imperatives of Wittgensteinian and Austinian linguistic sensitivity, or indeed the claims of ordinary-language philosophy; and—I want to insist—it begins to show why literature of the self and of the self's metaphysical predicament is necessary. It shows what cannot be said.[30]

---

[30] The saying–showing distinction derives from Wittgenstein's early work in the *Tractatus Logico-Philosophicus*; see esp. the Preface and 6.41–6.522, in which the relative insignificance of the sayable is emphasized. In the Preface he writes: 'Thus the aim of this book is to draw a limit to thought, or rather—not to thought, but to the expression of thoughts: for in order to be able to draw a limit to thought, we should have to find both sides of the limit thinkable (i.e. we should have to be able to think what cannot be thought)' (p. 3). In a widely influential passage in a letter to von Ficker ('Letters to Ludwig von Ficker', ed. Allan Janik, trans. B. Gillette, in C. G. Luckhardt (ed.), *Wittgenstein: Sources and Perspectives* (Ithaca, NY: Cornell University Press, 1979), 94–5) he writes that 'my work consists of two parts: of the one which is here, and of everything which I have *not* written. And precisely this second part is the important one. For the Ethical is delimited from within, as it were, by my book . . .', and goes on to refer to that which he has defined by remaining silent about it, warning von Ficker that, although the book may 'have much to say which you want to say yourself', he (von Ficker) may well—owing to the Tractarian silence—not 'notice that it is said in it'. Wittgenstein's conception of the very nature of philosophy is changing because of this issue. Because the logical form of propositions cannot be stated (contra his Russellian background), philosophical work must instead yield *clarification*, signifying 'what cannot be said, by presenting clearly what can be said' (*Tractatus* 4.115). And clarification—perspicuous presentation—is the aim of the later philosophy, where clarification is achieved not

There is, it will be obvious, far more to say about this. But the present task is to consider at least a few more episodes in Cavell's work as they pertain to the topic at hand.

In the course of providing context for his suggestion that religious life and expression in language might profitably be construed as a Wittgensteinian form of life[31] in his essay 'Kierkegaard's *On Authority and Revelation*',[32] Cavell quotes Wittgenstein's remark from *Philosophical Investigations* that 'One human being can be a complete enigma to another.' And he continues the quotation, in which Wittgenstein shows where we can intelligibly speak, in an ordinary voice, of matters of an incapacity or inability to understand a human being: 'We learn this when we come into a strange country with entirely strange traditions; and what is more, even given a mastery of the country's language. We do not *understand* the people. (And not because of not knowing what they are saying to themselves.) We cannot find our feet with them.' In this example, it is clear that in one sense we understand their words, yet in another, less shallow sense, we do not understand them at all.[33] Understanding is not reducible to one unitary phenomenon, and it is

---

through, but in the *absence* of, substantive (traditional) philosophical assertions. For a fine brief essay on the saying–showing distinction, see Glock, *Wittgenstein Dictionary*, 330–6. This distinction has been applied to aesthetics in various ways (suggesting, broadly, that the arts show distinctive varieties of meaning that cannot be said). See Susanne Langer, *Feeling and Form* (London: Routledge & Kegan Paul, 1953), and *Philosophy in a New Key* (Cambridge, Mass.: Harvard University Press, 1978), esp. 79, where she makes the reliance on Wittgenstein's early philosophical distinction explicit, and B. R. Tilghman's thoughtful *Wittgenstein, Ethics and Aesthetics: The View from Eternity* (London: Macmillan, 1991). Tilghman closes his searching study with a Tractarian sentence: 'Art does indeed, Wittgenstein would say, provide an experience not to be obtained by any other kind of activity: it shows the meaning of life' (p. 178). I offer an analysis of Langer's position in *Art as Language: Wittgenstein, Meaning and Aesthetic Theory* (Ithaca, NY: Cornell University Press, 1995), ch. 1.

[31] I consider the meaning of Wittgenstein's phrase 'form of life', and its significance for an understanding of the arts, in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), ch. 2.

[32] In *Must We Mean What We Say?*; these and following passages, pp. 172–3.

[33] In relation to this contrast between the shallower and deeper understandings of words, see Cavell's helpful remarks on (what I take to be) a parallel contrast in understanding the meaning of Wittgenstein's phrase 'form of life', in *This New Yet Unapproachable America: Lectures after Emerson after Wittgenstein* (Albuquerque, N. Mex.: Living Batch Press, 1989), 41: 'A conventionalized sense of form of life will support a conventionalized, or contractual, sense of agreement. But there is another sense of form of life that contests this. Call the former the ethnological sense, or horizontal sense. Contesting that there is the biological or vertical sense.' Those who misconstrue Wittgenstein as a sociological relativist seem to emphasize greatly the horizontal sense of his foundational phrase while de-emphasizing—or missing—the vertical.

certainly not reducible to what we would call, perhaps in the context of learning a foreign language, knowing the meanings of each of the words in their sentences; Wittgenstein's example shows this. Similarly, misunderstanding is not reducible to a unitary essence, or one isolable mental experience. Cavell, however, does not at this point draw out the content of Wittgenstein's words; he instead juxtaposes a passage from Kierkegaard, which begins with the unforgettable line 'Most men live in relation to their own self as if they were constantly out, never at home . . .'. The significance—or one line of significance—is clear.

Full engagement with the problem, with the nature of (as well as the articulation of) the metaphysical predicament or condition of the self that runs throughout Cavell's work, demands the recognition of a deep problem doubled: the asymmetry between the first and third person can be mirrored in an internalized version of this problem, such that the self does not understand, does not know, itself. This, one might say, is the internal psychological doubling of an external social problem, or the single-mind version of the other-minds problem, or to express it still another way, perhaps the solipsistic turning of skepticism on itself. And this, as in Wittgenstein's work, greatly disturbs the philosophical picture of the mind's introspective access to itself as the transparently accessible, epistemically privileged, circumstance of self-knowledge, or self-understanding. Thus Cavell writes, following both Kierkegaard and Wittgenstein: 'One may want to say: A human being can be a complete enigma to himself; he cannot find his feet with himself.' And emphasizing the difference in understanding Wittgenstein's example shows, he adds, that 'he understands his words, but he is foreign to his life'.[34] These passages of Cavell's not only continue the unearthing and removal of the misleading picture of the given nature of introspective self-knowledge as parsed by Wittgenstein. They also, in showing the complexity of the problem of other minds, of other-understanding, through their implications reveal part of the human complexity of any biographical—and, as we now also begin to see, autobiographical—undertaking.

Given its foundational nature, it is not surprising to see resonances of this doubled problem throughout Cavell's work. In 'Moral Perfectionism', in the context of writing on Emerson and Thoreau, and particularly in relation to Emerson's notion of the 'unattained but attainable self'[35]—itself an internal doubling—Cavell discusses Thoreau's

---

[34] Cavell, 'Kierkegaard's *On Authority and Revelation*', 173.
[35] Cavell, *Conditions Handsome and Unhandsome*, 8.

remark 'with thinking we may be beside ourselves in a sane sense'. The self, beside itself, is *precisely* the internalization of the structure of the problem of other minds. Cavell observes that Thoreau's remark characterizes thinking as 'a kind of ecstasy', but I believe one might also see in it the characterization of thinking as a kind of autobiographical reckoning, a kind of perpetual self-estimation. Indeed, elsewhere (in *The Claim of Reason*[36]) Cavell offers his own characterization, if not of thinking per se, then of philosophical work. Following Wittgenstein's deservedly much-discussed motto 'To imagine a language means to imagine a form of life',[37] Cavell writes:

In philosophizing, I have to bring my own language and life into imagination. What I require is a convening of my culture's criteria, in order to confront them and my words and life as I pursue them and as I may imagine them; and at the same time to confront my words and life as I pursue them with the life my culture's words may imagine for me: to confront the culture with itself, along the lines in which it meets in me. This seems to me a task that warrants the name of philosophy.[38]

The culture's—perhaps in Wittgenstein's sense, the form of life's— criteria are brought into creative self-confrontation, but in a way that is ineluctably autobiographical as well. Thinking, as Wittgenstein has shown us in his reflections on private language, in his remarks on introspection and consciousness, and in his remarks on thinking itself,[39] is inescapably social, and yet, at the same time, conducted within the first person in a way that would impel the voice of private experience, the voice of human separateness. For Cavell, it seems that philosophy itself is one kind of reenactment of these doubled problems: the self against the other (the culture's criteria); the impulsion to give voice to metaphysical autonomy against the criteria of Austinian ordinary language; and the self against, indeed, itself. Philosophical thinking is, for Cavell, in one distinct aspect, autobiographical; and although it may not follow necessarily, nevertheless it would not be surprising to find plausibility in the claim that autobiography is, in one distinct aspect, philosophical in turn. And as to Cavell's phrase 'In philosophizing, I

---

[36] Cavell, *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy* (Oxford: Oxford University Press, 1979), 45.

[37] The passage comes from *Philosophical Investigations* §19.

[38] Cavell, *The Claim of Reason*, 45.

[39] Each of these will be the focus of a chapter or section in the following (the private-language considerations in Ch. 4, Sect. 3; introspection in Ch. 3, Sect. 3; and thinking in Ch. 4, Sects 1 and 2).

have to bring my own language and life into imagination' in particular, one could hardly find a more succinct description of much literature and perhaps all autobiography; in Cavell's voice, although he may not state this explicitly, speaking for and of philosophy can be tantamount to speaking for and of (varieties of) literature.

The position of the self in self-reflection, the idea of the self inwardly doubled, the suggestion of Thoreau—in which Cavell rightly sees so much—that with thinking we may be beside ourselves, are intertwined ideas that can illuminate and deepen. They can also, if taken in other ways, severely mislead us in our efforts to understand more fully the nature of the self in its literary depictions, in its artistic self-presentations. It is not difficult to see that these ideas could strengthen the Cartesian or metaphysically dualistic misconception of the self against (or beneath) which Wittgenstein has argued, and in doing so strengthen the misconception of the self as it functions in the creation of an autobiography or any other self-revelatory literary undertaking. In short, they can, if we are not both cautious and patient in ways that Wittgenstein, Austin, and Cavell all encourage, powerfully work against the achievement of what Wittgenstein would call a perspicuous[40] representation of the facts, our practices, concerning the nature of the self. Here I want to move to still another episode in Cavell's engagement with the ever-evolving threat of other-minds skepticism, but in this case I would like more to apply Cavell's observations to the just-mentioned danger of misconstrual rather than to describe the encounters.

In the essay 'The Avoidance of Love: A Reading of *King Lear*',[41] Cavell at one point investigates the logic of the position of the narrator in a novel, in contrast to the absence of a narrator in drama. This can, I believe, reveal a good deal about our pre-Wittgensteinian intuitions concerning what we may too easily misconstrue as the *ideal*, not only of the nature of the autobiographical self, but more particularly of autobiographical truth.

Cavell emphasizes that 'no character in a play *could* (is, logically, in a position from which to) narrate its events'.[42] Cavell finds three principal

---

[40] See Cavell's discussion of this fundamental Wittgensteinian notion in his 'Epilogue: The *Investigations'* Everyday Aesthetics of Itself', esp. 380–1.

[41] In *Must We Mean What We Say?* This essay, along with other writings in Shakespearian interpretation, is collected in Cavell's *Disowning Knowledge in Six Plays of Shakespeare* (Cambridge: Cambridge University Press, 1987).

[42] Cavell, 'The Avoidance of Love', in *Must We Mean What We Say?*, 335–7. This position, that the character, as actor, is inextricably part of, and not separable from, the

reasons for this, and each I believe holds considerable significance for our understanding of the autobiographical situation. First, no character can or could possess the credibility of a narrator, not because this character's honesty may be in question, but for the far more metaphysically significant fact that the character, as *actor*, 'is part of what is happening; he is fixed in the present'. The *actor* cannot 'insert a break in [the present]', and 'if he narrates, then *that* is what he is doing', that is, narration is now the action being depicted on the stage. Cavell contrasts this—a striking contrast—to that of the narrator, who *cannot* act in the dramatic present, who *cannot* 'make anything happen'—and this is one source of the narrator's privileged epistemic position vis-à-vis the reader's credulity. I want to suggest that this contrast provides an intuition-shaping conceptual model—potentially a very misleading one—for our pre-Wittgensteinian thinking of the autobiographer, or for that matter any self that is engaged in truthful self-revelation. The autobiographical position, we too easily think, is that of the narrator; the past actions of that pre-narrating self (now recollected and reported) are the actions of the 'actor'. And this divide between the acting and narrating self naturally nourishes Cartesian conceptions of self-knowledge and introspective access: Only the autobiographical narrator has privileged access to the intentions and motivations of the actor. Truthful self-reportage would be, for a narrator on this model (assuming an unimpaired memory) transparently easy, and the ideal of first-person truthfulness would simply be the active and full (both of which would be, on this model, unproblematic) disclosure—here a distinctively metaphysical disclosure—of the mind and action of the pre-narrating past actor. However, the matter of the logic of the autobiographical situation, along with the attendant issue of first-person truth, is vastly more complex, as Wittgenstein's and Cavell's writings show. But this model can easily lead us to stop far short of such post-Wittgensteinian complexities. And, indeed, an autobiographer, like any first-person sensation-reporter (a rather unordinary phrase itself), *is* engaged in the *action* of self-narration; there is no 'perch' above

proceedings, and thus metaphysically cannot rise above the situation upon which we might ask or expect him to pass Olympian judgment, is comparable to Nietzsche's pithy observation that the value of life cannot be estimated by any living human, precisely 'because he is a party to the dispute, indeed its object, and not the judge of it' (*Twilight of the Idols*, trans. R. S. Hollingdale (Harmondsworth: Penguin, 1968), 30). (Nietzsche adds that the judgment cannot be provided by any dead person either—'for another reason'.)

life's analogy to the dramatic stage upon which to sit (and from there, as Cavell has rightly observed, we could *do* nothing anyway). It is perhaps worth noting that this runs perfectly parallel to the claim that the ordinary-language philosopher makes to the metaphysician: Intelligible uses of language will be, for reasons Wittgenstein and Austin show at length, (verbal) acts *in contexts*, i.e. on, not above, the stage of speech.

Second, Cavell, taking note of the two narrations that do occur in *King Lear*, observes that those narrations *are* actions, and that they take place 'within the same continuity of causation and freedom and responsibility as every other act of the play'. This, as I am suggesting, is the truer position of the autobiographer. The narration is an action within the ongoing continuity of causation, freedom, and responsibility of the autobiographer—and thus it houses the complexities of meaning and of interpretation that are resident in any other verbal or written first-person report; the model of the omniscient narrator is a false ideal. It is notable that Cavell points out that in the play 'Edgar's choice to narrate then and there is as significant as the content of his narration,' and this is precisely what should be said from the vantage point of ordinary-language philosophy, for as Cavell puts it: 'Philosophy which proceeds from ordinary language is proceeding from the fact *that* a thing is said; that it is (or can be) said (in certain circumstances) is as significant as *what* it says; its being said then and there is as determinative of what it says as the meanings of its individual words are.'[43] It may be true that this way of expressing this point about ordinary-language philosophy itself gives rise to problems of ordinary-language philosophy, i.e. that it is a *fact* that a thing is said is dependent on the details of context of utterance; what actually constitutes a 'circumstance' of utterance will be similarly context-dependent; the very question of the *significance* of a phrase will arise, or not, in context; the issue of *determinants* of meaning will be occasioned by the use of 'determinants'; the very notion of the meaning of individual words is—as Cavell is as aware as anyone—philosophically troublesome in the extreme (not only in that it invites meaning-atomism rather than meaning-holism—but then this distinction too would not fare well under the scrutiny of ordinary-language analysis); and the notion of measuring, and then finding equal in weight, the meaning-determinative force of the 'fact' of saying along with the atomistic word-meanings is far from unproblematic.

[43] Cavell, 'The Avoidance of Love', 336.

But here again one experiences an impulse to give voice to a thought in the face of, or against the verbal proscriptions of, challenges presented by the ordinary-language approach. And in this case, the thought one feels impelled to try to express works in fact in favor of, in support of the value of, ordinary-language analysis. What one here wants expressed is a general point concerning (1) the contribution (as we may feel forced to call it generally) (2) that the very fact (if we can give ourselves license to call it that) (3) that the narrative is uttered at all makes to the meaning (as we dangerously call it) of (4) that utterance (as we generally categorize the self-revelatory narration). If this, incidentally, seems overly hampered by qualifications and doubts, I will only say that, by the standards of post-Austinian ordinary-language philosophy, it is reckless; one can use a more fine-tooth comb.[44] But the fundamental point I want to make, or try to make, is in truth a comparison: that the autobiographer, as a special case of the first-person narrator, is inevitably *in* the position of Edgar; that the fact that the autobiographer is narrating is part of, and not above, beyond, or outside the life—the causal and contextual continuum—of that writer or teller.[45] And this means, as Cavell points out next in his discussion of the logic of the fictional narrator, most significantly, that 'a "first-person narrative" is not a narrative'. The position of the first-person narrator indeed *cannot* be, or become equivalent to, the position of the narrator in fiction: The fictional narrator does not enter into the causal and contextual continuum, and not only does he or she not, he or she metaphysically

---

[44] See, for example, among the highest achievements of ordinary-language philosophical work, both O. K. Bouwsma, *Philosophical Essays* (Lincoln: University of Nebraska Press, 1965), and (here also for the most sustained writing in the field) Frank B. Ebersole, *Things We Know: Fourteen Essays in Problems of Knowledge* (Eugene: University of Oregon Press, 1967), *Meaning and Saying: Essays in the Philosophy of Language* (Washington, DC: University Press of America, 1979), and *Language and Perception: Essays in the Philosophy of Language* (Washington, DC: University Press of America, 1979). The latter volumes in particular display what one might call a vision of the significance of context for meaning that is impossible to articulate generally (this is itself signally instructive), but that is of the greatest therapeutic (in Wittgenstein's sense of that term) value in freeing us from deeply misleading (conceptual) pictures (again in Wittgenstein's sense of that term). There is an interesting sense in which work in this vein—and at this level—shows what it cannot say (see n. 28 above). One might usefully keep Wittgenstein's fundamental remark (from *Philosophical Investigations* §122) in mind while exploring this kind of work: 'A main source of our failure to understand is that we do not *command a clear view* of the use of our words.—Our grammar is lacking in this sort of perspicuity. A perspicuous representation produces just that understanding which consists in "seeing connexions". Hence the importance of finding and inventing *intermediate cases*.'

[45] I return to this particular point in much greater detail in Ch. 5 below.

cannot, precisely because he or she does not have a causal foothold in that world, thus precluding the very capacity to act.

The contrast between the first-person and the fictional narrator is in truth even more striking, precisely in that the fictional (or perhaps third-person) narrator cannot conceal, cannot willfully mislead as a causal intervention in the action, cannot be duplicitous, cannot employ self-protective descriptions of events, because he or she does not possess a self (in the requisite sense of the first-person narrator) to conceal, to protect, to hide behind dissimulation. Cavell expresses the point this way: 'The third-person narrator, being deprived of self-reference, cannot conceal himself; that is to say, he has no self, and therefore nothing, to conceal.'[46] Yet, despite the metaphysical impossibility of converging with that of the (imaginary) fictional narrator, it is nevertheless far too easy to hold up this latter position as the conceptual model for the autobiographical position: a model or picture of the self and its position that shapes our intuition and buttresses, if not directly causes, the philosophical misconstruals of the self Wittgenstein is combating, often by employing the methodological practice of returning words used in the philosophy of the self to their uses in the ordinary voice.

Third, Cavell observes that there is conceptual room for the activity of a reporter who is giving a report simultaneous with the event the report is describing (and it is thus written or said in the present tense), because the reporter is there at the event, while we are not. In a theater, Cavell notes, there is no such position, no such conceptual room, because, quite indisputably, 'We are present at what is happening'. Here once again, although perhaps well beyond what Cavell intended, this observation is significant for our understanding of the self-narrator, and in a way somewhat less evident than the previous two observations. Our already strong inclination toward Cartesian self-misconstrual can be further strengthened by the misapplication of this 'reporter-from-elsewhere' model to the case of the first-person narrator. Directly stated—with this model lodged in the intuitive subterrain—it becomes easy to conceive of the first-person narrator as a reporter of an occurrence from which we are absent—the private, mental occurrence, we think, from which we must always, as a metaphysical necessity, be absent, and to which there can and will forever be only one 'reporter', the mind present at and to its own inward occurrences. But again, this simple dualistic dichotomy separating the self-mind from the other

[46] Cavell, 'The Avoidance of Love', 336.

mind cannot withstand the scrutiny of linguistic analysis: the usages of the ordinary voice on these matters both refuse to acknowledge and are far too nuanced for this simple Cartesian dichotomy, and it is again literature that shows the philosophically most significant facts in contexts of linguistic usage, in human language-games. In truth, we can be present at, or absent from, countless lived experiences that we may well call 'mental', just as we may call some of them 'private' (and others not); the particular way in which we are present or absent, or, in some cases, both, will depend on *particular circumstances* and not on what can be taken to be a grand metaphysical truth of the self's metaphysically enforced solitary confinement, as given as the first brute fact of human experience. But the conceptual undertow dragging us back to this conception, this picture, of the self and its predicament is powerful, and, again, it is only strengthened by the misleading analogy of the narrator in fiction. The autobiographer and, much more broadly, the first-person narrator or teller is a very different kind of creature.

Again, Cavell's sustained investigation of the logic of other-minds skepticism thoroughly acknowledges the force of what I'm calling this conceptual undertow, and in the essay 'Being Odd, Getting Even',[47] Cavell offers memorable examples of the motivating utterances of the metaphysical voice, what he here calls the 'move to the metaphysical'.[48]

This move to the metaphysical is like saying that since it makes sense to suppose that I might lack any or all of my limbs I might lack a body altogether, or that since I never see all of any object and hence may not know that a given object exists I may not know that the external world as such exists. Ordinary language philosophy, most notably in the teaching of Austin and of Wittgenstein, has discredited such a move to the metaphysical, as a way of discrediting the conclusions of skepticism.

Such metaphysical utterances, such skeptical claims, are indeed examples of what Wittgenstein called 'language gone on holiday', the transgressions of the bounds of our language-games. But Cavell is here too, as in his commentary on Cook's Wittgensteinian linguistic analysis, unwilling to give up the seriousness, perhaps the human profundity, of the insight, born of the inward phenomenological feeling of human separateness or isolation, that in part motivates these conceptual–linguistic

---

[47] Cavell, 'Being Odd, Getting Even (Descartes, Emerson, Poe)', in *In Quest of the Ordinary: Lines of Skepticism and Romanticism* (Chicago: University of Chicago Press, 1988).   [48] Ibid. 110–11.

transgressions. Thus Cavell adds next: 'But in my interpretation of Wittgenstein, what is discredited is not the appeal or the threat of skepticism as such, but only skepticism's own pictures of its accomplishments.' The appeal and the threat, consistent both with Wittgenstein's vision of the organic growth and ever-changing evolution of our language (-games) and his insight into the myriad tricks that language plays on its users and the ever-evolving 'bewitchment of our intelligence by means of language', present themselves in ever-new guises. To employ the language of Wittgenstein's therapeutic analogy[49] for ordinary-language philosophical work, to diagnose one case is not to cure all cases—just as it is not necessarily to *cure* the case diagnosed.

In his essay 'Ending the Waiting Game',[50] Cavell offers an encapsulation of the view of the self we have been considering in its relation to autobiographical revelation and, more fundamentally, autobiographical knowledge. Drawing out the comparison of a number of skeptical philosophical views to forms of madness, Cavell, listing such positions as believing the world is only illusion, or doubting the external world, or claiming that our world is composed out of isolated bits of experience, includes the belief that 'each thing and each person is a metaphysical enclosure, and no two ever communicate directly, or so much as perceive one another'.[51] He shortly turns to Wittgenstein's analogy between philosophical work and therapy, noting that 'his late methods (he compared them to therapies) were to bring philosophy peace at last', and then quotes Wittgenstein's famous dictum from the *Remarks on the Foundations of Mathematics*: 'The philosopher is the man who has to cure himself of many sicknesses of the understanding before he can arrive at the notions of the sound human understanding.'[52] And that sound understanding is the possession of a perspicuous grasp of a field of our practices, a clear view not gained from ascending theoretical heights but of what lies before us on the level of praxis. It is the understanding achieved by the ordinary voice,[53] having, as T. S. Eliot has it, 'returned once again to the place it started but

[49] I will return to this in Ch. 7, Sect. 3, below.

[50] Cavell, 'Ending the Waiting Game: A Reading of Beckett's *Endgame*', in *Must We Mean What We Say?*   [51] Ibid. 126; this and the following brief quotations.

[52] *Remarks on the Foundations of Mathematics*, ed. G. H. von Wright, R. Rhees, and G. E. M. Anscombe, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967), 157.

[53] Against the uncomprehending view that a reliance on ordinary language is tantamount to anti-intellectualism, see Cavell's discussion in 'The Ordinary as the Uneventful (A Note on the *Annales* Historians)', in *Themes Out of School: Effects and Causes* (Chicago: University of Chicago Press, 1984); see esp. pp. 192–3.

knowing it for the first time'.[54] Moving between the poles we saw Wittgenstein initially identify, it has returned from—having felt, expressed, grappled with, and come to terms with—the impulsions of the metaphysical voice.

   Within this chapter it has been possible to analyze only a selection of the remarks, observations, and arguments Wittgenstein has made on language and mind as they pertain to the achievement of a clear—indeed philosophically peaceful—understanding of the autobiographical subject, the first-person narrating self (it will be the task of the remaining chapters to work through a good many more). And it has similarly been possible to review only a few sites of investigation of Cavell's larger philosophical undertaking as they pertain to this subject. But this may be sufficient at least to suggest a number of the further and increasingly intricate ways in which misleading conceptual models, or philosophical pictures, as Wittgenstein and Cavell after him used the phrase, can prevent a clear view. As we saw in a foundational or preliminary way in Chapter 1, much of the language of the self—the *ordinary* language of the self, for all its philosophical significance—is found in autobiography, in a different sense in biography, and in still different senses in various literary depictions of the mind and mental activity. Although we will have a fairly close look at parts of Augustine's *Confessions* in this regard in Chapters 4 and 5, it would take a lengthy and separate study to consider fully (although a number of these will come into play briefly along the way) the multifarious ways in which Nabokov's *Speak, Memory*, Mill's *Autobiography*, Stein's *The Autobiography of Alice B. Toklas*, Rousseau's *Confessions*, the autobiography of Bertrand Russell, autobiographical works by de Beauvoir, Nietzsche, Sartre, Voltaire, Kierkegaard, Vico, Henry James, Thoreau, and countless other pieces of life-writing—including Cavell's *A Pitch of Philosophy: Autobiographical Exercises*[55]—are all significant for the achievement of what Wittgenstein called a perspicuous presentation, a 'sound understanding'. But it is clear that the 'deeply mysterious "I" ' *has* been investigated at length in literary contexts with great philosophical resonance, and if

---

   [54] T. S. Eliot, 'Little Gidding', in *Four Quartets* (San Diego: Harcourt Brace Jovanovich, 1943), 59 (Part V, lines 240–3).
   [55] (Cambridge, Mass.: Harvard University Press, 1994.) Given the large-scale trajectory of Cavell's intellectual project, the move into autobiography—the long-term occupancy of the autobiographical situation and its exploration from the inside (without reintroducing Cartesian conceptions of selfhood)—seems not only philosophically motivated but in fact necessitated.

Wittgenstein left something of a Tractarian silence directly on the nature of the autobiographical self, he certainly provided the conceptual tools with which to clear our view philosophically so that we *can* see the multiform literary and linguistic practices that lie before us. Perspicacity presumes conceptual clarity, and given the analogies, the models, the pictures—the undertow—we have considered, it is far too easy for the self—devoid of the stabilizing and grounding influences of language *as used* and literary investigations—to view its image in a reflecting but darkening glass.

If many of these considerations weigh against a Cartesian picture of selfhood, are we thus being moved closer to its polemical antithesis, behaviorism? If not, we need to clarify *why* not, and once that is done, what significance do these reflections hold for our understanding of self-revelatory expressive language? And what significance would that, in turn, hold for our understanding of introspection? These are the issues in the next chapter.

3

The Self, Speaking

IF we look back over all that Wittgenstein has written to expose, analyze, and free us from the conceptual confusions and distortions in our understanding generated by the Cartesian or dualistic theory of the self that we have to this point considered, it seems possible to grasp, without excessive difficulty, exactly what Wittgenstein is opposing in those sectors of his investigations. And thereby we are offered an important part of what might accurately be called a *negative* definition of the self, i.e. we are given an understanding of what the self is *not* along with, perhaps more importantly, why we might have thought the self corresponded to those inaccurate conceptions in the first place.[1] But that negative achievement may be accompanied by a sense of loss. One way of voicing that sense is to ask the seemingly simple question: Are we, at the end of Wittgenstein's criticisms of Cartesian dualism, unavoidably reduced to one or another kind of behaviorism? Have Wittgenstein's reflections left us with any such eviscerated, or indeed vacated, view of the self?

I will try to investigate Wittgenstein's critiques of the reductive theory of the self, known as behaviorism, as it has been developed in response to—and thus shaped by the preconceptions of—the dualistic Cartesian model of the self and suggest that out of Wittgenstein's critiques there arise both a clearer conception of first-person utterances or descriptions and a clearer conception of introspection—indeed a nontheorized conception of introspection. And those, taken together, can naturally offer a clarified conception of autobiographical or first-person knowledge and the nature of its expression in self-referential

---

[1] I use the concept of the self, as a philosophical problem, in a broad sense to include the problem of the clarification of first-person psychological concepts in relation to the construals and misconstruals of 'the inner'; the more narrow problem of the self, i.e. the nature of the referent of the first-person pronoun, is I believe only one part of the question of selfhood (and in fact, for reasons which will emerge, is fully comprehensible only within a broader context).

language. This will involve: (1) an examination of the 'threat' of behavioristic reductionism that would, if unleashed, eliminate the very concept of the cognitively and affectively engaged subject (and thus an examination of the possibility that Wittgenstein may be a disguised behaviorist); (2) an examination of the supplanting of the theoretical model of object-and-designation as it has been misleadingly applied to inner-state descriptions with the very different notion of what have been abbreviatedly called 'avowals'; and (3) an examination of Wittgenstein's rethinking of the notion of introspection in accordance with the first two issues. Bringing the three issues together helps to show just how Wittgenstein's work on those topics undercuts, and frees us from, the polemical opposition (i.e. the Cartesian versus the behavioristic conceptions of selfhood) that traditional theory has produced and sustained, in both explicit and implicit forms. And I want to suggest that Wittgenstein's critique of that theoretical opposition, a critique which may itself defy theoretical encapsulation or succinct expression, can powerfully reorient our understanding and appreciation of the great value of the arts of self-representation (e.g. autobiography, self-portraiture, representations of human subjectivity in theater and film, etc.) as nonreductive philosophically significant investigations into selfhood and self-reflection. But the immediate question is: Have Wittgenstein's reflections, and particularly his critique of the Cartesian model of inner selfhood, left us conceptually bereft?

1. A BEHAVIORIST IN DISGUISE?

The answer we can quickly extrapolate to the question, fortunately, is an emphatic *No!*[2] But to see why the answer is the result of argument

---

[2] It is true that there are a few places in Wittgenstein's writings where he veers especially close to behaviorism: in, for example, *Philosophical Remarks*, ed. R. Rhees, trans. R. Hargreaves and R. White (Oxford: Blackwell, 1975), ch. vi, and *The Blue and Brown Books* (Oxford: Blackwell, 1958), 51–2. But in many other places the clear priority of the anti-behavioristic, noninferential description of a person's emotional state over the description of that person's behavioral movements as evidence for the emotional state is made clear: see *Zettel*, ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967), §225; *Remarks on the Philosophy of Psychology*, 2 vols (Oxford: Blackwell, 1980), i, ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe, §§1066–8. For a fine brief essay on the topic, with many helpful references, see Hans-Johann Glock, *A Wittgenstein Dictionary* (Oxford: Blackwell, 1996), 'Behaviour and Behaviourism'.

rather than dogmatic assertion, we need to see how the repudiation of behaviorism fits with Wittgenstein's conception of language, particularly in his above-mentioned escape from the fundamental and deeply mistaken notion that meaning is invariably a function of reference, or that all words really function as names, and that inner-state reports function as descriptions. A full examination of that notion in linguistic philosophy directly would take us far afield, but we can here consider those aspects of the issue most pertinent to self-description.

The relation of the issue to Wittgenstein's discussions of linguistic meaning—and the way in which language itself leads us to theorize parallel relations between (1) a name and its referent, and (2) an emotion-term and its inner referent—appears throughout Wittgenstein's later philosophy, and directly relates to the much-discussed issues of private language and of criteria. I offer a wider view of Wittgenstein's philosophy of language elsewhere,[3] but we can at least here see something of the relation of that work to his not systematically formulated argument against behaviorism. Indeed, that Wittgenstein's position on the self, contra-behaviorism, is not systematically presented in the manner of traditional philosophical disputation is precisely what has led to the widespread misunderstanding of his philosophy of psychology within analytic philosophy and beyond, i.e. the belief that he is a behaviorist in disguise (and moreover one who explicitly states that he is not). As we shall see, if his position *were* systematically formulated against its opposition, i.e. if presented as an articulated general *theory* of the self, it would not, for that very reason, constitute a late Wittgensteinian position. In that case it would only preclude the insights into the nature of the *questions* of the self that his indirectly presented position—which a reader must actively assemble as a conceptual mosaic—offers.

Recognizing the plausibility of the interpretation of his view as essentially behavioristic, Wittgenstein gives voice to it in the interlocutor's accusation in *Philosophical Investigations* §307: 'Are you not really a behaviourist in disguise? Aren't you at bottom really saying that everything except human behaviour is a fiction?' If he has removed the concept of introspection as the source of self-knowledge (although, as we shall shortly see, he in truth removes only a mistaken and metaphysically misled picture of introspection), and exposed the nonsensical nature

---

[3] In *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994); and *Art as Language: Wittgenstein, Meaning, and Aesthetic Theory* (Ithaca, NY: Cornell University Press, 1995).

of the idea of the metaphysically hidden, it could certainly seem so. And if he also claims that behavior is at the bottom of our knowing (e.g. that a person is in pain), then is he not, perhaps unwittingly, implicitly siding with the view that the mental is in truth reducible to the behavioral? But behavioral reductionism can be articulated in a number of ways.

The least subtle variety claims that there *is* only behavior; it denies that the mental exists. It is claimed that mental events are only epiphenomenal illusions, and should be excluded from any clearheaded account of—well—human behavior. First, however, Wittgenstein has suggested repeatedly that the mental and the physical are concepts, or categories of concepts, that reside within different language-games. And it is certainly not clear from any such claim that separation implies elimination; indeed, what is crystal clear is that in Wittgenstein's later philosophy he does not desire to eliminate anything (apart from conceptual confusion) but rather to 'leave everything as it is'. In *Philosophical Investigations* §305, in reply to the interlocutor's claim that we surely cannot deny that an inner process takes place in remembering, Wittgenstein replies: 'What gives the impression that we want to deny anything?' The *impression* of denial is explained, he claims, by his having argued against the picture of the 'inner process' to provide a 'correct idea of the use of the word "to remember." We say that this picture with its ramifications stands in the way of our seeing the use of the word as it is.' And denial is clearly not Wittgenstein's position in §306: 'Why should I deny that there is a mental process?' And once again moving between the philosophical and the ordinary poles, he adds: 'But "There has just taken place in me the mental process of remembering . . ." means nothing more than: "I have just remembered . . .".' But it is *not* the mental process, he goes on to reiterate, that he is denying: it is the metaphysically misleading power of a particular form of expression. He is not, of course, denying that a mental process may take place in someone else's head when that person remembers something. What he is denying is that such a mental process is a necessary criterion for the meaning of 'remembering'.

Second, it is clear that Wittgenstein is not siding with a behavioristic program of any eliminative kind; his position on the difference between pain-behavior accompanied by pain and pain-behavior without pain makes that clear. He says, indeed, 'What greater difference could there be?' (§304); in saying that, i.e. in not only preserving but laying great emphasis on the *difference*, he is implicitly laying emphasis on

the existence of something other than pain-behavior. Yet that too requires some subtlety, for the 'something other' cannot be helpfully characterized as an inner object. That characterization would satisfy the demands of the linguistic misconceptions, i.e. that words mean only through referring, so if they mean some*thing* by 'pain' then there must exist an inward object-referent, or that 'pain' names the inner object. Here it does become difficult to quiet the philosophical voice that senses loss in the face of behavioristic eliminativism. But Wittgenstein said, further in response to the claim concerning his really saying that at bottom everything except human behavior is a fiction, that if he speaks of fiction then 'it is of a *grammatical* fiction' (§307). And that fiction is precisely the philosophical picture of object-and-designation, which is implicit in the exchange in §304. Having laid emphasis on the difference, we find the interlocutor's philosophical voice expressing the loss: 'And yet you again and again reach the conclusion that the sensation itself is a *nothing*.' Wittgenstein's voice replies with vehemence: '—Not at all. It is not a *something*, but not a *nothing* either!'

Although vehement, that seems hardly the most helpful answer. It is in fact disorienting: *how* are we supposed to go on thinking in the face of that double-sided denial? The point, of course, is that we are *not* supposed to go on thinking, in the terms that the grammar of object-and-designation, the terms of the metaphysical picture, has laid down. 'We have only rejected the grammar which tries to force itself on us here' (§304). The paradox—that it is neither something nor nothing—disappears if we make 'a radical break with the idea that language always functions in one way, always serves the same purpose'. As we shall see, one of the multifarious ways in which language functions, and particularly in the case of pain-reports ('reports' is itself misleading, since the ordinary voice, as we discussed it in Chapter 2, would ask to whom, for what purpose, and in what context do we speak of *reporting*?), is as an *expression*; those can be properly understood as avowals rather than descriptions, and that change in view supports a change in our view of the inner—but this is to jump ahead. For the present we can see that disorientation—if it is the result of paradoxically running up against the limits of a misleading picture—is philosophically therapeutic: it is in the context of this particular discussion that Wittgenstein makes the now famous claim that the aim of his philosophy is to 'shew the fly the way out of the fly-bottle' (§309).

The fairly crude behaviorist view that would *eliminate* the mental is thus not a view with which Wittgenstein in any sense, explicitly or

otherwise, agreed. There is, however, another articulation of behaviorism that Wittgenstein occasionally seemed close to as well: the slightly more refined view that all propositions containing mental predicates can be reduced to propositions concerning behavioral dispositions. That can be considered from the differing vantage points of first-person and third-person utterances. The first-person case will be discussed shortly—although it can be said presently that Wittgenstein unambiguously states that there is an undeniable difference between the relations one has to one's own speech and to the speech of others: 'My own relation to my words is wholly different from other people's' *(Philosophical Investigations* II. x, p. 192). If true, then any attempt to assimilate the first-person case to the third-person case (as Carnap had suggested[4]), where we analyze first-person emotive or mental propositions into the third-person form and determine their truth-values through self-observation, cannot gain any real plausibility. Thus, Wittgenstein remarks: 'If I listened to the words of my mouth, I might say that someone else was speaking out of my mouth' (p. 192). The point concerns, again, the *strangeness*, the alien nature, of the philosophical voice sounding out the tenets and implications of behaviorism, which stands in striking contrast to the grounded voice. If we 'listened' in the way that behavioristic conception implies, we would be bizarrely self-bifurcated (and the corresponding understanding of autobiographical writing would be similarly bizarre)—but then the behavioristic program modeling the first-person on the third-person cases would fail anyway, since we would then want to know about the separated *listening* self, not the speaking self, thus again marking the difference, the irreducible asymmetry, between the first- and third-person cases that behaviorism would be attempting to eradicate. But that way of speaking *can* prove apt and meaningful:[5] ' "Judging from what I say, *this* is what I believe." Now, it is possible to think out circumstances in which these words

---

[4] The fundamental position here, in essence, was that any first-person statement concerning an emotive state or condition could be reduced to, or analyzed into, statements concerning one's own (first-person) behavior. And with that comes its natural concomitant position: that the observation of the self, by the self, provides the verification for the statement. For Wittgenstein's challenge to that idea, i.e. that it can so much as make sense to verify a first-person emotive statement by observing one's own behavior, see *Philosophical Remarks*, 89–90, and *Zettel* §539. And again, see Glock, 'Behaviour and Behaviourism', 56–7.

[5] The case is thus analogous to contextualizations of the philosophically dangerous word 'hidden'; in ordinary usage, as we saw in connection with the cases considered in Ch. 1, it does not carry dualist-metaphysical freight.

would make sense' (p. 192). But they will not display the meaning
the behaviorist would give them (were it possible for them to have
that meaning). And recall Wittgenstein's query in his discussion of
introspection: 'Do I observe myself, then, and perceive that I am seeing
or conscious?' (*Philosophical Investigations* §417). Without examining
at present the first-person case in detail,[6] it certainly appears that it
logically will not allow assimilation to the third-, and the resistance
betokens the ineliminability and the irreducibility of the mental[7] that
Wittgenstein is not only allowing but in fact preserving throughout his
investigations.

Looking at the third-person case directly, we find three similar anti-
reductive considerations. First, there is the problem of the allegedly
inferential nature of the knowledge we gain of the emotional conditions
of other people on behavioristic terms. If the behavioristic reduction
is to succeed, then the perception or recognition of an expressive
state is actually an inference drawn from particular bodily movements.
However, a cluster of insuperable difficulties for the behaviorist emerge.
In *Philosophical Investigations*, part II, §xi, page 210, we find this remark:
'One might say of someone that he was blind to the *expression* of a
face. Would his eyesight on that account be defective?' The answer to
the question is of course negative, which means that the seeing of the
expression is something *other* than just seeing (what we would, on this
behavioristic model, be led to call) physiognomic evidence. And that is
already implied by the very notion of being blind to an *expression*; one
can see the face fully well, yet see nothing in it, and that severs the link
between the behavior and our knowledge of the emotional state that
is allegedly wholly reducible to the behavior. That there is something

---

[6] Such an examination would require a fairly full consideration of the collection of
remarks on private language (to which I will return in Ch. 4, Sect. 3, below). For the
central writings (from which there stems a huge secondary literature), see *Philosophical
Investigations* §§243–315. For some helpful discussions of first-person issues, see Colin
McGinn, *The Character of Mind*, 2nd edn (Oxford: Oxford University Press, 1996),
'The Self'; G. E. M. Anscombe, 'The First Person', in S. Guttenplan (ed.), *Mind and
Language* (Oxford: Oxford University Press, 1975); and the papers collected in Quassim
Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994).

[7] The nature of the irreducibility has been stated, if in rather different terms, by
Donald Davidson; see *Essays on Actions and Events* (Oxford: Oxford University Press,
1980). In the present connection, see esp. 'Psychology as Philosophy' along with
'Comments and Replies'. Davidson's characterization of this irreducibility is, perhaps
against the initial appearances, incompatible with Wittgenstein's late view of the mind.
See Tim Thornton, *Wittgenstein on Language and Thought: The Philosophy of Content*
(Edinburgh: University of Edinburgh Press, 1998).

unaccountable for in terms of behavior, or physiological movements, is further underscored in the next remark in *Last Writings on the Philosophy of Psychology*, volume i (the writings from which the above remark in *Philosophical Investigations* was taken), where we find an entry with nothing but two quotations from the ordinary voice: ' "he has the eye of a painter," "the ear of a musician" ' (§764).[8] The painter's eye sees something well beyond what the nonpainter sees, yet it is perfectly conceivable to imagine that physiologically, or retinally, they see the same thing. And just so with the musician and auditory experience. The fundamental point is double-pronged: it is anti-reductive—what we see in a face or in a painting, or hear in a composition, is not reducible to what we see or hear in blunt physiological terms—and it is anti-inferential—one can fully well see (or hear) the 'evidence' and yet not reach the conclusion.[9]

Second, and worse for the behaviorist, is the preceding point but reversed: one can in truth have the conclusion without (again what we are on this model misleadingly forced to call) the evidence. We may well, and often do, know that a person is emotionally pained or in or near any other of countless emotional states or conditions *without* being able to specify with any detail or precision whatsoever—and particularly in terms of localized, and particularly facial, movements—*why* it is we know it. And indeed, giving a precise description, one thorough enough to warrant evidentially the inference we have allegedly drawn from the facial and bodily movements, is usually impossible. It is true that a painter in possession of a highly developed painter's eye may be able to approximate that kind of description, but if one even for a moment experiments with such cases it is immediately evident that such finely shaded descriptions, easier drawn than said, are like rationalizations: they come *after* the important fact, the fact of having already recognized the emotional state of the subject.[10] Moreover, consistent with the

[8] *Last Writings on the Philosophy of Psychology*, 2 vols, ed. G. H. von Wright and Heikki Nyman, trans. C. G. Luckhardt and Maximilian A. E. Aue (Oxford: Blackwell, 1982). For extremely helpful further discussions of this branch of Wittgenstein's writings, see Malcolm Budd, *Wittgenstein's Philosophy of Psychology* (London: Routledge, 1987), and Joachim Schulte, *Experience and Expression: Wittgenstein's Philosophy of Psychology* (Oxford: Oxford University Press, 1995).

[9] For a discussion of the point specifically in relation to aesthetic perception, see *Art as Language*, 'The Aesthetics of Indiscernibles'. For further related remarks of Wittgenstein's, see *Last Writings on the Philosophy of Psychology*, vol. i, §§766–7.

[10] See nn. 2 and 5 above for references in and beyond Wittgenstein's writings on this topic.

great majority of philosophical discussions of perception, I spoke just above of *knowing* that a person is emotionally pained etc., but the truth is considerably more nuanced—and this nuanced truth is shown in countless depictions of person-perception in the arts, perhaps especially in film, theater, opera,[11] and literature, but particularized depictions of emotionally comprehending person-perception are also provided by sculpture and painting.[12] The interesting questions, in actual human practice (and reflected in the arts and presented at a distance from the ongoing engagements of ordinary practice and thus amenable to close yet practically detached scrutiny—giving us our practices within a kind of 'laboratory' of observation[13]), do not invariably concern knowing, but rather sensing, suspecting, believing, half-believing, seeing-but-not-wanting-to-see, or any of very many other phenomena of emotionally informed person-perception that range across a vast spectrum.[14]

Philosophy, left to its own devices, has often suggested that there is one unitary and general question of other-knowledge that takes the form 'How can X know Y is in pain?' Philosophy, still with its own devices but aided by the diverse arts of self-representation that themselves constitute a vast and variegated collection of philosophically informative descriptions of instances of person-perception, would no longer find the unitary question sufficient to the task. Philosophy would—and through attending to the relevant arts could—instead assemble a Wittgensteinian overview, a conceptual mosaic of particular cases.[15] Specifically, I do not at this juncture want to put forward the

---

[11] I offer a reading of *Don Giovanni* in just these terms in 'Leporello's Question: *Don Giovanni* as a Tragedy of the Unexamined Life', *Philosophy and Literature*, 29/1 (Apr. 2005), 180–99.

[12] I discuss this point in connection with Rembrandt in Ch. 6, Sect. 2, below.

[13] That is to say, even though our practices are, as Wittgenstein says, what lies before us and it is an undistorted and conceptually clarified understanding of these that will prove therapeutic, there is still very good reason to turn to the arts, and in connection with present issues, to turn to the arts of self-representation. Artistic representations, in addition to giving us our practices at an aesthetic distance (I will return to this below) and thus allowing contemplative observation, can, in showing us a range of experiences beyond our own, extend and enrich the cases we have before us. And of course in the arts our practices are often presented to us with powerful and illuminating commentary upon the depicted practices (e.g. Dostoevsky) or with a descriptive detail that can be rare in the rush of actual events in life (e.g. Henry James).

[14] For one discussion of this epistemological continuum, see my examination of Henry James's tale 'The Tree of Knowledge', in *Meaning and Interpretation*.

[15] But of course the relation between philosophical problems of the self and the arts is not itself reducible to a single, unitary definition either; I will return to a number of the possible relations below (e.g. (1) the arts show what philosophy says, (2)

insupportable claim that art does, or could do, philosophy's job better than philosophy itself does, but rather that some artworks are indeed like Wittgensteinian philosophy in their capacity to deliver—and keep us mindful of—an understanding of human behavior that is not inferential in nature. It is true, it must be said, that there are cases in the theater where actors do deliberately produce bodily movements that would warrant the inference of an emotional state, but such cases are often derisively called 'indicating', which—as a move along a continuum from naturalness to melodrama—are, within behavior, like the return from the philosophical to the ordinary voice: that way of conceptually modeling expressive human behavior can make sense, but in rarefied contexts very unlike the natural and ordinary cases the behaviorist wants to explain—or explain away.[16]

Third, there is a fundamental problem in the attempt to reduce person-perception to body-perception, to see a human being as a body-in-motion. Although not perhaps initially evident, this way of putting the matter in fact argues against itself: If the behaviorist says that we should or do, un- or semi-acknowledgedly, perceive or see persons *as* bodies and on inferential grounds attribute emotions, thoughts, etc., to these bodies, the very term 'as' casts doubt on the behaviorist's position. To see an *x* as a *y* is to logically imply that *x* is not an instance of, or identical with, *y*; seeing that a thing *is* a certain thing is incompatible with seeing it *as* that same thing.[17] We can see the ambiguous line drawing from *Philosophical Investigations*, part II, §xi, *as* a duck or *as* a rabbit; one does not see a duck *as* a duck or a rabbit *as* a rabbit. To make any such claim would further falsify, or philosophically mischaracterize, our natural perceptions; to suggest that a person is perceived first as a body makes a similar mistake. We can indeed try

the arts therapeutically remove confused metaphysical pictures in philosophy, yielding conceptual clarification, (3) art itself becomes a manifest form or medium of philosophical thinking, (4) the arts themselves are metaphysically confused and await the clarifications of Wittgensteinian analysis, and (5) the arts, as case-evidence, confirm or disconfirm philosophical theses).

[16] See *Philosophical Investigations* §313. We will return to the case of acting in the discussion of Goethe in Ch. 6, Sect. 1.

[17] I discuss seeing-as, or aspect-perception, more fully and in relation to literary meaning in *Meaning and Interpretation*, ch. 4. For a lucid account that is ultimately of the causal–inferential kind, see Richard Wollheim, 'Seeing-as, Seeing-in, and Pictorial Representation', in *Art and Its Objects*, 2nd edn (Cambridge: Cambridge University Press, 1980); for an equally lucid critique of this position on Wittgensteinian grounds, see John Hyman, *The Imitation of Nature* (Oxford: Blackwell, 1989).

(and usually fail, fortunately) to see a human being as a body in motion, exhibiting behavioral evidence, but the result is undeniably—and here again the irreducible shows itself—*inhuman*. And, instructively (and illustrating one distinct category of philosophy–art relations), the inhuman character of the distinctive and fortunately unusual mode of person-perception is shown in works of art as well. With this in mind one might review Fernand Léger's film *Ballet mécanique* (where the title itself indicates the oddity), some of Meyerhold's theater, and perhaps the chillingly unsympathetic protagonist in Kosinski's novel *Cockpit* as cases in which the unnatural or alien person-perception at issue is shown, through its artistic depiction, as the exception that proves the rule (i.e. here the arts offer a depiction of one extreme philosophical conception of human perception that is, as we see at a glance, strikingly and sometimes uncannily unlike the norm, the settled mode or attitude of human interaction). Similarly, we might consider the tradition in early, silent film of employing formulaic character postures in order to 'signify' emotional states because of the absence of spoken language. Such postures can indeed be conceived as 'signs' that refer to emotional states rather than manifesting them or expressing them; moreover, they are signs about which spectators have to make inferences. But it is clear, again at a glance, and tellingly so, that such silent films employ a deliberate, calculated signal-system for communicating—not exactly emotional states, but the *idea* of—emotional states, and once again show the very great distance between this filmic communication-system and our ordinary mode, or settled attitudes, of emotionally engaged interaction. It is as though silent film performs the substantial philosophical service of presenting the precise picture of how expressive communication *would* proceed *if* the behavioristic model under consideration were the accurate one. The case is telling because we know immediately: Real life is not like that, any more than language is like, say, Morse code, or normal expressive gestures like pantomime.

In *Philosophical Investigations* §420 the philosophical voice insists: 'But can't I imagine that the people around me are automata, lack consciousness, even though they behave in the same way as usual?—If I imagine it now—alone in my room—I see people with fixed looks (as in a trance) going about their business—the idea is perhaps a little uncanny.' The return to the ordinary shows that such perception is very much the exception: 'But just try to keep hold of this idea in the midst of your ordinary intercourse with others, in the street, say!' Indeed, if we try to force such a way of seeing upon ourselves we will find that words

such as 'all their liveliness is mere automatism' become meaningless, or the result will be a feeling of uncanniness. 'Seeing a living human being as an automaton is analogous to seeing one figure as a limiting case or variant of another; the cross-pieces of a window as a swastika, for example' (§420).[18] To see a human being as an evidence-exhibiting body is not to see what it (and the very word 'it' here is conceptually misbegotten) *is*.[19] Wittgenstein finds a memorable way to capture this fundamental point by appealing to what we do, in real circumstances, in response to a person in pain. Is it not absurd, he asks, 'to say of a *body* that it has pain?' (*Philosophical Investigations* §286). Marking the essential distinction here, and in such a way that the irreducibility of the 'I' is evident, Wittgenstein further asks: 'In what sense is it true that my hand does not feel pain, but I in my hand?' And how do we know that it is not—against this variety of behaviorism—the body that feels the pain? We know it by *looking* at what we do in a way very much unlike the kind of looking that is articulated on the classical Cartesian introspectionist model. ' . . . if someone has a pain in his hand, then the hand does not say so (unless it writes it) and one does not comfort the hand, but the sufferer; one looks into his face.'[20] Indeed one need only consider any of the countless representations of both sympathetic and empathetic responses to pain (either physical or emotional) in film and theater; in such cases, if one did not look into the face of the pained person, that would constitute an emotionally significant form of evidence, and hardly an indication that one is admirably and fully attending to the pain. Similarly, a viewer of a representation of a physician who, avoiding the gaze of a patient, looked *only* to the injured part would likely use the word 'inhuman' to describe the character of the physician. Our practices, placed before us by the mimetic arts, make Wittgenstein's point.

Thus, three fundamental elements of third-person behaviorism, i.e. (1) the characterization of our knowledge of another's emotional state as inferential, (2) the characterization of behavior as evidence

---

[18] I will turn to the distinct role that this kind of imagination-assisted perception plays in our reconsiderations of our own pasts in Ch. 7, Sect. 3, below.

[19] See also *Philosophical Investigations* §284, where the genuine oddity of the very idea of ascribing a *sensation* to a *thing* (an idea foundational to the behavioristic conception of selfhood) is uncovered.

[20] See also *Zettel* §§540–1, where the suspicious idea of *attending* to our own behavior, and the idea of the 'primitive' pre-linguistic bases to a language-game, are brought into play.

which allegedly precedes our knowledge, and (3) the characterization
of person-perception as body-perception, are met with strong anti-
reductive arguments and counterexamples. It may indeed at times
appear that Wittgenstein's position on the self is approximating one or
the other of the formulations of behaviorism, but on closer examination
it turns out to be merely an illusory proximity; he is certainly not
a 'behaviourist in disguise'.[21] But with the Cartesian-introspectionist
model of the self, the conception of the self that would dismiss the
bodily in favor of the purely mental, removed from consideration, and
with its inverse, the behavioral model of the self that dismisses or reduces
the mental in favor of the bodily self, removed from consideration as
well, what—and we might want to ask this now with at least some
impatience if not outright urgency—*is* Wittgenstein's view of the self
and the inner, if indeed there is one? Settling that would seem an
indisputably necessary prolegomenon to the clarification of the nature
of autobiographical self-description.

It is true that Wittgenstein has said that the mental and the physical
fall into different language-games, but what, to ask our question in a
slightly different way, is the *relation* between them? How is Wittgen-
stein's conception of the 'I' assembled from those categorically distinct
language-games? Or indeed *is* the conception, in any such sense, assem-
bled? The language-games of physical objects and sense-impressions,
Wittgenstein says at *Philosophical Investigations*, part II, §v, page 180,
are different, and there is, he says, a 'complicated relation' between
them. 'If you try to reduce their relations to a *simple* formula you go
wrong.'

---

[21]   And there is what one might call another kind of argument, one mentioned explicitly
only a few times (see *Remarks on the Philosophy of Psychology*, 2 vols (Oxford: Blackwell,
1980), i, ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe,
§§129 and 314, and *Zettel* §492), that is perhaps best described as a cumulative
argument taken from the aggregation of Wittgenstein's many remarks on all of the
topics relating to the self. 'Behavior' is itself *vastly* more complex and variegated a
concept, and encompasses a great deal more, than explicitly envisioned by behaviorism.
Again, the work of novelists, playwrights, and filmmakers shows this consistently. In
such works of art the significance of context for verbal meaning is usually readily
apparent, just as is the significance of context readily apparent for the meaning of what
we might—cautiously—generically call 'behavior'. As Glock nicely encapsulates the
matter, 'When Wittgenstein speaks of the behavioural manifestations of the mental,
"behaviour" includes not just social expressions and gestures, but also what people do
and say, and the occasions for the use of mental terms. These form a highly complex
syndrome. What counts as a manifestation of sadness on one occasion, may not on
another' (*Wittgenstein Dictionary*, 57).

2. FIRST-PERSON AVOWALS

So when do we see this complicated relationship in action, in context? Wittgenstein gives a number of cases of things we say and do (not to imply that saying is invariably categorically distinct from doing; words are deeds) that show such connections. In saying 'I noticed that he was out of humour' (*Philosophical Investigations* II. v, p. 179), do we have a report about his behavior or his state of mind? Before answering, Wittgenstein shows by analogy what his answer will be; he inserts the example 'The sky looks threatening' and asks if it is about the present or the future. He then answers 'Both' and adds—most significantly—'not side-by-side, however, but about the one *via* the other'. He follows with a further example, in which a doctor asks, 'How is he feeling?' and the answer comes from the nurse, 'He is groaning.' It would be, well, extraordinary if the doctor were then to wax metaphysical, complaining, 'I asked about his inner states, not about his behavior!' Practice, again, does not obey the dictates of the dualistic theory Wittgenstein is undermining, either on the behavioral side or on the Cartesian side of the dichotomy.[22] Here too I believe it is literature, theater, and film that provide a great, and irreplaceable, service to philosophical understanding, specifically in the sense that we are shown by these arts cases of human interaction and the forms of understanding of others and of selves that neither illustrate nor correspond to Cartesian or behavioristic conceptions of the self. Indeed, such cases function as 'reminders', in Wittgenstein's sense, of what 'lies before us' as our lived practices that are distorted in the prism of theory, and when such cases are not within our imaginative grasp, we can find either or both of the opposed philosophical pictures of selfhood once again plausible. But again (and I will return to this), the arts do not do the philosophical work *by themselves*: it is philosophy—and I am suggesting here especially Wittgenstein's wide-ranging remarks as assembled for the particular purpose of gaining insight into the self

---

[22] For an illuminating exploration of a number of psychological phenomena (e.g. unconscious motivations and self-deception) in a manner that shows how the Wittgensteinian position is free of this double-sided snare, see Richard Allen, 'Psychoanalysis after Wittgenstein', *Psychoanalysis and Contemporary Thought*, 20/3 (1997), 299–322. See also the succinct and very helpful remarks on self-knowledge (being unlike knowledge of the other) in P. M. S. Hacker, *Wittgenstein: Meaning and Mind* (Oxford: Blackwell, 1990), 95–6.

and person-perception—that irreplaceably shows us where, and *how*, to look in literature, film, theater, and the other arts. For the present, we need to get clear about Wittgenstein's use of the word 'both', and particularly that small but underscored word 'via'.

Wittgenstein, as has been seen, has good reason to oppose the view—a view imposed by the picture of object-and-designation—that utterances or expressions that give voice to inner states *describe* them. They are not utterances that work in that way—that is, in that one way. Rather than descriptions,[23] they are avowals or emotive expressions, standing in a relation to the feeling that is unlike what we would expect coming to the question with the pre-reflective (or semi-reflective) belief that all words get meaning through reference, through naming. Indeed, Wittgenstein claims—and this seems amazing at first glance—that such utterances are not *about* the inner states. The essence of his view is contained in a brief passage of *Philosophical Investigations* §244, just after he has initiated the examination of the idea of a private language. He begins with a question that seems—and it is in truth not at all what it seems—fundamental to the subject and innocuous: 'How do words *refer* to sensations?' Clearly the concept of reference—along with all the encouragement that particular formulation of the question gives us

---

[23] But it is important not to state the claim too strongly or too generally that such utterances are avowals and not descriptions. What is intended here, of course, is that they are not descriptions *as philosophically conceived* in accord with the underlying dualistic picture of the self. It is unlikely that Wittgenstein would categorically deny that the utterance 'I have been emotionally out of sorts all day today' is a description of how one feels, or that we might in some context use that word to, in turn, describe the utterance. And it is important, while keeping open the possibility of description (properly understood) in the first-person case, to keep in mind the difference in logical status between descriptions in the first- and third-person cases; i.e. when one says 'I have just raised my arm and put my hand behind my back', it makes no sense—there is no move within the language-game—to ask 'How do you know?' By contrast, when another's similar action is described, someone may ask 'But how do you know?' and we can reply 'I saw him do it', or 'He told me he did it', and so forth. The objection is to the dualistic, metaphysical construal of 'description', and the effort to supplant that concept with that of 'avowal' should be seen in that light. Wittgenstein's position is, again, complex to its core, and the least desirable result of following the course of these reflections would be to generate yet another general theory of first-person utterances. (And, as we shall see, not *all* first-person statements about the mind are in the present sense—i.e. as articulated in response to the inner-description picture—avowals. General pronouncements—like my own 'Rather than descriptions, they are avowals'—consistently run the risk of becoming anti-theory theories; much of Wittgenstein has been chronically misunderstood in precisely this way.) The antidote, again, is a vigilant awareness of particular cases—which the arts unendingly provide.

to think that a word has to have an object, outer or inner, physical or mental, to refer to if it is to mean—is smuggled into the center of the subject immediately. But less obvious conceptual maneuvers are in evidence as well: the very framing of the question in terms of *words* and *sensations* bifurcates emotional expressions along the lines first brought to attention in *The Blue and Brown Books*, where we find a heightened sensitivity to the misleading power of the question (derived from the larger dualistic model of linguistic meaning) asking how the sign gets its life.[24] The framing strongly suggests, or at its most innocent quietly reassures, our dualistic linguistic conception, the metaphysical view that the inner sensation is one thing—indeed one *kind* of thing—and that the word that communicates it is another.

Wittgenstein is opposed to every bit of this. Giving mock-gullible voice to the seeming metaphysical innocence of the line, he adds '—There doesn't seem to be any problem here; don't we talk about sensations every day, and give them names?' That sentence houses two versions of the single misconception with which Wittgenstein does battle. First, it incorporates the presumption that in using any language concerning emotions we are thereby talking *about* them, and second, it incorporates the presumption that in going so far as to talk about them we must have thus first *given* them names. Both aboutness and name-giving imply that the relation between the word and the experience is contingent, and that the speaking mind makes the connection between name and inner referent, between inner object and its designation. In the next line, however, Wittgenstein interjects his own voice: 'But how is the connexion between the name and the thing named set up?' Wittgenstein often returns to the question of language-learning in *Philosophical Investigations* as a way of removing the philosophical mist surrounding mind–language issues, and he does so here: 'This question is the same as: how does a human being learn the meaning of the names of sensations?—of the word "pain" for example?'

In fact the questions are not quite the same; the latter question leads us to inquire into an account of learning the *use* of the word, and it at least can rid us of what he calls the 'stage setting',[25] the

---

[24] See *Blue and Brown Books*, esp. 4–6, 15–16, 34.
[25] See *Philosophical Investigations* §257: the issue is there discussed inwardly or in psychological terms, where the typically unwitting presumptions behind the idea of giving a *name* to a pain are exposed.

unwitting importation of the tripartite concept of name and of object and of the mind's deliberate imposition of a referential link between them. Thus, in the suggestion he gives—and the suggestion is a microcosm of the larger relation (if we can call it that) between (*a*) the self and emotionally expressive language and (*b*) the self and expressive behavior—we find words gradually taking the place of 'the primitive, the natural, expressions of the sensation'. This is an instructive suggestion for three reasons: (1) it places expressive language into a context of natural and primitive expressive gestures rather than a context of deliberate, stage-set, ratiocination on the part of the speaking self; (2) it reawakens and mobilizes our strong intuitive sense that, as gestural expression is naturally occurring expressive behavior, it would be strange indeed to suggest that primitive and natural behavioral expressions are learned cognitively through one-to-one correlation memorizations, e.g. screaming in one way means great pain, grimacing means moderate pain, screaming another way means great fear, smiling in one way means love, smiling in another way means malice, and so forth through a truly vast catalogue of such imagined correlations; and (3) it suggests that the very model of object and designation, of a separation between, for example, pain and its expression, is a conceptual illusion, that there may be another way of conceiving of the relation. But the last suggestion in turn is easily misconstrued under the influence of the misleading philosophical picture. Wittgenstein thus gives the example (*Philosophical Investigations* §244): 'A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.' The philosophical voice then asks, unwittingly forcing the suggestion back into the old model, the old picture, 'So you are saying that the word 'pain' really means crying?' and Wittgenstein's voice replies, most significantly for the careful articulation of this new conception: 'On the contrary: the verbal expression of pain replaces crying and does not describe it.'

Description, implying a separation between the describer and the described, would drive a wedge between pain and its expression. But, as Wittgenstein now asks pointedly in the next section, 'For how can I go so far as to try to use language to get between pain and its expression?' Thus, in first-person cases, avowals—or emotive expressions of the self—function in the way that instinctive or natural expressive behavior functions (and we are here escaping the prejudice that all language works in the same way), and not like the descriptions we might give of any

object upon close visual or tactile scrutiny.[26] So the suggestion, indeed a radical recontextualization of our conception of verbal expressivity, clearly goes against the traditional introspectionist model of the self and the correlated conceptions of our interpretive knowledge of the self's content and our outward expression of that inner content.

But doesn't that suggestion, against what we might expect at this stage, give new life to behaviorism, the other polemical opponent here? If we witness the behavioral manifestations of others and draw conclusions about them on that basis, have we not arrived at precisely what Wittgenstein was *against* a moment ago? The answer is negative, despite a fleeting semblance. When the nurse above sees and hears the patient groaning, she does not (1) observe a body in movement, nor does she (2) hear a statement *about* pain being uttered, nor does she (3) interpret an external sign that gets its life by referring to the inner object of pain. Each of those misconstruals fails to acknowledge, here again, what Wittgenstein described as the distinctive *stance* or *attitude* that we have toward other human beings. While it may be true that the patient's behavior tells her something about his condition (and that essentially is the element that makes it appear that Wittgenstein is sliding into his opponent's camp), that is not, again, *inferential* knowledge. And if she gives voice to the words in the example in *Philosophical Investigations*, part II, §iv, page 178, 'I believe that he is suffering', she does not exhibit an hypothesis concerning his state on the evidence of his behavior, but rather *responds* to his actions, be they linguistic, gestural, or instinctively behavioral (not to suggest that those are invariantly separate categories), as a human being. Her attitude toward him, as Wittgenstein puts it on the same page, is 'an attitude towards a soul'.

[26] The claim that first-person avowals *function* in the way that instinctive and natural expressive behavior functions, although therapeutic in one way, is nevertheless problematic in another. One might say instead that first-person avowals are logical extensions of instinctive expressive behavior to evade the issue of the similarities or differences in the functions of utterances, since the similarities or differences are, in detail, context-dependent. The naturally expressive behavior of an infant itself has numerous and different functions, e.g. to express urges, needs, curiosities, pains, etc., so it can be misleading to suggest a *singular* function for any utterance-type. Also, emotive expressions build upon, refine, and extend primitive expressive behavior, thus enabling extended and even new functions (the expression of warmth for another, for example). One might thus say that they have the same logical status, perhaps, as primitive expressions, but may function differently. The somewhat less particularized point here, expressed in terms of the sameness of function, is that first-person expressive avowals can be better (and newly) understood when placed with instinctive or natural behavior than with descriptions of objects given subsequently to close inspection.

Thus, behavior is relevant—as part of the criteria[27] for the state—but its role is easily misconstrued under the influence of the various (but related as branches stemming from the same trunk) dualistic theoretical pictures of inner and outer, object and designation, sign and its life, and inner object and outward designation. Expressions of states like pain, or *avowals*, are not for Wittgenstein *descriptions in the metaphysically implicating way*; language does not get between pain and its expression, and the nurse's attitude toward the third-person case only reaffirms those points.[28] But it is important to bear in mind, again, that nothing that has been said would fully assimilate the first- to the third-person case: there is always the possibility of third-person dissimulation, of pretending,[29] and that does not make sense in the first-person case (one can fool others, but one cannot *oneself* be fooled about the state by the self).[30] And there is as well still the sense in which the

[27] Again, this is a separate and major component of Wittgenstein's philosophy. In brief, one might consult *Blue and Brown Books*, 24–5, and *Philosophical Investigations* §§290, 353–4, and II. vi, p. 181, and xi, pp. 222–9. But the issue fans out throughout Wittgenstein's later philosophy, and there is an extensive secondary bibliography on the theme. For clear and historically informed discussions, situating the concept of a criterion into the larger story of the development of Wittgenstein's philosophy, see P. M. S. Hacker, *Wittgenstein's Place in Twentieth-Century Analytic Philosophy* (Oxford: Blackwell, 1996); see p. 336 for references.

[28] It is worth pausing to contrast the notion of an avowal with the earlier 'picture theory of meaning' of the *Tractatus Logico-Philosophicus* (trans. D. F. Pears and B. F. McGuinness (London: Routledge & Kegan Paul, 1961)). On the early picture theory, propositions are meaningful just and only because they record or express a thought which itself constitutes a representation of a particular and determinate state of affairs in the world. An avowal, as it emerges in Wittgenstein's transitional and later philosophy, *is* a meaningful utterance, but in a way that is *not* descriptive, that is not dependent for its meaning on picturing a state of affairs. It is, in short, a concept, within Wittgenstein's development, that clearly marks the break with the picture theory of meaning. For a discussion of the distinctive nature of the 'attitude' referred to here and its significance for aesthetic understanding, see my *Meaning and Interpretation*, 131–8.

[29] Dissimulation and pretending would thus be understood *not* as giving one-half of the genuine case, i.e. only the external behavioral indications of an inner state that in the pretending case is not there, but rather more along the lines of 'indicating' in the theater as mentioned above. The latter way of construing 'pretending', or dissimulation, avoids the misleading additive model, in which the genuine (nonpretending) case is conceptualized falsely as the two isolable parts, inner and outer, appearing on their separate stages simultaneously.

[30] Which is not to make the very different claim that self-deception is impossible. (I return to this topic below.) The Wittgensteinian construal of self-deception, as Richard Allen shows (see n. 22 above), is every bit as delicate a matter as one might now expect. Its full elaboration (which would also be aided by artistic studies of the phenomenon) does not in any event bring about the assimilation of the first- to the third-person case; one does not deceive oneself by exhibiting overt behavior that one then, as though

nurse's statement 'He is groaning' *is* a description of the patient, whereas the groan itself is not a description of a secret entity in an inner realm of inwardly perceived objects.[31] *Responding* to a first-person utterance of that general type as a description is ordinarily unthinkable. Describing often brings with it, or there naturally arise as moves within descriptive language-games, the possibilities of (*a*) more closely observing, (*b*) considering and (*c*) reconsidering, (*d*) striving for accuracy, (*e*) correcting oneself, (*f*) comparing, and many related possibilities, and shortly before Wittgenstein makes those clear in *Last Writings on the Philosophy of Psychology*, volume i, §51, he asks: 'Does someone crying out "Help!" want to describe how he is feeling?' He answers simply: 'Nothing is further from his intentions than describing something.'

Avowals give us a very different picture, a corrective, to the conception of object-and-designation as it is applied to the inner life, and understanding them breaks the hold of the idea that the meaning of what we say is invariably a function of our describing facts of a case; in *Philosophical Investigations* §292, we find, 'Don't always think that you read off what you say from the facts; that you portray these in words . . . '.[32] Avowals break the hold of misconceptions of self-expression, from both the first- and third-person perspectives, and thus inform us about the nonunitary nature of the self's relation to language. Is 'The sky looks threatening' about the present or the future? Again, *both*, and not side by side, i.e. not in a way corresponding to those dualistic pictures, but in multiform ways as they occur in ordinary experience. And in words such as 'I noticed he was out of humour', we have a sentence concerning both a person's behavior and his state of mind, *not* side by side, with the mental and the physical separate but coterminous

viewing another (in a mirror, say), observes and (if the deception is successful) finds convincing. The actual cases are far more interesting than this dualistically motivated mischaracterization would suggest. See esp. Allen, 'Psychoanalysis after Wittgenstein', 315–18.

[31] See *Philosophical Investigations* §§289–91. In reference to the meaning of first-person pain utterances, Wittgenstein—going against the presumption that a *justification* for such an utterance is necessary to its meaning and that the justification will take the form of an inner object to which the pain refers—asserts: 'To use a word without a justification does not mean to use it without right' (§289).

[32] To contextualize the remark to grasp its meaning better, see also *Philosophical Investigations* §§290–1: 'Perhaps,' Wittgenstein suggests, 'the word "describes" tricks us here' (§290), and what we call descriptions are in fact 'instruments for particular uses' (§291).

in the phrase, but one, again, *via* the other. For Wittgenstein, the mental, and the emotive, are *manifest* in behavior; most importantly, they are not reducible to behavior, but nor are they metaphysically separate from it. It is, I believe, true that that claim, from a traditional philosophical perspective, and from a considerable distance, appears at a glance far *closer* to behaviorism than to Cartesianism, and for that reason Wittgenstein's repudiation of behaviorism has been much less well understood. But the deeper and fundamental position on the self (being incrementally developed as his thought unfolds) is, in relation to traditional philosophical categories, again radical in the extreme. It is the overcoming, and the undermining, of the conception of the Cartesian or introspectionist self on the one hand and the behaviorist conception on the other, and it thus constitutes a radical critique of conventionally entrenched philosophical and psychological theory. Representations of selfhood in the arts, to the extent that they show these self-informative particularized practices—and thus resist the impulse to theorize the self in unitary terms—substantially contribute to this critique.

In *Philosophical Investigations* §357 Wittgenstein puts it compactly: 'If one sees the behaviour of a living thing, one sees its soul.' The inner—but calling it that without Cartesian or introspectionist impli-cations—suffuses the outer—here calling it that without behavioristic implications; it is the dualistic characterizations of the interlocutor's philosophical voice (against the grounding influence of the ordinary voice) that are myths of theory. And in those connections Wittgenstein also famously asserts in *Philosophical Investigations*, part II, §iv, page 178: 'The human body is the best picture of the human soul.' His position is thus *not* Cartesianism, *not* introspectionism, *not* behaviorism, *not* any kind of reductionism, eliminativism, nor—despite the appearance of the claims just mentioned if read in isolation from their context—monism in response to dualism. We have seen, in the discussion of first-person avowals and their third-person receptions, something approximating a positive basis on which to proceed, but that discussion is as much designed to loosen the grip of a philosophical picture—or a conspiring cohort of them—as it is to advance any genuinely positive thesis.

So we must ask, *is* there any more to a genuinely positive position concerning the nature of what we call the inner self? And whether there is or not, we also need to know, if the inner is not metaphysically hidden and if first-person ascriptions of psychological predicates are expressions or avowals, how the familiar phenomena of hidden dimensions or aspects of a person (as we saw them in Chapter 1), of self-deception (as

we will see it examined by Kierkegaard), and of self-discovery (as we will see it in Augustine) are so much as possible on a Wittgensteinian view. Those are clearly common in our ordinary language; how do we characterize those language-games in a way that does not fall back into picture-driven conceptual confusion?

It is time to turn to Wittgenstein's alternative, non-picture-driven conception of introspection, the introspection that we see undertaken in life, depicted in literature, theater, and film, and in a kind of double-exposure, both enacted and represented in autobiography.

3. REAL INTROSPECTION (AND KIERKEGAARD'S SEDUCER)

The foregoing catalogue of misconceptions, motivated by metaphysical pictures embedded in language, presumes that there is an epistemological problem with regard to the self's inner objects. Although skepticism concerning third-person knowledge is also misleadingly motivated by those pictures, it is at least true that in third-person cases of emotional ascriptions we *can* be wrong, we can make mistakes—and thus there is one kind of epistemological problem here (but in truth unlike the problem envisaged within those preceding philosophical positions). But there is no problem *of that kind* in the first-person case, precisely because we do not introspect upon, or observe, or perceive, or unveil the hiddenness of, our own sensations or emotions; Wittgenstein argues that we simply *have* them. Thus, he claims in this connection, it is 'wrong to say ''I know what I am thinking'' ', although right to say 'I know what you are thinking'; it is to this contrast that he adds his famous parenthetical remark 'A whole cloud of philosophy condensed into a drop of grammar' (*Philosophical Investigations* II. xi, p. 222). The unmediated sense of 'having' in place of observing, perceiving, etc. also condenses a cloud into a drop, but to say it, corrective in one sense, brings with it a danger of conceptual relapse. If we *have* those *im*mediately, are we not reverting to first-person incorrigibility of the Cartesian kind? The answer is No,[33] but to avoid the relapse

[33] One exceptionally helpful article that shows that the answer is No is Sydney Shoemaker, 'Introspection and the Self', in Peter A. French, Theodore E. Uehling, and Howard K. Wettstein (eds), *Studies in the Philosophy of Mind*, Midwest Studies in Philosophy, 10 (Minneapolis: University of Minnesota Press, 1986), repr. in Cassam

we need to become mindful of our actual practices, the examples of self-knowledge and of self-deception, and of the genuine problems of

(ed.), *Self-Knowledge*. Shoemaker, showing a characteristically heightened sensitivity to the power of tacit or unexamined presuppositions, shows (among other valuable things) that the alarm that we may feel upon contemplating the possible truth of Hume's denial that we have introspective awareness of a self could be significantly (and, I would add in connection with the present discussion, very tellingly) lessened. 'If we did not take it for granted that we do perceive *something* by introspection, and that this introspective perception is the source of our introspective knowledge, it would not be so likely to strike us as significant or disturbing that we do not introspectively perceive any self or mental subject' (Cassam (ed.), *Self-Knowledge*, 124). One way of expressing the importance of this point (and it is considerable) is to say that it is only so long as we model introspective knowledge or awareness on *perceptual* knowledge, on the model or empirical picture of external-object perception but turned inward, that we will stay locked in the puzzle about how the 'inner sense' that is required by this model works, and how to describe the perceptual relation we would thus have to ourselves (or, to put it more accurately, to our inner self-object, where that is construed as an object of inner perceptual focus). Another way to put this would be to say that Shoemaker here, by taking the perceptual model for introspection extremely seriously and subjecting it to the closest scrutiny, admirably shows how extremely little it does for us, i.e. how little it explains that cannot, as he says, be equally well explained without it. In response to the objection that he may have overlooked 'the point that only one self could be the object of my introspective perception' (p. 128), where he has already made the observation that 'introspective observation of a self being angry is not going to yield the knowledge that *I* am angry unless I know that that self is myself' (p. 128), he writes: 'that is a piece of self-knowledge I could not get by introspective observation; for unless I already know that this self is myself, observing that it perceives itself is not going to tell me I observe it. So it remains true that, if I am to get self-knowledge by introspective perception, I must have some that I have not got by introspective perception' (p. 128). And so, if in standing back from these exacting points we feel all the more doubtful about the applicability of the perceptual model to introspection (and perhaps more sympathetic to the nonobservational character of some self-knowledge—although that is a separate topic for another day), then these reflections (yielding Shoemaker's most sharply focused conclusion, 'At best, the hypothesis that there is introspective self-perception seems to explain nothing that cannot be equally well explained without it', pp. 128–9) provide a powerfully therapeutic service in removing the presumptive allegiance to the blinding picture of what we might, as a shorthand, call object-observational introspection.

See also a closely related discussion of Donald Davidson's, in his 'Knowing One's Own Mind' (in Cassam (ed.), *Self-Knowledge*), where he articulates the virtues of getting 'rid of the metaphor of objects before the mind'. He continues: 'Most of us long ago gave up the idea of perceptions, sense data, the flow of experience, as things "given" to the mind; we should treat propositional objects in the same way. Of course people have beliefs, wishes, doubts, and so forth; but to allow this is not to suggest that beliefs, wishes, and doubts are entities in or before the mind, or that being in such states requires there to be corresponding mental objects' (p. 62). I would add that the mere recognition that this is after all a *metaphor* in play here goes a good way toward freeing us from the grip of the conceptually domineering presupposition. And it opens the way to seeing the philosophical relevance of the close study of artistic depictions of beliefs, wishes, and doubts as they are experienced, expressed, and described by human beings *in situ*.

self-knowledge that are, again, experienced in life and best depicted in literature, theater, and film, i.e. in the representational, and particularly self-representational, arts.

Self-deception, in particular, provides a telling phenomenon where conceptual relapse of the kind just described can seem so natural that indeed one has to struggle against the 'undertow' dragging us back to a Cartesian picture of the mind and mental activity. This struggle, itself perhaps reminiscent of Wittgenstein's later struggle against the pull of his own earlier views, i.e. the linguistic atomism and the picture-theory of meaning of the *Tractatus*,[34] is against the picture of self-deception as the hiding from the self of an inner object already hidden from others. Self-deception is often, and seemingly naturally—if extremely misleadingly—modeled on a common picture of other-deception, i.e. the self is thought to deliberately hide or conceal the truth it knows from itself just as one who knows the truth deliberately conceals it from another. On these conjoined and misleading models, the hidden content, in the case of other-deception, is kept away from the other's inspection; in parallel, but inwardly, the hidden content is kept away from the self's introspection.

The above modeling[35] of the phenomenon has led to the standard formulation of the puzzle of self-deception in the philosophy of mind,

[34] I discuss the *Tractatus* in its attempted application to artistic meaning in *Art as Language*, 'Art and the Unsayable'.

[35] Another way of bringing out the extent to which this standard and widely accepted—and I believe instructively erroneous, as we shall see below—model of self-deception relies upon a prior unacknowledged endorsement of the picture of the inner Cartesian theater is to focus on the word 'knows' in 'S knows that P and believes not-P'. 'Knows' here actually implies recursive awareness (otherwise there would not be a contradiction-problem of self-deception), so that 'S knows that P' means 'S knows that she knows P', where indeed S knows her belief by (1) introspecting on the immediately accessible and transparent 'P', and where (2) she knows it is she who is introspecting. It is, as we will see, this entire structure of the problem and its Cartesian undergirding to which the Wittgensteinian view provides an alternative (or better, escape). To allow the very formulation of the problem in these terms is thus not genuinely to clarify what is philosophically at issue with the diverse phenomena of self-deception, but to front-load the Cartesian picture in such a way that escaping its influence in any subsequent attempt to contend with the problem *in these terms* is almost impossible.

I was pleased to learn, after completing this study, that Richard Moran makes a parallel point about the larger question concerning the nature of the relation of the mind to itself, its contents, and its operations, in his exceptionally lucid and helpful *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001). He writes: 'The legacy of Cartesianism has been decisive in the philosophy of mind not only in the positive influence it exerted in the centuries immediately following Descartes, but just as much in the force of its repudiation in the twentieth century. Nowhere

i.e. how can S know that P and believe not-P? And this in turn naturally invites a broadly Freudian layering or dividing of the mind so that the other-deception model can be 'interiorized' with one layer or division of the mind concealing P from S's introspective gaze at another level or in another division. Indeed this model of self-deception so readily suggests itself, it is so easily presumed, that it is just here that one might pointedly ask how the very phenomenon of self-deception is so much as *possible* on the Wittgensteinian view. If Wittgenstein is dismantling the traditional dualistic model of introspection, and if indeed, having followed the passages on avowals versus inner descriptions, we understand that we now just *have* emotional states rather than perceive them inwardly, how then *could* we deceive ourselves? And since, as a fact of life, we know that we can and do indulge in self-deception, is not this blunt fact itself a

---

is this clearer than in the question of the mind's access to itself and its operations. Recent philosophy typically rejects the picture of the mind as immediately transparent to itself, and then tacitly takes this rejection to be equivalent to rejecting the very idea of introspective access, thereby ceding the very concept of first-person awareness to its Cartesian interpretation' (p. 4). To cede the concept in this way is, to put it in the terms of the present discussion, to allow the picture to dominate; as we have seen in the first section of this chapter it easily does so both in its affirmation and, as Moran is discerningly calling the right kind of attention to in his study, its rejection (in the ways that, as he is observing, have become typical since the mid-twentieth century). Freedom from the picture in its rejection as well as its affirmation (that is, of the 'inner picture', as we saw Wittgenstein call it in Chapter 1, Sect. 2) is, as we will see with increasing clarity as the following chapters unfold, an intricate matter that does not reduce to polemical position-formulas. This is just like—if here in microcosm—the present problem of escaping the 'S knows that P and believes not-P' formulation but in a way that both unearths its Cartesian substructure and resists opting for an equally schematic or formulaic alternative reductive structure. Incidentally, Moran hits another methodological nail on the head by calling for significantly increased attention to the agent in elucidating more fully the first-person perspective; much of the debate concerning self-deception has proceeded by focusing abstractly on the beliefs but not concretely on the believers. Moran writes: 'A more complete characterization of the first-person perspective will require bringing the agent more explicitly into the picture, and doing so will involve taking the discussion into a range of issues concerning the agent's perspective of deliberation and self-interpretation that have not been at the center of recent discussions of self-knowledge' (p. 33). That is one way of describing what my attempts to turn to, and integrate, literary examples of self-knowledge and what I am calling autobiographical consciousness throughout this book are designed to do. (I should perhaps note, as a task for another day, that I do have a number of points of disagreement with Moran's elucidation of the expression of first-person knowledge—most having to do with his on occasion perhaps not *fully* acting on his own extremely good advice concerning the need for a new-found focus on the agent. It is, as I am suggesting throughout this book, only the most acute, exacting, and sustained concentration on the contextual particularities that will deliver the freedom, or again the radical methodological departure, from the pictures and their conceptually linked polemical repudiations that Wittgenstein is working through and beyond.)

strong argument against Wittgenstein's position? But the strength is not in a *reductio* of this kind, but rather, again, in the conceptual undertow.

The counter-Wittgensteinian argument works only if both the dualist-introspectionist conceptual model of self-deception is accurate and self-deception occurs as the model suggests. A study of cases—again supplied by the self-representational arts—shows that, while self-deception does most assuredly occur, it need not be explained, and indeed is not best explained, in terms of the model. Self-deception can be more plausibly, and less problematically, characterized as miscontextualization, precisely where the miscontextualization is *motivated*, and not merely accidental or the result of insufficient attention. If one does not, for any of countless contextually specific reasons, want to recognize and acknowledge a truth about oneself or one's circumstances, one can piece together a mosaic of cases, examples, and particular details that together add up to a false picture. And as part of this self-shielding process, one can pay only selective attention to one's past and anticipated thoughts and deeds, drawing lines of continuity among particularities that bring out a pattern that is convincing enough. It is also, not surprisingly—and consistent with the unfolding Wittgensteinian conception of real introspection—within such language-games of self-deception that we can experience what we revealingly call the 'nagging feeling' that we are avoiding, not hidden inner objects concealed from introspection, but many relevant particularities of thought and deed that would add up to a very different picture. A similarly familiar phrase in such contexts, 'the tip of the iceberg', is better explicated as a repository of experience that would greatly outweigh the evidence to which we are selectively attending than as a doubly hidden inner object. It is also within such language-games that we encounter the distinction between explaining and 'explaining away'; for example, a person fearing spousal infidelity may explain away repeated absences as a result of an unusually demanding schedule; a mother fearing a drug-addicted son may explain away the paraphernalia of drug use as the left-behind possessions of others; a person fearing a medical condition may explain away advancing symptoms as fleeting bodily manifestations of short-term stress; and so forth. In each of those cases (all of which have been portrayed in film, e.g. in Bergman; in theater, e.g. in Beckett; and in fiction, e.g. in Dostoevsky, again at a safe aesthetic distance, allowing us to take in their nuances, contemplate them, and learn from them with a reflective calm often unavailable within the contexts of our own lived experience), it is not so much that S knows P but believes not-P as it

is that S is actively assembling a mosaic of particulars that serves the underlying motivation to believe that the spouse is faithful, that the son is 'clean', and that one is not really ill. These, properly understood, are instances of motivated miscontextualization, not of inward concealment *as metaphysically construed*; if we speak of inward concealment, within these games this will mean that real self-exploration has not yet been undertaken, that real introspection, as a process of open and nonselective 'mosaic' construction, has not been accomplished. This is precisely the kind of contextualizing, constructive, self-interpretive work that we see undertaken and accomplished in great works of philosophical autobiography such as Augustine's *Confessions* (to which we will return in Chapters 4 and 5).

Works of this kind show us—or remind us—how to discriminate between different kinds, or perhaps categories, of self-ascriptions of psychological predicates; my claim is not that artists *are* philosophers and that artworks are, or are created to function as, philosophical texts: they are not, since the fundamental work of art is not conceptual clarification. But self-representational works of art—in fiction, auto-biography, poetry, film, theater, opera, and the visual arts—do manifest, or enact, instances of real introspection and they delineate the numerous psychological concepts in play within contexts of self-understanding and self-interpretation. And we are particularly well positioned to learn from such cases in the arts, precisely because, while the relevant psycho-logical concepts and enactments of introspection are presented within the stream of life depicted within the artwork, those concepts and enactments, here again, are not as we perceive them in motion within the stream of our *own* lives, i.e. it is in this precise sense that we have the luxury of viewing them from a safe, indeed aesthetic, distance.[36] Some self-ascriptions of psychological predicates, as artistic examples show, are incorrigible[37] and some are not: one cannot be mistaken about being in pain in the way that one can be mistaken about being in love. (One, for example, has a difficult time thinking of a single case of the former

---

[36]   This is, I believe, the kernel of truth within the earlier 'aesthetic distance' theories; see Edward Bullough, '"Psychical Distance" as a Factor in Art and as an Aesthetic Principle', *British Journal of Psychology*, 5 (1912), 87–117, repr. in George Dickie, Richard Sclafani, and Ronald Roblin (eds), *Aesthetics: A Critical Anthology* (New York: St Martin's Press, 1989).

[37]   This is a dangerous word to use; it does mean that there is no room to speak of error, but *not* because the inner object is immediately accessible to Cartesian introspection, as we have seen.

in the arts; aesthetic investigations into the latter are too numerous to mention.) Human representations in the arts explore, and cultivate our awareness of and sensitivity to, seemingly infinite gradations of human psychology, and they correspondingly expand our conceptions of that psychology by giving us enormously complex and nuanced character-izations of emotional states. Furthermore, the arts of self-representation do not only mimetically depict a preexistent state or condition of psychological and emotional being, they can also—and this is especially true of autobiography—provide instances of self-exploration and self-discovery. In such cases the artist discovers new dimensions of him- or herself, or achieves new depths of self-understanding, through the very act of artistic creativity or composition itself.[38] And there opens a two-way street between philosophy and the arts of self-representation and particularly self-investigation: the arts assemble reminders and particular cases that taken together show how the Wittgensteinian positive conception of introspection is true to those human practices, and the Wittgensteinian conception of real introspection shows us, at the same time, how to understand—which often comes to showing us how not to misunderstand—those cases and how to discern their philosophical significance. But we still need to understand more fully the positive conception of introspection itself.

We have seen Wittgenstein's objections to the very idea of description, both for what it says about mythical psychological acts and for what it implies about mythical inner entities. Yet, once again moving from the philosophical to the ordinary voice, and as we began to see above, there can be such a thing as describing in such cases. The man crying 'Help!' is not describing how he is feeling, but in *Last Writings on the Philosophy of Psychology*, volume i, §§48–50, Wittgenstein observes that there are intermediate cases, 'transitions', between what we would and would not call a description. And with ordinary usage as his criterion, i.e. retaining a greatly heightened sensitivity to the difference between the metaphysical voice and the ordinary—and with an attendant mindfulness of the capacity the former has to assume falsely the appearance of the genuine

---

[38] It is difficult to find a better illustration than Augustine's *Confessions*, in which this *active* sense of retrospective understanding worked out in and through the writing itself is almost constantly in evidence. I offer a sketch of this great autobiographical undertaking in M. Jolly (ed.), *The Encyclopedia of Life Writing* (London: Fitzroy Dearborn, 2001), 59–61; as mentioned, we will return to some of the details of Augustine's text in Ch. 4 and, particularly with regard to the problem of verifying retrospective self-interpretations, Ch. 5.

meaningfulness of the latter—he writes: 'the phrase "description of a state of mind" characterizes a certain *game*. And if I just hear the words "I am afraid" I might be able to *guess* which game is being played here (say on the basis of the tone), but I won't really know it until I am aware of the context.' Thus, after all that has been said, there *is* something called 'a description of a state of mind', but it will be one of a number of interrelated language-games, and the context is essential for knowing to what language-game the words and phrases we are hearing belong.[39] But such descriptions do *not* in any sense proceed in terms of the philosophical picture of the description of an object in a hidden, inner room. They proceed, rather, by sensitively attending to the precise utterances made within those games. This requires a contextually nuanced awareness of precisely which game is being played, and of the contextually seated usages of the words within those games. In such cases, a self may well describe its state of mind, but in an innocuous and comprehensible sense, i.e. where the description is given in the ordinary voice, within a language-game where the very word 'description' does not import the dualistic metaphysical freight detailed in the foregoing. And it is within such games that the self can indeed experience a distinct kind of discovering-through-describing, where the self is trying various formulations of self-description that show their special plausibility[40] or fittingness[41] in, and *only* in, the self-interpretive contexts where mosaics of particulars are assembled in varying ways. It is precisely such cases that show the real meaning of the phrase 'self-description', i.e. in a manner that is at once nonCartesian and nonbehavioristic.

Knowing thyself is thus not a matter of introspecting, in the metaphysical sense of that term, on the inner objects contained in one's private Cartesian interior. It is, rather, a matter of introspection very differently understood, a matter of reflecting on oneself and one's

---

[39] For an examination of Wittgenstein's phrase 'language-game', within his philosophy of language and then as the phrase can cast light on literary cases, see my *Meaning and Interpretation*, 'Language-Games and Artistic Styles', and then in terms of the significance of tone and context for hearing how to *take in* utterances, see 'Circumstances of Significance' and 'Aspects of Interpretation'. For a reading of a literary case that (in this latter sense) describes a state of mind, see my 'Dencombe's Final Moments: A Microcosm of Jamesian Philosophy', *Henry James Review*, no. 18 (1997), 223–33.

[40] I return to this issue (as self-interpretive rightness) in Ch. 5, Sects 1 and 2, below.

[41] 'Fittingness' is yet another dangerous word, in that the metaphor can lead us to believe that the external word fits, or corresponds in its 'shape', i.e. in its descriptive correspondence, to the contours of the inner object, which is of course precisely the picture being combated here.

actions—one's words, deeds, gestures, thoughts, second thoughts, hopes, fears, aspirations, doubts, wishes, needs, and countless other things that take a central (or perhaps informatively peripheral) place when recalling the actions and utterances and the context within which they took place. Self-discovery, according to this Wittgensteinian conception of introspection, might consist of a number of different, but interrelated, phenomena: (*a*) the relatively simple act of explicit reflection itself taking place for the first time, i.e. one may never have thought deeply, seriously, and openly about such and such an action before; (*b*) acknowledging the way in which a given action ties up with the context in which it took place and which one had refused to do before, e.g. one may have simply refused to admit to oneself before that a given action of one's past was part of a pattern of malicious behavior (which one now, through seeing a new network of interconnections, is in a position to admit); (*c*) interpreting a past action in a new way (or, as we say, seeing it in a new light, to which I will return in Chapter 6, Section 3) by seeing a new connection between it and an aspect of the context within which it occurred, or a new connection to future planned or anticipated behavior, e.g. one might suddenly say 'Yes, I was in love with her even as far back as then', upon seeing how a given action in the past—one that seemed at the time meaningless with regard to one's emotional attachments—formed part of a pattern of loving behavior toward someone that is only visible retrospectively and that takes on moral force only retroactively; (*d*) working out how one felt about someone by imagining counterfactual contexts, e.g. where one reflects 'If I had still been with X, I would not even have noticed Y—therefore I do not now think that I could have been really in love with Y as I thought back then'.[42] These varieties of self-reflection, in which we place our behavior, our thoughts, and our emotions in contexts of previous and subsequent similar episodes and occurrences, depend on the fact—hidden by the Cartesian misconception of introspection—that very often the content of such behavior, thought, and emotion is *relational*.[43] The meaning, the significance, the import of countless such human occurrences are not given in the occurrence *simpliciter*; meaning in action, as in language, is profoundly and ineliminably contextual,

[42] Cases of these four kinds are meant merely to begin to suggest the extended range of possibilities; to mention one example, in many novels of Iris Murdoch we very often find explorations of interrelated, and often overlapping, cases of life-transforming contextualized introspection of just these kinds.

[43] I will return to this matter in connection with Davidson in Ch. 7, Sect. 2, below.

and just as words and utterances hold significance within—and only within—their contexts of utterance, so the content of real introspection—again, hopes, fears, thoughts, second thoughts, and everything else mentioned above (and much more)—is discernible only within the specific language-games of thought and behavior as these organically unfold in context. Again, this is the kind of truth that is readily shown by our practices outside philosophy but which one readily forgets inside philosophy;[44] the misleading picture of the mind and its contents would suggest that such mental events are isolable and given in psychological experience with clearly delineated boundaries. As artistic representations and enactments of self-investigation show (again, as manifestations or enactments of psychological—and particularly introspective-interpretive—concepts, not as deliberate instances of philosophically motivated conceptual clarification),[45] they are not.

To see better that psychological experience is not atomistic, is not isolable, we need only recall that an action, a thought, a memory, a desire—like a word—takes on an inflection or distinctive shading with repetition. William James, we might remember, within a larger context of arguing for the nonatomistic nature of conscious experience (the 'stream of consciousness'), emphasized that the exact repetition of one given experience, *strictly speaking*, is impossible. Taking a relatively simple case, James gives as an example the hearing of thunder: if we hear it again, the very fact that it is a reoccurrence, a second clap, places the second in relation to the first, and thus changes the second or gives it an aspect the first did not, and could not, have had. The

[44] But it does not follow from this truth being readily shown within our practices that we therefore, independently of philosophy, grasp the concept of real introspection clearly and in a way that allows us to state it perspicuously. It is a leitmotif of Wittgenstein's philosophy that, despite the implicit clarity of our practices, we are not routinely able to give surveys or perspicuous overviews of the 'logic' of our deeds, the philosophical 'grammar' of our practices. Doing that, as Wittgenstein shows repeatedly, can prove extraordinarily difficult. We are taught, or we learn from within our form of life, to use concepts, but we do not thereby learn how to describe perspicuously, or give a conceptually clarified account of, *how* we use them. It is precisely that task that falls, on Wittgenstein's later conception of philosophical method and progress, to the philosopher. If such a great gulf did not yawn between our practices—what lies before us—and our philosophical descriptions of them, there would in fact be no conceptual confusion, and philosophy as conceptual clarification would have no *raison d'être*.

[45] But there are exceptions in which artists *are* knowingly and intentionally working as (Wittgensteinian) philosophers, pursuing conceptual clarification explicitly, unlike the great majority of cases in which, again, philosophically intricate psychological concepts are implicitly manifest in the work; the fiction of Iris Murdoch is, I believe, knowingly philosophical in precisely this sense.

point applies with equal force, James rightly suggests, to any experience given to consciousness, and we might apply the point to the experience of works of art as well: like the inflection or distinctive shading a word takes on with repetition, so our hearing, seeing, or reading again of a work will occur at a different point in the stream of life. Thus, an experience—a word, a work, or a remembered action, thought, or desire, or a memory itself remembered—is inextricable from its context. 'Repetition' is a word we use in ordinary discourse, but it too is easily philosophically misconstrued; if we think of repetition crudely, as simply another instance of the same experience, we can slide easily into thinking of experience in atomistic, isolable terms—which as we have seen nourishes the misconception of the mind holding isolated pieces or bounded contents of experience upon which it introspects. But that distorted picture of the mind and its contents, like the Cartesian-introspectionist misconception that the Wittgensteinian position and expansive conception of real introspection unearths and ultimately supplants, has its antidote in our contemplation of lived, relationally intertwined experience as James's writings on the stream of consciousness say it and as the arts of self-representation show it.

It is also now perhaps clear that the picture of atomistic or bounded, isolable mental experience James is combating substantially contributes to the misconception of what we call 'hiddenness' in human affairs as we examined it in Chapter 1. We can, and do, find particular persons to be open, or the reverse, to us: as Wittgenstein mentions in *Philosophical Investigations*, part II, §xi, page 223, some people can be transparent to us while others can prove a complete enigma. Individuals can be deeply enigmatic—by being both wholly unforthcoming while having a range of experience utterly unlike our own, for example[46]—but, for reasons now deeper and more complex than those we considered in connection with the war memoirists in Chapter 1, they are not so by virtue of *metaphysical* hiddenness. Some people can be open to us, others hidden (and many on the broad continuum stretching between those poles), but those qualities do not, in life and in autobiographical self-understanding, correspond to the metaphysical categories of inner and outer. Real concealment in human behavior, as the studies of the

---

[46] Or they may prove enigmatic as a result of a pattern of interpretive data—a mosaic of assembled particularities—that is itself consistently interpretively ambiguous, i.e. the mosaic, the picture formed by connecting the dots of particulars, is itself ambiguous. Henry James provides an exemplary case; see 'The Lesson of the Master', discussed in my *Meaning and Interpretation*, 104–29.

phenomenon in literature and film amply show, is not a matter of metaphysical privacy, and our descriptions of persons as open or hidden or enigmatic do indeed hold meaning, but not of the kind required by the dualistic introspectionist picture.

Speaking of the claims of the metaphysical voice, recall, as we saw in the discussion of Ellen Glasgow in connection with privacy in Chapter 1, Section 3 above, that Wittgenstein said in *Last Writings on the Philosophy of Psychology*, volume i, §974, 'Nothing is hidden here; and if I were to assume that there is something hidden the knowledge of this hidden thing would be of no interest.' That remark does go well beyond what we are discussing presently into his removal of the private object as the meaning-determining referent within the context of the private-language argument, but the tenor is for our purposes clear enough. And the aspect of that remark most germane to the present issue was shown in the following remark in the ordinary voice: 'But I can hide my thoughts from somebody by hiding my diary. And in this case I'm hiding something that might interest him.' We may well encounter people who find us enigmatic — to reveal more of the significance of this remark too by recontextualizing it here — but to 'say that my thoughts are inaccessible to him because they take place within my mind is a pleonasm' (§975). So descriptions, properly understood, are possible, just as is a kind of introspection conducted as self-reflection — but only where 'self' is not misconstrued. Consider then, but now in this light, Wittgenstein's example from *Philosophical Investigations* §585:

When someone says 'I hope he'll come' — is this a *report* about his state of mind, or a *manifestation* of his hope? — I can, for example, say it to myself. And surely I am not giving myself a report. It may be a sigh; but it need not. If I tell someone 'I can't keep my mind on my work today; I keep on thinking of his coming' — *this* will be called a description of my state of mind.

The words uttered with a sigh or with, say, a stern and threatening tone are in truth phrases with different meanings, they are moves in different games.[47] Yet they *look* the same. One may be a manifestation of hope; the other may also be a manifestation, but perhaps of anger.

Predictably, meaning is not stable — as indeed it would be if such words and phrases uniformly *were* reports on, or descriptions corresponding to, inner states constituting inward referents. In §586 Wittgenstein thus says: 'The exclamation "I'm longing to see him!"

[47] See *Philosophical Investigations* II. ix, pp. 187–8.

may be called an act of expecting. But I can utter the same words as the result of self-observation, and then they might mean: "So, after all that has happened, I am still longing to see him." ' The most telling remark Wittgenstein makes next: 'The point is: what led up to these words?' What led up shows the language-game, the context of usage, indeed nothing less than the meaning of the phrase. And the investigations we make in biographical contexts, often in 'What-did-you-really-mean-when-you-said . . .?' form, do not take us into the private inner realm, they take us into what we might well rightly, ordinarily, call the private life of a person, where what is and is not classified as *private* will be contextually determined as well, and where, as we have seen Wittgenstein argue repeatedly from different angles, the private will never exemplify the metaphysical sense of 'privacy' upon which the mistaken conceptions of the self rely. Asking 'What led up to these words?' is one particular way of manifesting the concern expressed in his earlier claim 'An "inner process" stands in need of outward criteria' (*Philosophical Investigations* §580).

For the human understanding, the comprehension, of such private— such *sensibly* private—matters, we need to consult outward criteria. And that sense of the understanding of the private is, again, true of the first-person case just as it is of the third-person case; it is the positive form of introspection we need to undertake to gain self-knowledge. Or at least such introspection, or contextualized self-reflection and reconsideration, will *sometimes* be what is called for in gaining self-knowledge; it is certainly not, in a precontextual way, a *rule* of first-person knowledge. 'Does it make sense to ask "How do you know that you believe?"—and is the answer: "I know it by introspection?" In *some* cases it will be possible to say some such thing, in most not.' Reminding us of the context in which true introspection operates, Wittgenstein adds: 'It makes sense to ask: "Do I really love her, or am I only pretending to myself?" and the process of introspection is the calling up of memories; of imagined possible situations, and of the feelings that one would have if . . .' (*Philosophical Investigations* §587). And in a different part of the book, Wittgenstein, following his question 'Are the words "I am afraid" a description of a state of mind?' (*Philosophical Investigations* II. ix, p. 187), writes: 'We ask "What does 'I am frightened' really mean, what am I referring to when I say it?" And of course we find no answer, or one that is inadequate.' Again it is the final observation that is the most telling: he adds: 'The question is: "In what sort of context does it occur?"' (p. 188). No answer will prove satisfying if the concept of reference ('what am I referring to?') enforces the

metaphysical pictures, nor will an answer prove satisfying if we ask
what the words 'I am frightened' *really* mean, in a way presupposing
that the meanings—what we are asking for as the meaning, what we
generically want—of the words are not only separable from, but fixed
prior to, the occasion of their use. 'Let the use of words teach you their
meaning' (*Philosophical Investigations* II. xi, p. 220). And, again, it is
the uses of words within dramatic and literary contexts that allow us
the previously described reflective distance, the relatively detached point
of view outside the bounds of the unfolding language-game. With *this*
aesthetic distance (articulated not in the traditional way but rather in
terms of the witnessing of the relevant psychological events and actions
in the stream of life of the represented fictional world), the spectator is
enabled, with the perceptual scrutiny of the connoisseur, to attend not
only to the subtleties of the linguistic action as we have seen above, but
now also, indeed, to be taught (or shown, in such a way that we can learn
them more accurately, more exactingly)—precisely in Wittgenstein's
sense—the meanings of the words. All of that contributes to the
positive conception of introspection and the positive (or noninferential)
conception of third-person understanding; those conceptions are anti-
reductionist and wholly independent from their dualistic metaphysical
employment as driven by conceptual models, by metaphysical pictures.[48]

There is then a clear sense in which this aspect of Wittgenstein's work
provides a substantially more positive understanding of the self, and
self-knowledge, and it can be added to the earlier remarks on avowals
as part of the attempt to answer the question asking whether there *is* a
positive conception of the self that Wittgenstein develops over the course
of his work following his early observation on the deeply mysterious
nature of the 'I'. But can these positive points be summarily presented
as an overarching *theory* of the self? The answer, of course, is an anti-
reductionist No,[49] which is necessitated by all that Wittgenstein has said

[48] As we have seen, in part, in Ch. 2, sect. 3, a sustained example of precisely this kind
of introspection, carried out with an equally sustained mindfulness of the significance of
Wittgenstein's philosophy for any autobiographical endeavor, is found in Cavell, *A Pitch
of Philosophy: Autobiographical Exercises* (Cambridge, Mass.: Harvard University Press,
1994).

[49] One might ask why an affirmative answer cannot be given, characterizing Wittgen-
stein's theory of the self in general, and introspection in particular, as the 'contextual
theory'? Such a characterization of the view developed would be both (1) inapt—in
that a theory is generally regarded as unifying into a single comprehensive conceptual
assemblage a large and diverse collection of seemingly disparate particulars, whereas
the general and metaphysically confused conception of the self and of introspection is

along the way of these 'sketches of landscapes'. Indeed, Wittgenstein is engaged in the philosophical struggle of clarifying the 'logic', the philosophical grammar, of the mental concepts we employ in our language-games of self-understanding. Rather than positing and then theoretically seeking a hidden explanandum of selfhood, he is assembling an 'album', a mosaic, of fully clarified parts of that understanding—in short, a perspicuous overview. Thus, one can reread all of Wittgenstein's remarks pertaining either directly or indirectly to the self (and that is a difficult line to draw, since conceptions of linguistic meaning are inextricably involved with conceptions of the mind and self), or one can, to employ an image from the early philosophy in reference to the later, kick the ladder away once one has climbed it, freeing oneself of the general theoretical questions that, as we have seen, both are born of, and then further engender, misconceptions.

If the misconceptions are removed, what then lies open to view? The multiform human practices and activities that involve speaking of or acting in connection with the self, problems of self-knowledge, of self-understanding, the experienced asymmetry between the first- and third-person cases, the *real* questions of understanding other minds, of knowing another's thoughts and feelings, and so forth. Those very practices are both the engagements of life and—often—the content of literature, theater, film, and most centrally autobiography. We contextually see situated selves depicted—often with great subtlety and a fine attentiveness—in Greek tragedy, in Homer, in Virgil, in Dante, and—with a distinctive originality—in Petrarch, in Milton, in Shakespeare, in Proust, in Dostoevsky, in Tolstoy, in Kafka, in George Eliot, in Henry James, in Woolf, and so forth through lists extending far into theater and film that we would not want to try to complete. What we would learn from looking to those works of art with the fundamental philosophical question concerning the nature of the self

---

clarified out of existence by attending to actually disparate particular cases; and (2) deeply misleading—in that the impulse to construct theories was precisely the habit of mind with which Wittgenstein was engaged in constant battle. A closer look at Wittgenstein's later methods (at least), and most notably his conception of philosophical progress, would show how prismatically distorted is the interpretation of Wittgenstein as a propounder of theories (or anti-theory theories) on traditional philosophical issues. See *Philosophical Investigations* §§109–33, and particularly in connection with the confusions inherent in approaching these questions with psychological theory, II. xiv, p. 232. For an exceptionally fine introduction to his later methods and the philosophical spirit of the investigative enterprise, see Marie McGinn, *Wittgenstein and the Philosophical Investigations* (London: Routledge, 1997).

and the correlated understanding of self-knowledge in mind would not
be worth learning if it could be encapsulated.[50] We can similarly learn
from Augustine in his *Confessions* (as we shall see in Chapters 4 and 5),
from Rousseau, from Simone Weil, from countless others writing in the
same genre, and from close examinations—indeed close readings—of
self-portraiture also. But as I have indicated above, the *exact* terms in
which Wittgenstein's philosophy of psychology will be helpful to our
understanding of the arts of self-representation (counting autobiography
as central to these), and the ways in which those works of art will be
conversely philosophically helpful, will be describable only *in situ*, only
in particular (and hence a *theory* of the relation is, again, instructively
unavailable).

There are cases in which the arts show in tangible form what phil-
osophy does, or would, claim; for example, Nabokov's autobiographical
Speak, Memory shows the direct and powerful significance of memories
(in Russia) upon later perceptions (in America), and that in turn shows
the relational, nonisolable nature of mental phenomena as briefly intro-
duced earlier and as we will see in greater detail below. There are cases in
which the arts provide a corrective, or again an antidote, to false or con-
fused philosophical notions; for example, a close study of Rembrandt's
great series of self-portraits will pointedly and powerfully remind us that
the acquisition of experience is not hidden in the phenomenological
interior, but is seeable—indeed knowable—in the physiognomy of the
face. There are cases in which the contents of another's mind are perhaps
more forcefully expressible in a (seemingly) indirect way, e.g. through
painting the self, rather than in (seemingly) directly describing it; e.g. Pi-
casso's final, disturbing, death-obsessed, and fearful self-portrait. There
are, as has been much discussed in recent years, cases in which the
works of art become themselves a medium of philosophical thought,
e.g. Duchamp, Warhol, Cage, in such a way that the work itself *is* the
aesthetic thought and not the end-product subsequent to it—which in
turn teaches us a philosophical lesson about the anti-private ontology of
thought. There are cases in which either (1) the arts themselves would
engender or reinforce a conceptually confused construal of the self,
or (2) our metaphysically confused interpretation of the work of art is

---

[50] In relation to the claim that the very ability to encapsulate knowledge of that
kind is a symptom of its diminished or debased value, see Henry James's 'The Figure
in the Carpet'; I discuss the tale in those terms in *Meaning and Interpretation*, 'Against
Reductionism'.

allowed to sustain itself by the very content of the work; examples of each category might be found in Cindy Sherman's photographs, or Philip Roth's autobiographical fiction, or in late Godard (all reinforcing the notion that the inner is literally internal, hidden, and inaccessible to the outer). In such cases Wittgenstein's philosophy of psychology can play the role within aesthetic interpretation that it plays within philosophy itself: it reorients our thinking away from entrenched misconception and toward conceptual clarification. This, at its strongest, can lead to new interpretive stances towards the art, e.g. seeing Sherman's photographs as self-revelatory essays on the ineliminably *situated*, circumstantially placed, and hence relational nature of the self. And, of course, as is clear from much of the foregoing, there are countless cases in which the arts of self-representation, in many genres or art forms, provide case-evidence that keeps us mindful of variation (and thus wary of facile generalization) and that thus cultivates the acuity of our perception of self-expressive nuance.

One case, a work that is at once a philosophical and literary work, brings together the issues of self-deception, introspection, hiddenness, and self-discovery. In Kierkegaard's 'Diary of the Seducer' from *Either/Or*, we are given a portrayal of the first of Kierkegaard's three modes of life (the aesthetic, the ethical, and the religious). In it the seducer compares the aesthetic mode favorably to the ethical: in the former everything is 'light, beautiful, transitory'; in the latter everything becomes 'harsh, angular, infinitely boring'.[51] But as we see as the diary unfolds, and as the judge, the exemplar and defender of the ethical, writes to the seducer, it emerges that whether he lives in sensuous immediacy or in a deviously calculating way, he shields himself from the truth that his life represents emptiness and meaninglessness. And his self-deception is precisely of the kind we have discussed above: it is not so much that he knows P while believing not-P, but rather that he assembles a picture of himself, emerging from assembled particulars that have been selectively chosen from his autobiographical memory, that portrays himself as the happy master of duplicity, manipulative moral ambiguity, posturing, and smug superiority to those lesser seducers who employ cruder methods and whose machinations show less finesse.

The judge, however, offers a forceful description of the seducer's life, also assembled from particulars of his behavior, his thoughts and deeds,

---

[51] In *A Kierkegaard Anthology*, ed. Robert Bretall (Princeton: Princeton University Press, 1946); this passage, p. 58.

entirely at odds with the self-concept of the seducer. He writes, 'Life is a masquerade, you explain, and for you this is inexhaustible material for amusement; and so far, no one has succeeded in knowing you; for every revelation you make is always an illusion . . .',[52] and he goes on to proclaim, 'In fact you are nothing; you are merely a relation to others . . .', and asks, 'Do you not know that there comes a midnight hour when every one has to throw off his mask? Do you believe that life will always let itself be mocked? Do you think you can slip away a little before midnight in order to avoid this?' This warning, this call to high seriousness from diversionary frivolity,[53] is itself something the seducer has already shielded himself from with his foregoing disparaging words about the ethical mode; the shield is not one proposition in one part of the mind suppressing another incompatible proposition, but rather the result of an established pattern of self-deception as willful blindness to the collection of actions and related thoughts in one's life that would paint the contrasting picture of the self.

Self-discovery is thus also represented by Kierkegaard in terms perfectly consistent with real interpretation: the judge writes: 'I have seen men in real life who so long deceived others that at last their true nature could not reveal itself.' And with this warning in place, the judge goes on to speak of the danger—as a result of paying drastically insufficient attention to the emergent patterns and moral trajectories of one's past and the resultant assembled mosaic that active and highly particularized reflection yields—of the seducer's self fragmenting into a 'multiplicity'[54] and thereby losing 'the inmost and holiest thing at all in a man, the unifying power of personality'. Self-discovery is the gradually emergent result of meticulous and full or nonselective attention; the judge writes (also in a manner strikingly consistent with the Wittgensteinian position) that in every person 'there is something which to a certain degree prevents him from becoming perfectly transparent to himself'. The introspective transparency, presumed on the traditional dualistic or Cartesian conception of introspection, is a myth; real introspection, of

---

[52] *A Kierkegaard Anthology*, 99; this and following quotations.

[53] It is, incidentally, within Kierkegaard's portrayal of this frivolity that we find a perfect analogue to the state of full imaginative involvement, yet safe detachment or aesthetic distance, that the arts allow us and that, as described above, give us the contemplative space to derive the philosophical significance contained within the arts of self-representation. The seducer writes: 'Today I have written a love letter for a third party. I am always happy to do this. In the first place it is always interesting to enter into a situation so vividly, and yet in all possible comfort' (ibid. 59).

[54] Ibid. 100–2; this and following quotations.

the kind articulated above and shown in Kierkegaard, is what is required to 'win what is the chief thing in life—win yourself, acquire your own self'. And throughout his diary, one finds passages that hover on the line separating self-deception from self-discovery: reflecting upon the fleeting nature of physical beauty, the seducer writes: 'I could become quite melancholy over this thought, and yet it is no concern of mine. Enjoy, do not talk. The people who make a business of such deliberations do not generally enjoy . . .'.[55]

And hiddenness is portrayed—indeed investigated—by Kierkegaard as well. One place human hiddenness manifests itself is in the seducer's interactions with the woman to whom he becomes engaged. First, he notices that engagement (for him a commitment made to be broken) suits his purposes perfectly, in that it has the tone of the ethical, the weighty, but 'it does not have ethical reality in the stricter sense, as marriage does'.[56] He plots from the very beginning that he will manage things in such a way that it will ultimately be she who breaks the engagement, and proceeds to reflect on his skill as an aesthete (in Kierkegaard's sense). But the hiddenness of his motives, his plans, indeed his character, is most assuredly not of the metaphysical kind corresponding to the pre-Wittgensteinian view of the self. It is, instead, shown perfectly by Kierkegaard throughout the diary, and described perfectly by the seducer within it: he writes, 'Someone has said that it takes a little more than honesty to get through the world. I should say that it takes something more than honesty to love such a girl. That more I have—it is duplicity.'[57] What is hidden; as is shown here, is not *metaphysically* hidden; it is, one might say, of the same ontological *kind* as that which is overtly revealed to his fiancée, Cordelia, but carefully kept from her. The traditional picture of selfhood would itself, as we have seen, seduce us into a misconstrual of hiddenness; the situating of the concept into a fully imagined context of usage, as Wittgenstein advises and as Kierkegaard has done, rightly construes hiddenness as a matter of *duplicity*, and not of traditional dualism of the kind we considered in Chapters 1 and 2. Just as the logic of first-person predicates is better shown (if in a dangerously general way) by characterizing them as avowals rather than descriptions (as metaphysically construed), so the logic, the grammar, of hiddenness is better shown by characterizing it as duplicity rather than ontologically guaranteed concealment. And all of the mental phenomena (so easily

[55] Ibid. 76.   [56] Ibid. 58.   [57] Ibid. 64.

miscast, as Wittgenstein has shown) at work in Kierkegaard's text, i.e. self-deception, self-discovery, and hiddenness, both contribute to and buttress the notion of real introspection—indeed one could plausibly argue that this is the fundamental concept of Kierkegaard's text as it is both represented and, in light of the deeply autobiographical nature of Kierkegaard's writings, enacted.[58]

There is perhaps one way, however, of generalizing—although it will inevitably miss greater things than it captures—concerning the insight we might gain from the preceding Wittgensteinian philosophical–critical undertakings. It may fairly be taken to be clear at this point that we are inclined, owing to a cohort of interrelated misleading philosophical pictures, to envision the meaning of a word as a ghostly, hidden, fixed inner referent. Seeing the expansive—and no doubt expanding—vocabulary of the self in context could free us of that blinding misconception. That would be one hardly negligible benefit. But we are also inclined to see, indeed, the *self* in terms much like that picture of a word's meaning. *Last Writings on the Philosophy of Psychology*, volume i, closes with the observation (§979): 'The idea of the human soul, which one either sees or doesn't see, is very similar to the idea of the meaning of a word, which stands next to the word, whether as a process or an object.' Both conceptions—and they are tandem conceptions—are philosophical myths, but to say that is again most assuredly *not* to claim that the self and meaning do not *exist*: that would be to meet an assertion with its extreme antithesis, an antithesis that shares its misleading preconceptions. It *is* to say that *those* construals are wrong, misled, generated by the tricks language repeatedly plays on us. Other conceptions, like the expansive ordinary or grounded conception of introspection in place of its metaphysical counterpart, or like avowals in place of descriptions, or like contextualized, ordinary descriptions in place of the philosophical sense of the term, although difficult to capture succinctly, are what can genuinely illuminate. Such illumination just is the result of aesthetic investigations into the arts

---

[58] This is the kind of general pronouncement concerning the content of a piece of literature that holds philosophical significance (Kierkegaard's texts are this and the reverse, depending on the aspect upon which one centrally focuses) that calls for support with much more detailed and particularized passage-by-passage investigation. In a fine brief study, Patrick Gardiner rightly claims: 'All in all, Kierkegaard's analysis of aestheticism is conducted with a psychological subtlety and an elaborate attention to detail that defy brief summary' (*Kierkegaard* (Oxford: Oxford University Press, 1988), 47). Precisely the same is true of Wittgenstein's analyses of self-knowledge, self-awareness, and self-understanding throughout his writings on mind and language.

that depict the self. For those reasons the achievement of conceptual clarification, on (1) the misconceptions of person-perception articulated in behaviorism, (2) the proper construal of first-person expressive speech, and (3) the real or positive nature of introspection, are all of the first importance for our understanding of the epistemological power of autobiographical writing and the arts of self-representation. And that clarification, the result of a Wittgensteinian critique of the theory of the self and its relation to language, theory that manifests on one side as metaphysical claims about subjectivity (Cartesian introspection), and on the other as scientific psychology (behaviorism), preserves a secure place for nonreductive philosophical humanism in the study of human perception, interaction, expression, and self-understanding. Moreover, it shows how Wittgenstein's philosophy itself holds meaning through the way in which it is *used*, i.e. as a form of conceptually clarifying therapy that dissolves perplexity rather than adding to theory of one form or another.

Beyond that, the positive conception of introspection (obscured by both Cartesian and behavioristic views of the self and not explicitly clear from Wittgenstein's own writings) allows us to appreciate the value of some distinctive artistic acts of self-representation, a value we see much more clearly with a perspicuous grasp of Wittgenstein's later philosophy of psychology. Some artists, particularly good at exploring the contextually situated self and showing the indissoluble interconnections of utterance, thought, behavior, and context, are themselves exemplars of the human capacity to introspect; they show what introspection actually is within the artistic process of representing it.

Perhaps Wittgenstein's writings on the self, a subject to which he returned over many years, exhibit their own double-aspect. In one sense, it is a frustratingly unfinished task—and a fundamental one—that leaves us without a developed account: we have only a few positive remarks surrounded by massively powerful but *negative* considerations—we are shown largely what not to think. But in another sense, perhaps appearances are here again deceiving, perhaps what we are left with is more positive than may initially appear. For as the Preface to *Philosophical Investigations*, as we have seen, makes clear, there is a 'wide field of thought' that is crisscrossed in every direction, and by now it is beginning to appear that much more is directly relevant to our understanding of the self than we might at first have believed. Furthermore, perhaps the correct view of the self requires the achievement of a philosophical overview just as is required for other

problematic philosophical concepts like knowledge, meaning, certainty, verification, proof, criteria, understanding, and so forth. If so, perhaps Wittgenstein has left not so much a lacuna as an intelligent, Tractarian silence on a subject that must be seen during what he called in that Preface 'long and involved journeyings'. And as to what lies open to view before us, particularly in the multiform artistic representations of selfhood and first-person understanding, we can now see that such things *are* open to view only if our vision is not prismatically distorted by theories, i.e. by metaphysical pictures resident in our language, by misleading analogies, by false dichotomies, by false conceptual models. All of those would cut short those journeyings, thus precluding philosophical–critical perspicuity, and in that sense Wittgenstein's negative definition is—if in a distinctive sense—transmuted into a positive one: it grants the self the conceptual freedom,[59] the ability to think about ourselves and each other in a way unburdened by metaphysically motivated misconceptions, that is prerequisite to the fulfillment of the Socratic injunction.

With these reflections behind us, how then do we form a conception of the self caught in the very act of self-reflection? And what of the *image* that the mind all-too-readily creates and as quickly endorses of its own self-investigation?

---

[59] Here we see at least something of the conceptual similarities or affinities between Wittgenstein's methods and psychoanalysis—but there are real and certain limits to the comparison. See John Wisdom's foundational writings on this linkage in *Philosophy and Psycho-analysis* (Oxford: Blackwell, 1953), esp. 169–81; Jacques Bouveresse, *Wittgenstein Reads Freud* (Princeton: Princeton University Press, 1995); and Frank Cioffi, *Wittgenstein on Freud and Frazer* (Cambridge: Cambridge University Press, 1998). See also Wittgenstein's *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, ed. C. Barrett (Berkeley: University of California Press, 1972), 41–52.

4

The Self, Thinking

Despite all that Wittgenstein has said against what is generally characterized as mentalism—in essence the view that the self is fundamentally a Cartesian point of consciousness, in which a full catalogue of mental events and acts takes place—it is difficult to relinquish the idea that the self is, whatever else it may not be, the inner repository of *thinking*.[1] Indeed it is not clear that one should even *want* to relinquish this conception, since the alternative of a reductive behaviorism seems forever looming in a self-threatening way. But Wittgenstein, as we have seen (in Chapter 3, Section 1), is no behaviorist; nor is he a behaviorist in disguise.[2] So for Wittgenstein, at least, there is a way of relinquishing the deeply rooted conception of the self as the locus of hidden, inner thoughts, yet without eradicating the undeniable asymmetry between the first- and third-person cases and without thereby sacrificing the whole idea of the self to behavioristic explanation. What Wittgenstein calls for, of course, is a wide and deep unearthing of another pernicious misconception that drives one way of thinking of the self—indeed one way of thinking of thinking. And this matter is, as we shall see, far from unrelated to our understanding of autobiographical knowledge.

1. IMAGINING THOUGHT

It is in a sense natural, natural to *philosophical* thinking, or thinking in a metaphysical voice, to construe thought alone as the most private

---

[1] In its modern formulation this view of the self as the inner thinking thing has been most influentially, and perhaps most clearly, articulated by Descartes; see *Meditations on First Philosophy*, esp. 'Second Meditation (The Nature of the Human Mind, and How It Is Better Known than the Body)', in *The Philosophical Writings of Descartes*, ii, trans. J. Cottingham, R. Stoothoff, and D. Murdoch (Cambridge: Cambridge University Press, 1984), 16–23.

[2] As we saw above, in *Philosophical Investigations* §307, Wittgenstein's imagined interlocutor queries: 'Are you not really a behaviourist in disguise? Aren't you at bottom really saying that everything except human behaviour is a fiction?'

element in human experience. The very thought that we alone really
know our own thoughts—the idea that initiates other-minds skepticism
and that can lead to the epistemological self-deprivation called solip-
sism—is an attempt to encapsulate a wide range of experience into a
single proposition, i.e. just that the self is fundamentally an inward,
metaphysically private thinking essence contained within its own inner
world of pre-linguistic, pre-expressive pure cognition, and thus only
contingently related to the external, social, public world. That wide
range of experience, as literature as well as the other arts show us often
more clearly and always more fully than does philosophy, includes
multifarious human engagements such as thinking to ourselves, keeping
our thoughts to ourselves, keeping our own counsel, and on the other
hand letting another know our thoughts, revealing our thoughts to
another, telling secrets that are to be kept by the person to whom we
tell them, and so forth through a vast index of thoughts that are, on the
misleading philosophical model or picture that we have now considered
in a number of its manifestations, either kept in or, again, contingently,
let out. One employs the word 'contingently' here, precisely because
thoughts are indeed thought to be essentially and invariably private in
the first instance, and only revealed, or expressed, or told, or shown, and
so forth (here again there is a vast index that is also more thoroughly
investigated by literature and the arts than by philosophy) as a secondary
matter. The attempt to encapsulate this vast range of human experience
in that one proposition, again that the self just is a private think-
ing thing, however, fails. As Wittgenstein's investigation into thinking
shows, like many other philosophical errors he subjects to scrutiny: (1) it
misleadingly attempts to impose uniformity on great diversity; in doing
so (2) it mischaracterizes thinking, replacing the many things it is with
a philosophical *picture* of what it is; and (3) it elevates one particular
aspect of thinking—the aspect of its frequently inward nature—to a
defining principle of all thought. And this elevation of what one might
call a particular to what one will then call a universal thus precludes
the attainment of what Wittgenstein calls a 'perspicuous overview' of
the concept of thinking. And of course, if we want to understand the
contribution the actual facts of thinking make to our conception of the
self, we must first see those facts clearly, and not obscure our vision of
them by listening to the mischaracterizations of the metaphysical voice.

In *Philosophical Investigations* §316 we find this assertion made by the
errant interlocutor: 'In order to get clear about the meaning of the word
"think" we watch ourselves while we think; what we observe would

be what the word means!' Naturally, from what we already know of Wittgenstein's arguments against the Cartesian and Jamesian conceptions of introspection, we can expect that he will oppose this idea. And the expectation is very quickly fulfilled: he replies, '—But this concept is not used like that.'³ The concept of thinking has its many and diverse uses, and to understand it, indeed to understand its grammar—a point to which we will return below—we need to investigate its uses in its contexts. But the impulse to believe that we should after all be able to get clear—indeed *transparently* and with *direct* or unmediated inner access to the phenomena—about the meaning of 'thinking' by looking at what the word 'thinking' must refer to is nevertheless difficult to ignore. The impulse is fueled by two misconceptions, each given lengthy and detailed dissections in various parts of *Philosophical Investigations* and beyond:⁴ the first misconception, familiar to us from the previous chapter but emerging now in both a new guise and a new context, is the belief on a broader scale that a word invariably gets its meaning through reference—it means what it refers to; the second misconception, on a narrower scale and for us new, is that the closest scrutiny of a phenomenologically isolated single instance of thinking will give us the meaning of the word 'thinking'. And it is this latter misconception that Wittgenstein addresses in his next remark: he adds, parenthetically, that attempting to divine the meaning of 'thinking' in that way would be analogous to attempting to make out the meaning of the word 'mate' in chess if we—without knowing how to play—were to observe closely the last move of a game.⁵

---

³ Wittgenstein's advice to consult usage as a means to the end of grasping meaning is, despite the frequent reduction of this point to a slogan, hardly a simple, unitary, or even straightforward matter. In addition to the discussion of the genuine (i.e. use-grounded) conception of introspection in Ch. 3, I attempt to cast some light on meaning–use connections throughout *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994).

⁴ This first misconception is among the most sustained topics running through *Philosophical Investigations*; the second, as a psychologized version of the reference theory of meaning, naturally connects directly to the private-language argument. For a discussion of that argument's significance for artistic meaning, see *Art as Language*: Wittgenstein, Meaning, and Aesthetic Theory (Ithaca, NY: Cornell University Press, 1995), ch. 6. We will return to the matter in the third section of this chapter.

⁵ Although it is true that Wittgenstein stands in opposition to (or more accurately, undercuts) William James's conception of introspection, there are nevertheless many striking affinities between Wittgenstein's philosophy and pragmatism. I point to a number of those in 'Contours of Experience: The Foundations of Dewey's Aesthetic Thought', in J. Conant (ed.), *The Cambridge Companion to Dewey* (forthcoming). In the present instance, the anti-atomistic point concerning the patent absurdity of trying to learn the meaning of the word 'checkmate' by closely observing only

   Such an effort on the chessboard, it is clear at a glance, would constitute a truly mind-numbing exercise in futility. But there are slightly less obvious dimensions of Wittgenstein's remark that should be brought out. In Wittgenstein's writing any analogy to a game should put us on the alert; the analogy here has two strands of special significance. The first strand is the connection to the concept of a language-game, chess itself being the real-game counterpart to its linguistic analogies, as discussed in the first parts of the *Philosophical Investigations* and elsewhere;[6] the fundamental point is that words function as tools within circumscribed contexts, and that their meaning, again, will be determined not by a ghostly mental component attached to the otherwise lifeless outward sign, but rather by the role the word takes within the context, within the game. As I have suggested elsewhere, here too literary works, not surprisingly, often show a good deal more of that, and more of the nuances of significance, than direct treatments of those topics in the philosophy of language.[7] The second strand is the resonance with the example of games that Wittgenstein, and now many others following him, have employed in discussing the problem of universals; the fundamental point here is that the actual class of all games exhibits vastly greater diversity (and most significantly, no single property in common that justifies their inclusions within the class of games) than we expect when coming to them with metaphysical expectations concerning classificatory uniformity in mind. Wittgenstein, at various locations in his far-reaching investigations, showed both strands to be informatively true of the concept of thinking as well.

---

the final move in chess is deeply compatible with James's writings concerning the (anti-atomistic, anti-Lockean) stream of thought in which the very idea of a *discrete* experience is called into question; see 'The Stream of Thought', in J. Stuhr (ed.), *Classical American Philosophy* (New York: Oxford University Press, 1987), excerpted from *The Works of William James: The Principles of Psychology*, ed. F. Burkhardt, 3 vols (Cambridge, Mass.: Harvard University Press, 1981). I will return to this issue in Ch. 6, Sect. 3, below.

   [6] For Wittgenstein's discussion of the language-game, see *The Blue and Brown Books* (Oxford: Blackwell, 1958), esp. the discussions on pp. 17, 25, 81, 108, and 172, and *Philosophical Investigations*. I attempt to offer an elucidation of the concept, following Wittgenstein's various employments of the phrase, in *Meaning and Interpretation*, 9–44.

   [7] I have pursued the relevance of literary cases for issues in the philosophy of language, and most centrally the attainment of a perspicuous overview of the multiform concept of meaning, in 'Dencombe's Final Moments: A Microcosm of Jamesian Philosophy', and in the interpretation of Henry James's 'The Author of Beltraffio', 'The Lesson of the Master', 'The Figure in the Carpet', and 'The Tree of Knowledge', in *Meaning and Interpretation*.

Having said that it displays a fundamental misunderstanding if we try to clarify the philosophical problem about sensation by studying the headache we may have presently (*Philosophical Investigations* §314), he marks the misleading parallel between the assertions that (1) the expression of pain is a cry, and (2) the expression of a thought is a proposition (*Philosophical Investigations* §317); they are in truth disanalogous, because the purpose of the proposition is misconstrued if we see its purpose as 'conveying to one person how it is with another; only, so to speak, in his thinking part and not in his stomach'. It is perhaps worth noting in passing that this misleading analogy can exert a remarkably pernicious influence on our thinking about expression, or more particularly the putting across of thoughts, in the arts. If the work of art is construed as one kind of, or on the model of, a proposition,[8] then the belief that the content of the 'proposition'—the artwork—is metaphysically hidden and can be inferred only indirectly through a process of inductive criticism (a process that runs perfectly parallel to the alleged inductive process of getting to know the contents of the mind of another within philosophical discussions of the overcoming of other-minds skepticism) will come naturally to us, and indeed seem inevitable. But that is another (if closely related) story; the present purpose is to consider Wittgenstein's remarks on thinking as they pertain to the conception of the self, and it is with the disanalogy between the two cases—(1) the expression of pain and (2) the expression of thought—that it begins to become apparent that here too is yet another site of battle, upon which Wittgenstein will engage dualism as it would misleadingly shape our thought about thinking.

The very idea of suddenly understanding is one that possesses the power instantly to revitalize our conception of the hidden, Cartesian self; it is, we are inclined to think, a process that occurs inwardly, and indeed that occurs only within the metaphysical confines of our psychological interior. And that is precisely the kind of phenomenon the nature of which, so we too easily believe, verifies the conceptual picture of the Cartesian interior. But the question 'What happens when a man suddenly understands?' (*Philosophical Investigations* §321) is, Wittgenstein insists, a 'badly framed' question; most importantly,

---

[8] The modeling of the work of art on varying conceptions of the proposition has very likely generated more philosophical confusion in aesthetics than any other single source; I attempt to disentangle a number of these threads, and then to show how a Wittgensteinian conception of language might shed light on artistic meaning, throughout *Art as Language*.

against the revival of the Cartesian conception, he insists that the answer
to that question will *not* unfold in the terms implied by the question
as framed, that is, it will not unfold in terms of pointing 'to a process
that we give this name to'. Nor, he adds, will the question be answered
in the vocabulary antithetical to the Cartesian picture, i.e. behaviorism;
the specific facial movements, sudden alterations in breathing patterns,
and so forth are 'not answered by such a description' (*Philosophical
Investigations* §322). And the result of this, Wittgenstein suggests, is
usually to posit the mysterious existence of an inner experience or
occurrence that is indefinable yet available to introspective scrutiny
if not to verbal articulation: 'this misleads us into concluding that
understanding is a specific indefinable experience'. That premature
and conceptually confused conclusion blocks the way to what we
should actually undertake in order to obtain a satisfying answer to the
question asking what happened in suddenly understanding; we should,
Wittgenstein suggests, ask 'how do we *compare* these experiences', and
'what criterion of identity *do we fix* for their occurrence?' (*Philosophical
Investigations* §322). And that kind of undertaking is of a kind, of
course, that the twin beliefs seen just above, that is, that such words and
phrases mean through referring (in this case to inner mental objects or
processes) and that a close introspective scrutiny of them (or it) will tell
us what we need for the answer, would wholly preclude. Here again,
literary cases could contribute enormously to such an undertaking, *if*
we first remove the obstacle given in the form of misleading analogies,
presumptions concerning meaning and reference (specifically, again,
that all words function like names), and seemingly obvious but in
truth deeply erroneous beliefs concerning the epistemological value of
introspective scrutiny as motivated by dualistic metaphysical pictures.

   Yet why have we been given, in the reflections on thinking thus far,
again only *negative* considerations, i.e. why have we been told only what
not to think? The frustration we may feel here, coupled with a desire
to have a good Aristotelian definition of the subject matter in hand
before we proceed any further, is expressed in *Philosophical Investigations*
§327: Beginning with the question (dangerously framed) 'Can one think
without speaking?', Wittgenstein adds: '—And what is *thinking*?—well,
don't you ever think?' And then giving voice to conceptual relapse, adds:
'Can't you observe yourself and see what is going on?' The misleading
sense of simplicity is here too nourished by the notion of introspective
access: 'It should be quite simple. You do not have to wait for it as
for an astronomical event and then perhaps make your observation in

a hurry.' But the relapse is treated with the ordinary (in opposition to the metaphysical) voice inquiring, reasonably, in §328: 'Well, what does one include in "thinking"? What has one learnt to use this word for?' That carries the implications that we will have our answer to the question 'what is thinking?' *if* we know where to look, and that we have before us the resources to determine what we do and do not include under the concept. Again the impulse to narrow the question to one concerning exclusively an inner process is strong, and again it is met with a reminder of our practices; Wittgenstein shortly asks, if a person takes a measurement in the middle of a train of thought, but says nothing to himself during the measuring, has he interrupted the thought? This question is designed to erode the presupposition that thinking will invariably be explicated in terms of an inner *process*; indeed the direction Wittgenstein is taking here leads toward the conclusion that intentional verbs like thinking (and willing would be another, although it, like thinking, will display its own idiosyncrasies) do not necessarily refer to, and they do not get their meaning from, phenomenological processes or inward events that take place for a specifiable duration and that are thus available as inward objects amenable to inspection.[9]

   In subsequent sections Wittgenstein makes explicit the relation between our philosophical thinking of thinking and our philosophical thinking of language, suggesting that, quite against our inclinations, meanings do not go through the head in addition to verbal expressions (adding the densely compressed remark that language itself is the vehicle of thought). He further investigates the relations between thinking and speaking, showing that a crude additive model of the kind we might embrace under the influence of dualism will never suffice,[10] and, most tellingly for present purposes, observes that while it is true that we do in special cases 'call it "thinking" to accompany a sentence by a mental process, that accompaniment is not what we mean by a "thought"'. That is to say, the additive (thought-plus-speech) model may in truth apply in some rare cases, but we are still able to, and almost always do,

---

   [9] See in this connection Wittgenstein's discussion in *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958), II. xi, pp. 217–19, where he says, memorably, 'if God had looked into our minds he would not have been able to see there whom we were speaking of', and he adds a few remarks later: 'The language-game 'I mean (or meant) this' (subsequent explanation of a word) is quite different from this one: 'I thought of . . . as I said it.' The latter is akin to "it reminded me of . . . "'. I pursue these issues in Ch. 5.
   [10] For a fuller discussion of these points in connection with artistic meaning, see *Art as Language*, ch. 5.

identify thoughts without recourse to those distinct processes (*Philosophical Investigations* §332). Further eroding the plausibility, or rather further shattering the illusion, of the necessary phenomenological interiority of thought as process, he invites us simply to imagine people who could only think aloud, just as some can only read aloud (*Philosophical Investigations* §331). Here of course we can well imagine thinking taking place, but without the separability of the inner from the outer; and if that is imaginable, then the internal, as we are impelled to conceive the matter, does not necessarily precede the external.

A very good specific example of the kind of philosophical diagnoses being undertaken here, i.e. our coming into an understanding of the way in which a philosophical picture motivates seeing the entire matter of thinking in one narrow way to the exclusion of all others either different or incompatible with it, or where the dualistic picture is rendered *seemingly* inevitable as a result not of the facts of the case but rather as a function of a particular picture-driven way of seeing the cases before us, is found in *Philosophical Investigations* §334. Here Wittgenstein contrasts (1) the way we actually use the phrase 'So you really wanted to say . . .', i.e. using it to lead someone 'from one form of expression to another', with (2) the way we are philosophically motivated to construe it, i.e. as clear evidence for the (alleged) fact that 'what he really "wanted to say", what he "meant" was already *present somewhere* in his mind even before he gave it expression'. In truth, of course, 'various kinds of things may persuade us to give up one expression and to adopt another in its place', and making such an adoption does not imply the prior mental existence of the newer expression. That is, however, all too easily forgotten if we come to these subjects with the expectation of finding neat and narrow conceptual uniformity along with the host of misleading elements previously identified.[11] The misexplanation, an explanation that is developed in strict accordance with an underlying metaphysical picture and then used circularly or in turn as evidence for the truth of that picture, is seen clearly in the next section, where the case of making an effort, in writing a letter, to find the right expression for our thoughts is discussed; there again it becomes all too easy to posit the existence of the thought as the essential criterion for the correctness of the expression. But thinking, and what we say about thinking, does

---

[11] Or if we come to it, as Judith Genova has helpfully elucidated the matter, with a way of seeing that is unitary and reductive; see her *Wittgenstein: A Way of Seeing* (New York: Routledge, 1995).

not at all correspond to that template; one can very easily identify, for example, that one has had a thought similar to what another has just said, or indeed that one 'was about to say exactly' what another has just said in an animated conversation, without the employment of, or recourse to, any such mentalistic criterion. Thinking, Wittgenstein is making out, is not what we are inclined, or easily can be inclined, to think it is.[12]

One particularly memorable way of expressing the problem that all too readily emerges is given in *Philosophical Investigations* §338, where we find the matter presented in terms of *images*: when we think about the experiences of *wanting* to say something, 'we grasp at the *image*' of speaking, just as in thinking of wanting to dance we grasp at the image of dancing. But even a second thought about the actual experience of wanting to speak shows that the image, the pre-verbal or pre-sonic mental duplicate and exact predecessor of the actual speech to follow, *need* not exist at all.[13] This imagistic way of thinking of wanting to speak—where wanting to speak is taken as a kind of paradigm of thinking—is only a simplifying, unifying myth. Indeed, what is required in order to see the misleading nature of this imagistic characterization of thinking is the positive, philosophically unobjectionable variety of introspection, i.e. a full, patient recounting of human practices. It should be no secret that literature generally, and autobiography more particularly, are conceptually irreplaceable sources of such investigations.[14]

---

[12] This inclination is precisely expressed in the words of the interlocutor in *Philosophical Investigations* §337: 'But didn't I already intend the whole construction of the sentence (for example) at its beginning? So surely it already existed in my mind before I said it out loud!' A moment later Wittgenstein's voice adds: 'But here we are constructing a misleading picture of "intending", that is, of the use of this word.'

[13] Thus, when Wittgenstein says in *Philosophical Investigations* §341, that 'speech with and without thought is to be compared with the playing of a piece of music with and without thought', he means with the *actual* contextualized practices of playing with and without thought, and not with a philosophical picture of such playing, i.e. with or without a pre-sonic mental predecessor of the playing. I have tried to lay the foundations for such a comparison in 'Music and Imagination', in *Art as Language*.

[14] One can readily imagine such philosophical-literary investigations extending vastly beyond the discussion of Augustine to follow; (real) introspection, self-reflection, thinking, self-interpretation, etc. are in this sense shown (if not said) in Dostoevsky (as we shall shortly see), in Proust, in Tolstoy, in Petrarch, in Mann, in George Eliot, and in countless others. And it should perhaps also be said that literature has already, and to a much greater extent, been investigated for its philosophical significance in connection with moral issues (in Martha Nussbaum's writings on Henry James, for example), with skepticism (in Stanley Cavell's writings on Shakespeare), with politics

Nevertheless, one may, at this stage of following Wittgenstein's thoughts on this and related topics, greatly desire a clear, succinct statement that gives a solid, unambiguous answer to the question (now adding a word to the front of the question along with an emphasis) 'So what *is* thinking?' It is in §339 that we get what initially appears to be a startlingly straightforward answer: 'Thinking is not an incorporeal process which lends life and sense to speaking, and which it would be possible to detach from speaking, rather as the Devil took the shadow of Schlemiel from the ground.' But as is plainly evident, this answer, although succinct, direct, and clear, is still negative: Wittgenstein once again[15] seems to be surrounding the implied positive definition with negative statements which together render fully explicit what thinking is not. But there is a deeper reason for dissatisfaction with this answer, and Wittgenstein, moving down to the next layer of analysis within this section, says '—But how "not an incorporeal process"? Am I acquainted with incorporeal processes, then, only thinking is not one of them?' Here we have an instance of precisely the self-critical stance, itself a manifestation of a deeply held suspicion of philosophical generalization, that one needs when confronting definitional statements in mental matters. Wittgenstein wants to call into question the very terms—or more particularly the underlying source of the impulse to employ those very terms—in which the negative assertion is framed. He thus offers a layered diagnosis of that impulse: 'No, I called the expression "an incorporeal process" to my aid in my embarrassment when I was trying to explain the meaning of the word "thinking" in a primitive way.' Embarrassment, the feeling of having no answer where it seems by all rights we should have one, motivates the grasping of the term 'incorporeal': then the still deeper-layered self-analytical realization that the explanation called for is primitive dispels the embarrassment, quells the felt need for the succinct definition, and thus allows a redirecting of effort. Displaying the freedom afforded by that sequence of self-critical

---

(in Richard Rorty's writings on Orwell), and with Kantian ethics (in Richard Eldridge's writings on Conrad and Wordsworth). See M. Nussbaum, *Love's Knowledge: Essays on Philosophy and Literature* (New York: Oxford University Press, 1990); S. Cavell, *Disowning Knowledge in Six Plays of Shakespeare* (Cambridge: Cambridge University Press, 1987); R. Rorty, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989); and R. Eldridge, *On Moral Personhood: Philosophy, Literature, Criticism, and Self-Understanding* (Chicago: University of Chicago Press, 1989).

[15] As with his investigations into the nature of the language, Wittgenstein shows that only a removal of the veils of misconstrual will afford a clear view of the practices that lie before us.

thoughts, Wittgenstein proceeds in the second half of this extraordinarily significant section to reconsider, in a manner wholly consistent with his resonant question (seen above) 'Where do we say it?', where we might actually have use for such terms. 'One might say "Thinking is an incorporeal process", however, if one were using this to distinguish the grammar of the word "think" from that of, say, the word "eat".' But here we move down to still another layer of self-critical analysis: 'Only that makes the difference between the meaning look *too slight*.' And then, as always, using analogy to locate more precisely and to illuminate the philosophical significance of the point, adds: '(It is like saying: numerals are actual, and numbers are non-actual, objects.)' The sense one gets here, in addition to the similarities in method to psychoanalysis,[16] is that the language Wittgenstein himself has been led into here is being diagnosed as a kind of rupture of the language-game, a semi-legitimate (in that its use requires repeated self-undercuttings) employment of terms as misleading as they are leading.[17] Thus, the section closes with this assertion: 'An unsuitable type of expression is a sure means of remaining in a state of confusion. It as it were bars the way out.'

The belief—often unwittingly held as a deep presumption before we initiate any inquiry into thinking—that thinking must have an essence, and that its essence must be a determinate process or occurrence that resides inside the private mental world of the thinker (in short, the self, where this self-concept is given content by the act of thinking) leads to an appeal, also seemingly natural and relatively unproblematic at first glance, to the law of excluded middle. Wittgenstein thinks his way into this matter in §352, where he identifies the appeal as 'a queer trick' that 'our thinking plays us'. We are inclined, given the underlying presuppositions, to say 'Either such an image [or process] is in his mind or not; there is no third possibility!' Wittgenstein notes that this appeal easily arises in other regions of philosophy,[18] where we quickly come

---

[16] It is, at least in the present context, helpful to bear in mind that Wittgenstein's relation to Freud was complicated: Rush Rhees reports that, upon reading Freud, Wittgenstein 'sat up in surprise', exclaiming, 'Here was someone who had something to say,' even going so far as to describe himself (although one needs a nuanced understanding of this admiring remark) as 'a disciple of Freud.' See *Wittgenstein's Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, ed. C. Barrett (Berkeley: University of California Press, 1972), 41–52. See also Ch. 3. n. 59 above.

[17] And the leading as well as the (more obvious) misleading can be equally philosophically objectionable, analogous to leading a witness in a court of law.

[18] For example, we find this in the philosophy of mathematics, where we believe in accordance with the law of excluded middle that either the decimal expansion of π yields

to believe that *any* question asked of a given subject is appropriate, and beneath this that any such question is indeed intelligible in the first place—so that the case is either like one description or like another as contained in the question. Wittgenstein in fact diagnoses such appeals in terms of 'an inability to turn our eyes from this picture', that is, the philosophical picture embedded in, and enforced by, the particular framing of the question. But the question *ought* to be, in such cases, the deeper one asking 'Does reality accord with the picture or not?', i.e. is the picture one of the numerous misconstruals that we have already considered or another of that kind? Is it primitive, does it oversimplify, does it elevate one case to a paradigm and delete all other multifarious ones from consideration, does it lead us to crave a unitary encapsulation, and does it manifest an impulse to terminologically mischaracterize in the interests of illusory conceptual neatness or indeed essence? Those questions are layered beneath the law of excluded middle on the surface, and even if we stay at that level, Wittgenstein suggests, 'all the time we *feel* that it is not so'. The philosophical picture embedded in the question, embedded in the very terms of the question and its conceptual framing, requires indeed something more analogous to psychoanalysis than to essentializing theory.

But then, we may still want to ask: is thinking nevertheless *in essence* very much like, or in fact constituted by, saying something to ourselves? Is this not the process, despite anything we might say concerning the layered self-diagnosis of the questions we feel impelled to ask, that invariably defines thinking, and that either occurs or does not? Consider Wittgenstein's closing passage of §361: 'What is it like to say something to oneself; what happens here?—how am I to explain it? Well, only as you might teach someone the meaning of the expression "to say something to oneself." And certainly we learn the meaning of that as children.' And then placing himself squarely against the belief that the ineliminable essence of thinking is a process, adds: '—Only no one is going to say that the person who teaches it to us tells us "what takes place".' Yet the impulse to succinctly define persists, and we might well say of such a case of teaching a child what it means to say something to oneself, that teaching is to *impart* the meaning to the child, getting the

the group '7777' or it does not, and that there is no third possibility (and also perhaps that 'God sees, but we don't know'). See *Philosophical Investigations* §352, where the diagnosis of the misleading power of the philosophical picture behind the question (and the insistence on there being no third possibility) is given.

child to give 'himself the correct ostensive definition', i.e. to introspect on that unmediated, directly accessible process of inward 'speaking' or inward 'pre-speaking'. 'And this', Wittgenstein concludes, 'is where our illusion is.'

Wittgenstein is calling, as we might at this point in the discussion well expect, for a sustained mindfulness of the dangers of mischaracterization and of the impulse to create and employ psychological pictures. But in the end does this caution, this heightened conceptual wariness, require a kind of quiescence? And are we here too left with answers that, despite various efforts, are ultimately only negative: are we told only what not to think about thinking? The answers to these questions are, appropriately, somewhat complicated and nonunitary. We have been told repeatedly that the way to analyze, to understand, indeed to assemble a perspicuous overview, of a subject designated by an intricate—and indeed tricky—intentional verb like 'thinking' and its cognates is not to catch inwardly a process in the act, but to describe as fully as possible, and thus examine *in context*, the vastly intricate employment of those mental concepts. And one has to remove the conceptual blockades that would stand in the way of such lengthy and involved journeyings.[19] This would constitute, in so far as it clears ground necessary to a clear view, *positive* progress. In §340 we find: 'One cannot guess at how a word functions. One has to *look at* its use and learn from that. But the difficulty is to remove the prejudice which stands in the way of doing this. It is not a *stupid* prejudice.'

One preliminary way of eroding the blockade to genuine understanding is to reconstrue the word 'thinking' as a *tool* rather than as the name of an inner process: 'Look at the word "to think" as a tool' (§360). Naturally such investigations should proceed by appealing not to the law of excluded middle, but to literature: *look* at its uses, and examine how

---

[19] I refer here again to the Preface of *Philosophical Investigations* (p. ix), the methodological import of which we are now in a position to see more clearly. It is in this Preface that Wittgenstein explains the *necessity* of his compositional style; writing in the form of sequential philosophical remarks was 'connected with the very nature of the investigation. For this compels us to travel over a wide field of thought criss-cross in every direction.—The philosophical remarks in this book are, as it were, a number of sketches of landscapes which were made in the course of these long and involved journeyings.' That much we have seen, but he also there asserts the value of approaching the same (he must mean, given his own thorough investigations into the significance of context for meaning, the same, *very broadly speaking*) points from many different directions. It is this last point in particular that constitutes—to say the least—salutary methodological advice and that we see so powerfully and beneficially enacted in the passages discussed here.

the word is used as a *tool*. There remains, however, at the very least one lingering question, one doubt about this philosophical-methodological advice. Does it mean that we must accept *only* multiplicity, *only* diversity, *only* particular cases? Here the most positive answer we have yet encountered emerges. While it is true that the question 'What is thinking?', where this is asked in special connection to the larger inquiry into the self, will not be given the simple unitary answer that the essentializing philosophical impulse might have wanted, still such an investigation would deliver—if prejudices concerning both naming as an invariant determinant of meaning and introspective ostensive definition are first removed—nothing less than, indeed, the essence of the subject. For as Wittgenstein puts it in §371, '*Essence* is expressed by grammar.'

2. AUGUSTINE AND THE AUTOBIOGRAPHICAL SITUATION

An autobiography, it is true, can be many things—one of which is surely the philosophically interesting circumstance in which the subject takes itself as its own object of investigation. This circumstance—the inner 'logic' of the autobiographical situation—naturally mobilizes all of our philosophical, or pre-philosophical, intuitions concerning the alleged nature of the self as primarily (in Descartes's phrase) a thinking thing. Indeed it is difficult to get so far as to conceive of the meaning of the *word* 'autobiography' without picturing the self in a manner consistent with the one Wittgenstein's reflections have been unearthing, exposing the motivations for, and ultimately undermining. The self is thus construed in a distinctly dualistic way—such that in autobiographical compositions the self is thought to settle into the self-assigned task of revealing the contents of the inner self—contents transparent *only* to that writing self whose task it is to inspect and report truthfully on the (ontologically hidden) contents of the pre-written self. On this model—a model deeply enforced by this unanalyzed or semi-reflective conception of autobiographical *revelation*—autobiographical truth is thus construed in terms of correspondence, but correspondence turned inward: that autobiographical sentence or proposition is true which corresponds to the inward fact of the case as transparently known only to the writer. Similarly, autobiographical honesty is construed in precisely these terms as well: the honest autobiographer is thus conceived

as one who, not unlike a positivist of the inner world, continually verifies, word by word and line by line, the direct correspondence between the actual pre-expressed or pre-narrated state of inner affairs and the propositional depiction of that state in language.

One of the authors for whom Wittgenstein had profound respect is the author of the hugely influential work of autobiography which, we may easily believe at first glance (and, as we have seen, a theory-suffused glance it may be!), shows this larger conception of the autobiographical situation to be true, or at least very much along the right lines. Augustine's great *Confessions*[20] indeed looks like a work capable of once again resurrecting the conception of the self as the inner repository of thinking—where thinking is pictured in all the pre-Wittgensteinian ways previously discussed. In short, we face the question: is it?

Naturally, a full examination of everything of broadly philosophical significance in that book would require at least a book of its size. But it is possible within our scope to examine a few passages of Augustine's autobiographical project as they directly pertain to the question of the nature of thinking and to the larger conception of the very nature of autobiography engendered by that prior conception of the thinking self.

If the conceptual picture of the self as a hermetically sealed thinking thing were correct, we would expect to find an introspecting Augustine, poised between (1) the inner, pre-narrated self and its written, outward correspondent, and (2) the experienced present and the remembered past, where the productive to-and-fro between the intertwined polarities yields both self-knowledge[21] (inwardly) and self-revelation (outwardly). Yet even at the very beginning of Augustine's self-narrative, we find that picture implicitly unsettled. Writing of his perception of infant behavior, he says: 'I have myself seen jealousy in a baby and know what it means. He was not old enough to talk, but whenever he saw his foster-brother at the breast, he would grow pale with envy. This much is common knowledge' (p. 28). It is not only, first, that this pre-linguistic infant behavior is knowable externally by others before it

---

[20] Augustine, *Confessions*, trans. R. S. Pine-Coffin (Harmondsworth: Penguin, 1961). For an extremely helpful study of Augustine's philosophical thought, see Christopher Kirwan, *Augustine*, The Arguments of the Philosophers (London: Routledge, 1989); see particularly ch. iii: 'The Nature of Speech', 35–9, in which Wittgenstein's employment, and criticisms, of Augustine's linguistic ideas figure prominently.

[21] But not, according to this picture of the self, *new* self-knowledge; the self and its content are here allegedly transparent. However, the integration, or uniform narration, of that transparent self would be as close to new self-knowledge as conceptually allowed by this picture—in that it is newly narrated.

is knowable (in terms of self-recognition and linguistic accessibility) to the (infant) self, thus destabilizing the inner-to-outer model; or, second, that it is indeed 'common' knowledge of a kind that antedates any deliberate expressive revelation on the part of the infant; but more so, third, that Augustine himself, reflecting on his own infancy, *deduces* how his early pre-reflective life must have been, thus subverting the introspectionist model at the very outset. In this connection he writes: 'I do not remember that early part of my life, O Lord, but I believe what other people have told me about it and from watching other babies I can conclude that I also lived as they do' (p. 28).

But here, one may quickly object quite reasonably, he is talking about *infants*, and infants do not write autobiographies, and *that* is the circumstance we want to illuminate. Although that is true and a fair objection, we should remain mindful that (1) Wittgenstein often returns to cases of early language-learning to cast light on linguistic meaning and, more pointedly, to expose an intuitively supported misconception for what it is, and (2), infant or not, the thematic seed has already been planted of attaining self-knowledge decidedly *not* in accordance with the conceptual picture in question. Moreover, Augustine also severs the judgment of the truth of an autobiographical claim from inwardly derived knowledge: he adds next, 'But, true though my conclusions may be . . .', where those conclusions are, again, much more derivations than inner observations. He in fact establishes and retains a skeptical tone concerning the true or tru*est* verbal formulation; shortly after the above passage, having written 'The next stage in my life, as I grew up, was boyhood,' he quickly queries, 'Or it would be truer to say that boyhood overtook me and followed upon my infancy' (and ever-mindful of metaphysically curious implications, adds '— not that my infancy left me, for, if it did, where did it go?', p. 29). The 'inner positivist' would not need to speculate concerning the truest formulation, nor would there be logical room to do so: the inner truth would be transparent, and its verbal correspondence correct.

Indeed there are numerous instances of skepticism of this kind, along with a blindness to the meaning or the significance of one's own present action, which may or may not be uncovered at a later time of one's life. Augustine witnesses such blindness[22] in others, notably in the case of

---

[22] Augustine gives himself good reason to call into question the very conception of inwardly determined meaning, the conception with which Wittgenstein opens *Philosophical Investigations*; moreover, Augustine appears to do this without realizing it,

a lawyer speaking in court who, secretly 'long[ing] for fame as a fine speaker will stand up before a human judge, surrounded by a human audience, and lash his opponent with malicious invective, taking the greatest care not to say '"uman" instead of "human" by a slip of the tongue, and yet the thought that the frenzy in his own mind may condemn a human being to death disturbs him not at all' (p. 39). Here, obviously, meaning and significance are—to put it one admittedly dangerous (in that it can instantly revivify the behaviorist picture) way—*external*: the speaker's *words* do a gruesome work about which the speaker is unconcerned or unaware. Augustine is thus aware, and here implicitly records, that biographical reflection upon the meanings of the speaker's words is not circumscribed by the intentional world of the speaker. And more complexly, the speaker's words are chosen not in the interest of justice (as it may appear) but in the interest of fame as a rhetorician, and this *is* the truth of the motivation for the utterance (albeit known only autobiographically to the speaker). Here, the purpose for which the words are chosen is *not* expressed *within* the linguistic utterance, and this drives a wedge between the intention in speaking and the content of what is spoken—which is, if in broad terms, inconsistent with the inner–outer correspondence picture of the self thinking.

And Augustine witnesses his own blindness to the meaning of his actions at the time of the action repeatedly; in these contexts he frequently speaks of 'losing himself' and the great gulf between the reality of the situation and his false representation of that reality to himself. As is now fairly evident, if the picture of the self thinking that is in question here were accurate, such losses of self of the kind he retrospectively diagnoses, and such believed but false self-representations, would be extraordinarily puzzling if not metaphysically impossible. Indeed, if the self's epistemic circumstance truly were as this picture of the autobiographical situation suggests, phrases like 'losing oneself' and 'false self-representation' would be much like linguistic analogues to M. C. Escher's architectural drawings: drawable or sayable, but impossible. Yet Augustine was 'foul to the core' while 'pleased with [his] own condition'; he 'lost himself' due to his having been

---

and thus to that extent ironically exemplifies the precise thing he is observing in others in these and similar passages. Indeed, others (and Augustine himself in retrospect) are seen to have said and done diverse things whose meanings they or he did not grasp co-temporaneously with the utterance or the action.

'inflamed with desire for a circuit of hell's pleasures'. And retrospectively redescribing his thought and action in a way his earlier thinking-and-acting self would have found incomprehensibly false to what he then perceived as the facts at hand, he 'ran wild with lust' and thus, with 'adolescent sex welling up within' and exuding 'mists which clouded over and obscured [his] heart', in truth 'could not distinguish the clear light of true love from the murk of lust' (p. 43). Augustine, an author who initially conjures images of an autobiographer who transparently knows his own thoughts and who thus writes a great text of correspondingly true self-revelation, in fact—in autobiographical *practice*—employs images of mists, clouded-over semi-perceived realities, whirlpools of incomprehension, of self-misunderstanding, of having been 'deafened' to the truth and 'fettered' (p. 43) away from it, of being swept away unawares by the body's appetites, and so forth through countless similar rhetorical devices that, collectively taken, both deepen and thoroughly problematize the picture of the thinking self transparently facing itself, taking itself as its subject matter, seeing itself clearly in its reflexive mirror.

And indeed, a layer above this, any reader familiar with Augustine's concern—his deep, sustained, detailed concern—with the full articulation of his earlier years spent in pursuit of pleasures of the flesh, as he puts it in one of very many ways, 'being tossed and spilled, floundering in the roiling sea of my fornication' (p. 43), will very probably arrive at—at the least—a doubt concerning Augustine's own autobiographical clarity, and indeed honesty, with regard to his having transcended interests in these matters.[23] It does not take a post-structuralist to know that speech—linguistic representation—can itself make that which is absent seem (more) present, and Augustine's repeated articulations of these lustful stages on his life's way (from which he famously prayed for deliverance—'but not yet') may betray an interest he both retains and does not acknowledge to himself (thus rendering the interest or desire itself both present and absent). Such interpretive ambiguities on the part of the reader recapitulate the self-interpretive ambiguity of the autobiographer: it is now perhaps only necessary to point to the obvious incompatibility between these layered complications—rising initially out of (1) the author's self-interpretation; then out of (2) the text's challenge to our dis-confirmation of the authorial assertion; and then out of (3) the reader's interpretation of the previous two levels—and

---

[23] See e.g. p. 55: 'and it would be all the sweeter if I could also enjoy the body of the one who loved me', etc.

the unifying and simplifying metaphysical picture we may have initially brought to those far more unwieldy and complicated facts of the autobiographical situation.

There are countless further cases that either reinforce the complicating themes we have seen or add new ones in Augustine's text. Motive-skepticism abounds: a murderer is described as 'choos[ing] to be cruel and vicious without apparent reason' (p. 48), yet Augustine makes clear that there was in truth a reason, although unknown both to the murderer and to the teller of the tale (p. 49). He continually separates the significance of the act from the thought taken in performing the act, e.g. in referring to his own earlier crimes of theft (p. 51). He retrospectively distinguishes between falling in love and falling in love with the idea of love, a thought unavailable to him while living without this distinction in what he calls the 'hissing cauldron of lust' of Carthage (p. 55). He emphasizes the great significance not only of the embodied nature of human life, and he thanks God not for giving him a Cartesian consciousness transparent unto itself, but for 'my body [and] its five senses; you furnished it with limbs and gave it its proper proportions; and you implanted in it all the instincts necessary for the welfare and safety of a living creature' (p. 28), and he emphasizes *the meaning-determining nature of external context*: in his early love life, while he expected only joy, pleasure, and happiness stemming from his pre-experiential cognition on the matter, in contextualized truth he found himself 'lashed with the cruel, fiery rods of jealousy and suspicion, fear, anger and quarrels' (p. 55). A great gulf thus separates the pre-experiential reflective expectations and the embodied, lived experience in context—and the autobiographical meanings of those actions are inaccurately identified in advance and incompletely grasped *in situ*; only retrospection in these cases makes genuine epistemic advances, and as we have seen, even then the matter is largely inconsistent with pre-Wittgensteinian philosophical expectations. Augustine himself expresses this contextualism and motive-skepticism succinctly: 'the appearance of what we do is often different from the intention with which we do it, and the circumstances at the time may not be clear' (p. 67).

There are indeed many further issues of philosophical significance in Augustine's text. Like Aristotle before him, he discusses the initially curious fact that we enjoy the depiction of suffering in tragedy, and these passages (pp. 55–7) serve as an inoculation against the belief that specific thoughts will be directly associated with specific emotional charges. Thus, any attempt to develop an 'atomistic' conception of thoughts

and linked emotions is precluded (and such an analysis would naturally
be consistent with the disputed conception of thinking): his discussion
'shows that sorrow and tears can be enjoyable' (p. 56), that the context of
the thought determines the emotional impact of the thought-experience,
and if indeed 'the circumstances at the time may not be clear', then we
live with a self-interpretive skepticism born of the recognition that, as
contexts are ever-changing and evolving, the significance of our thoughts
and actions are similarly changing and evolving. That does not mean that
autobiographical honesty and autobiographical truth[24] are not *possible*,
but it does mean that, here again, these autobiographical virtues are not
explained in the simple and unitary ways (i.e. inner correspondence)
suggested by—or indeed conceptually enforced by—the picture of
thinking Wittgenstein is unearthing and repudiating. Augustine is also,
somewhat like Wittgenstein, repeatedly suspicious of the formulations of
metaphysical or philosophical questions and the implications they carry;
he refers in one discussion to the misleading power of his 'own specious
reasoning' (p. 62), and he of course finds the metaphysical *question*
concerning the nature of time puzzling (p. 263 ff.; see Wittgenstein's
discussion in *Philosophical Investigations* §§89–90). Augustine records
the failings of introspection as a source of self-knowledge; in a moving
passage in which he reflects on the death of his close friend, he reports
that his 'eyes searched everywhere for him, but he was not there to be
seen' and Augustine now hated the places they had enjoyed together,
for those places 'could no longer whisper to me "here he comes!" as
they would have done had he been alive but absent for a while' (p. 76).
He thus writes: 'I had become a puzzle to myself, asking my soul again
and again "Why are you downcast? Why do you distress me?" But
my soul had no answer to give' (p. 76). And he repeatedly indicates
the uncertainties of self-understanding: further in the same passage, he
speculates: 'Perhaps this, too, is why I shrank from death, for fear that
one whom I had loved so well might then be wholly dead' (p. 78).

   Augustine, like Wittgenstein, is no behaviorist: he certainly never
says or implies that our self-understanding is *reducible* to behavioral

---

[24] Interestingly, Augustine's statements imply that a speaker's conviction cannot serve
as an index of truth; in his discussion of his having fallen in with a group of sensualists
who, 'with glib tongues', took 'Truth and truth alone' as their motto, he identifies
all of this retrospectively (where reconsideration equals moral and epistemological
recontextualization of the kind we will consider in Ch. 6, Sect. 3, below) as deception
(see pp. 60–1). Judgments of truth are thus, for Augustine, far more an external matter
than his introspectionist image would begin to suggest.

descriptions, nor does he extend the inferential self-descriptions derived from externally observed behavior from infancy to the adult case. And, as his autobiographical practice indicates, he does not challenge the fundamental asymmetry between the first- and third-person cases. But his practice also shows that, despite the philosophically motivated image we may have of the author of the *Confessions*, he is certainly no proto-Cartesian either. The grand sweep of his autobiographical project indeed traverses too wide a field to fit into that reductive and unifying vision of the self and its reflexive knowledge, and not unlike Wittgenstein's investigations, its philosophical significance crisscrosses at countless locations. Thinking, as we have seen, for Wittgenstein and for Augustine is profoundly contextual; it is not pre-linguistic, pre-expressive pure cognition. In fact Augustine's project affords the reader a perspicuous overview of the very many different kinds of things covered by the phrase 'the self, thinking'. And in doing so, i.e. in showing how concepts of self-reflection and self-description actually work in their layered intricacies—in showing what Wittgenstein called their grammar—it reveals, if in a nonreductive manner, the essence of reflexive thought. And without a grasp of that, we cannot hope to understand how autobiography works.

To conclude these two sections before progressing into the third: (1) private thought is not, as Wittgenstein powerfully argued and as Augustine showed in autobiographical practice, private in the *metaphysical* sense of that term as elucidated in previous chapters; (2) getting clear about the meaning of the word 'thinking' is not, contra the dualist-introspectionist picture, achieved by watching ourselves, by inward observation, when we think; (3) the facts of the autobiographical situation show that the self is not transparent unto itself—indeed *thought*'s unitary and Cartesian self-image is greatly distorted; (4) thinking, as investigated and enacted throughout the *Confessions*, displays no single property common to all cases in its class; (5) the relation between a thought and a proposition is shown by Augustine to be anything but direct and immediate; (6) Augustine shows over the course of his book what genuine introspection is (in a manner fully consistent with Chapter 3 above)—against the Cartesian and Jamesian misconception of ghostly inner-directed mental-object scrutiny; and (7) Augustine is too sophisticated a philosopher to fall into the trap of attempting, as Wittgenstein said, 'to explain the meaning of the word ''thinking'' in a primitive way'. Augustine, to be sure, *thinks* about his life, but having followed him through this variegated investigation, a reader will not

then be likely to attempt to *define* self-thinking essentialistically (through inner-ostensive definition) as inward 'pre-speaking'—and again this, as Wittgenstein concluded, 'is where our illusion is'.

  Wittgenstein suggested that we 'look at the word "to think" as a tool' rather than as the *name* of an inner process, and thinking of it as a tool, as we saw above, is one way to remove the conceptual blockade to *looking at* the use of a word and learning from that. The difficulty, Wittgenstein said, is to remove the prejudice—and it again, as we also saw him remark, is certainly not a *stupid* prejudice—'which stands in the way of doing this'. Augustine indisputably made positive progress on this score, and although, as I have suggested, he may protest too much in regard to his life of sensual indulgence, he also shows that there are more things in autobiographical thinking than are dreamt of in that thinking's metaphysical mischaracterization, in thought's image of itself. It is Augustine's use of that distinctive category of thinking, where the subject takes itself as its object of investigation—but in ways free of the interrelated conceptual ailments we have seen Wittgenstein diagnose—that shows its meaning. But Augustine is not, of course, the only author to have composed literary work of the most immediate relevance to our questions concerning the nature of autobiographical language.

3. WITTGENSTEIN UNDERGROUND (AND DOSTOEVSKY'S *NOTES*)

It was Walter Kaufmann who said of *Notes from Underground*, 'The atmosphere of Dostoevsky's *Notes* is not one of soft voices and dim lights: the voice could not be shriller, the light not more glaring.'[25] But light cast on precisely what? Kaufmann shortly thereafter answers this question. Of the compact and powerful work that he describes as 'one of the most revolutionary and original works of world literature', he says that by shifting the focus away from the tradition of literary naturalism that maintained and continually sharpened the description of the material circumstances of both private and public life, Dostoevsky's *Notes* cast its harsh illumination on 'man's inner life, his moods, anxieties, and his decisions'. With language suggesting a Cartesian point of inward-directed

---

[25] Walter Kaufmann (ed.), *Existentialism from Dostoevsky to Sartre*, (New York: Meridian, 1975), 13; this and following quotations.

consciousness, hermetically sealed from the exterior world and transparently knowable unto itself introspectively, he adds that these inward events 'were moved into the center until, as it were, no scenery at all remains'. The removal of *all* 'scenery' would leave only what we might then be led to picture as the pure self, a phenomenological interior that, more than any contingent material or external thing (such as a body, physical environment, or human context), defines the essence—'essence' here taken in the pre-Wittgensteinian sense—of selfhood.

And the picture into which we are led by these reflections truly is Cartesian at this point on the conceptual landscape as well, precisely because the light that is cast emanates from no external source, and its strength, its 'wattage', will thus be a direct function of the power of the introspective effort. Dostoevsky's *Notes*, approached this way, will seem to offer a paradigm case of autobiographical self-description of a kind that details the inner workings of a mind in pursuit of knowledge of itself, knowledge of both its contents and its structure. Indeed Nietzsche, after recording the extraordinary joy he felt upon recognizing even at first glance a deep affinity of sensibility, describes Dostoevsky's *Notes* as 'a kind of self-derision' of the Socratic injunction to know thyself. This is, in any case, the fairly widely received impression of this work, which might be called its philosophical–literary image. But does a close reading of Dostoevsky's text support this image of the Cartesian mind caught in the act of hermetic self-scrutiny any more than did Augustine's? And if not, does its not doing so similarly disclose a number of further features of the self-interrogative phenomena we are attempting to clarify throughout this study?

It is true that we learn, within the space of the work's first twenty-one words, that the underground man is, according to his self-description, sick, spiteful, unattractive, and suffering from a diseased liver. This, if anything, does seem self-focused. But does it, at the very outset, lay the foundation upon which to build a model of a mind so internally engaged as to sever its ties to externalities, leaving *exclusively* the pure inward self? Not in any clear way: the sickness could be of the body or of the mind; the spitefulness would be discernible more in relation to others than it would be as an internal state unto itself; the unattractiveness could as easily be of the flesh as well as of sensibility; the liver is, inescapably, a matter of embodiment. This work, at its very point of departure, is neither exclusively of (what we initially picture as) the inner life, nor of the outward: it instead makes problematic the very dichotomy upon which the *Notes*' image rests.

While it is true that the *Notes* depicts what can only be called a form of hyper-self-consciousness (and as Nietzsche saw, a self-derisive form of that), the question is what precisely this comes to. The underground man writes: 'I used to be in the government service, but am no longer' (p. 54), thus positioning himself as a present self looking back at himself across the distance of time. This itself is hardly *hyper*-self-consciousness; it is the retrospective point of view taken by every autobiographical writer. He adds: 'I was a spiteful official. I was rude and took pleasure in being so. I did not take bribes, you see, so I was bound to find recompense in that, at least.' But now he adds a line of self-description that looks at himself from a distance, but a distance not opened by retrospection.[26] It is opened, in the writer's present, by an ironic detachment from his own sentences, and here he advances to the 'hyper' level. He adds: 'A poor jest, but I will not scratch it out.' He thus negatively judges, at the time of writing, his own little jest, and then at the same time yields to his second-order impulse, overriding his implicit first-order impulse to scratch it out, to leave it in to spite himself. This in itself suggests a nonunified consciousness—against that which we would expect in the Cartesian interior—in that he is writing the jest, judging that jest, recognizing and yielding to an impulse to override that motivation, and recognizing that impulse as self-spite.

But Dostoevsky's portrayal of his mental condition is also more than that. The underground man immediately adds: 'I wrote it thinking it would sound very witty,' thus adding still another layer, giving specificity to the intention fueling the initial jest. Most importantly, he adds to this the remark: 'but now that I have seen myself that I only wanted to show off in a despicable way, I will not scratch it out on purpose!' This adds still another layer of self-description, specifically his self-punishing reason for leaving in the jest, one that is now left in 'on purpose' rather than leaving it in merely passively, i.e. he is now taking an opportunity for self-spite that has arisen within the sentence by happenstance. Furthermore, it also adds a description of a condition of inward *non*transparency, shown when he writes: 'but *now* that I have seen myself', clearly indicating that he is learning of his motives *as he writes*. That is, he recognizes within his own sentences as they

---

[26] For an engaging discussion of this issue, and particularly in relation to Sartre's claim that the attempted act of consulting one's inner evidence in order to identify one's state or condition is itself a sign of self-alienation, see Richard Moran, 'Self-Knowledge: Discovery, Resolution, and Undoing', *The European Journal of Philosophy*, 5 (1997), 141–61. I return to a closely related discussion of Moran's in Ch. 5, Sect. 3, below.

unfold the motives, the alleged purely mental events, that should be the objects upon which his harsh introspective light is directed to the exclusion of all the other external 'scenery'. As Dostoevsky has in fact shown in this case—one that with insufficient care might be taken as a literary depiction of hyper-Cartesianism—the light is directed, but on the *externality* of his own writing (more accurately, upon his writing-in-the-process-of-writing). He is positioned not as the sole owner of a hermetically sealed point of consciousness, but rather as a mind: (1) positioned in relation to a remembered past that is not given transparently and immediately in introspection but rather one with a problematic significance that he must *work out*; (2) positioned with an ironic distance from his present self; and (3) positioned in relation to his present sentences as they appear not with immediate inward transparency, but rather with a layered complexity that belies the simple notion that the autobiographical or self-descriptive sentence stands in a one-to-one relation to a mental state only contingently expressed in language. Like Augustine before him, but with particularities that bring their own distinct light, Dostoevsky goes on to show that the relation between the self and its language is far more interesting and far more complex than any one-to-one inward-correspondence model could accommodate.

This is not to say that we do not encounter fairly stark contrasts in this small book between what we can intelligibly call the inner and the outer: we do. However, on close inspection, they do not correspond to the traditional metaphysical dichotomy that places the intangible inner as the essence of selfhood, and the corporeal outer as its contingently related seat. In discussing what he calls 'the chief point about my spite' (p. 54), the underground man, as he puts it, locates 'the real sting of it . . . in the fact that continually, even at the moment of the acutest spleen, I was inwardly conscious with shame that I was not only not a spiteful but not even an embittered man, that I was simply scaring sparrows at random and amusing myself by it' (p. 54). Here, we might say, he is inwardly conscious of one thing while outwardly displaying the other. Yet this description of what he sees about himself is far too simple, too crude, to capture what this depiction of mental life really contains.[27] Dostoevsky shows this to be, for the underground man, an irritating conflict between a role he plays and, not his self-introspected moral state incompatible with that role, but rather his humanly engaged social

---

[27] I will return to the process of gradually dawning self-understanding in Ch. 6 below.

reactions to the kindness of others (he refers to his being genuinely touched upon being given a cup of tea with sugar). And when the underground man goes on to describe his awareness of the conflict between the played role and who he really is, he describes himself as being 'conscious every moment in myself of many, many elements absolutely opposite to that' (i.e. opposite to the spiteful person he can deliberately appear to be), where these elements 'swarming' within him are both memories of past nonspiteful human interactions and present impulses to continue with just such interactions.

There is, then, an intelligible contrast here between the outer role and the inner person, but we make sense of the phrase 'inner person' here as a self that is recognizing its manifest social proclivities, not as a self introspecting upon its inwardly hidden pre-linguistic mental content. Indeed, Dostoevsky shows that the underground man understands his own interior life in just this way. And it is with the layered, relational conception of self-awareness (again, relations to others, to his past self, to his present self, and to that self's present words) that he plays with self-descriptive paradox (and we as readers recognize this move in his language-game and see that this move is made possible by the moves preceding it): he writes, 'I was lying when I said just now that I was a spiteful official. I was lying from spite.' The spite that motivates—or that he pretends to harbor as motivation, given what he has just said about his truer response to the tea with sugar—the lie is manifest *in* the language he is presently writing and not hidden in a metaphysical interior. It is *language* that not only conveys, but more strongly constitutes, the content of the inner self of which we can and do make sense. And so closing Part I of this study of self-consciousness, Dostoevsky has the underground man write: 'But what can a decent man speak of with most pleasure? Answer: Of himself. Well, so I will talk about myself' (p. 56). This only underscores that the self-revelation, the harshly lighted self-scrutiny depicted here, is inescapably linguistic. And it begins to suggest that language, as Wittgenstein showed, is unavoidably public.

Even in Wittgenstein's early writings of the *Tractatus Logico-Philosophicus* we see intimations of this necessarily public view of language.[28] In *Tractatus* 3.326 he wrote: 'In order to recognize a symbol by

---

[28] We saw in Ch. 1 the picture of the self, derived from Schopenhauer, influencing Wittgenstein's thinking at this early stage and moving him close to a solipsistic conception of private experience. It is of interest that these remarks concerning the public nature of

its sign we must observe how it is used with a sense,' which he followed in 3.327 with 'A sign does not determine a logical form unless it is taken together with its logico-syntactical employment,' and then in 3.328 we find the bluntly direct 'If a sign is *useless*, it is meaningless.' This emphasis on use, emerging in incipient form in Wittgenstein's early writings and then reaching full maturation in *Philosophical Investigations*[29] in passages we have already considered, is, to say the least, helpful to keep in mind while considering the particular achievement represented by Dostoevsky's *Notes*. In *Philosophical Investigations* §43 Wittgenstein, we should here recall, famously wrote, 'For a *large* class of cases—though not for all—in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language,' and later in that work (consistent with the 'tool' analogy above) he invites us to 'Look at the sentence as an instrument, and at its sense as its employment' (§421). Given the foregoing chapters, we are now in a position to comprehend more fully what it means to say that these remarks 'de-psychologize' our conception of meaning: it is syntactical employment, and not a ghostly inward pre-linguistic event that is only arbitrarily associated with an outward verbal sign, that gives meaning; it is usage in the language that gives sense, that ensures intelligibility, and not an event *before* that language is used; and a sentence seen as an *instrument* brings a shift in our expectations concerning what constitutes a sentence, i.e. once seen as instrument it is not thus easily seen as an outward correspondent to an article of inward pre-linguistic consciousness. Dostoevsky's underground man is *in* language; his thinking, as depicted with the greatest literary subtlety by Dostoevsky, is not prior to that and of an ontological kind different from it. His layered senses are in the sentences, and he learns *from* them rather than invariably imparting life *to* them.[30] Very late in his life Wittgenstein wrote what was perhaps his most compressed articulation of his long-developed point: '*Practice* gives the words their sense.'[31]

Dostoevsky's underground man shows a mind—if of a particularly ironic, self-mocking, hypersensitive sensibility—engaged in the *practice*

language, and the necessity of use for meaning-determinations, are already at work in his thinking, feeling the pull of another way of seeing these issues.

[29] Ludwig Wittgenstein, *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958).

[30] See Wittgenstein's discussion in his *Blue and Brown Books* 1–30.

[31] *Culture and Value*, ed. G. H. von Wright, trans. Peter Winch (Oxford: Blackwell, 1980), 85.

of autobiographical reflection, and as we shall shortly see, it is impossible to imagine that practice taking place anywhere but within language. But, having said this, it is equally important to bear in mind that autobiographical self-scrutiny, as depicted with such mimetic fidelity by Dostoevsky, does not reduce to language *when language is construed as an exclusively outward phenomenon*. That mischaracterization, as we now know, would reduce the position exemplified in Dostoevsky's text to behaviorism, and would lose the ineradicable asymmetry between the first- and third-person case. Wittgenstein spoke, in this connection, of the distinctive relation we, as speakers and writers of our own words, have to our language that nobody else can or could ever have. This, despite its explicit rejection of a self-as-other behaviorism, is at the same time not a reversion to Cartesianism.

The underground man, having resolved to talk about himself, quickly goes on to describe consciousness as an illness: 'I am firmly persuaded that a great deal of consciousness, every sort of consciousness, in fact, is a disease. I stick to that' (p. 56). Such consciousness, as Dostoevsky shows within the underground man's self-descriptions, while relational in the ways considered above, is, as a closer reading will show, not reducible to behavior for the following reasons: In writing of returning to his little corner of Petersburg 'acutely conscious that that day I had committed a loathsome action again' and that 'what was done could never be undone', he found himself 'secretly, inwardly gnawing, gnawing at myself for it'. He finally reports that this inner crescendo of self-recrimination concluded in 'a sort of shameful accursed sweetness, and at last—into positive real enjoyment!' (p. 57). The consciousness depicted here, if contorted by perverse self-recrimination, is again relational in its clear linkages to the external world: the recrimination necessarily refers, for its intelligibility, to its object—the loathsome action of that day, where that action was by definition a public action. That consciousness, as described with such exactitude with line-by-line layering of complexity, is not Cartesian. But the gnawing depicted there is, as explicitly described, 'inward', and this may give a Wittgensteinian pause. Not, however, for long: the phrase 'inward gnawing' suggests, here in its own distinct way, an inward–outward distinction within a context of hypersensitive self-scrutiny, where 'inward' has a syntactical employment signifying a repeated going-over in the imagination of the loathsome act such as it was, and not where the word is granted meaning by a mental act of private ostension. It has its use within Dostoevsky's incrementally expanding language-game of closely observed self-description and derives its sense

from that—not from a generic metaphysical dichotomy between the inner and the outer that is prior to language. Rather, that distinction is, indeed, an instrument, a tool, in Dostoevsky's language. Seemingly designed to show the necessarily public or socially interactive nature of language in opposition to the picture of hermetically sealed private experience, Dostoevsky gives the underground man the concluding line 'I have spoken of this because I keep wanting to know for a fact whether other people feel such enjoyment' (p. 57).

In subsequent passages we find the underground man cultivating his own finely honed appreciation of despair; these passages are instructive in present terms precisely because we witness him *working out* his appreciation in a manner even more extreme than what we saw in Augustine. That is, he does not seal himself into his introspective interior to make what for him is a kind of twisted moral progress, but rather we see him writing toward what he arrives at as the 'intense enjoyments' of despair. It is evident, once one is minded to look at the text in this way, that Dostoevsky has shown introspection for what it is, and not for what it is pictured to be under the influence of a grand metaphysical dichotomy. And the underground man again accentuates the ineradicable role of the external in identifying the precondition for what he regards as the correct appreciation of despair: the full grasping of the 'hopelessness of one's position', where that position is external and relational rather than internal and psychically hermetic. This is only underscored in his following words concerning the man of acute consciousness coming, not out of nature, but out of a retort (to a real or imagined insult, offense, or slight). And such a 'retort-made man' (p. 59), as he terms any individual such as himself, is a possessor, for better or worse, of a form of consciousness created in a crucible of public interaction, not private, solipsistic, internally lighted reflection.

In still another now famous passage, Wittgenstein, referring to both his views on language and on mind, wrote: 'We may not advance any kind of theory. There must not be anything hypothetical in our considerations. We must do away with all *explanation*, and description alone must take its place' (*Philosophical Investigations* §109). Seen under this aspect, this is precisely what Dostoevsky provides vis-à-vis the exacting description-sans-theory of a mind taking itself as its own object of study. Such an exacting description, one that for Wittgenstein 'leaves everything as it is' (*Philosophical Investigations* §124), shows in the most minute detail the phenomena in question. And that process of detailed showing, as Wittgenstein says, 'gets its light, that is to say its purpose,

from the philosophical problems' (*Philosophical Investigations* §109). If the philosophical problem is that of understanding—gaining an undistorted, perspicuous view of—autobiographical self-examination and, more precisely, the role of our language in that autobiographical endeavor, then it is this problem that casts light, harsh or otherwise, on Dostoevsky's exacting descriptions. And—to connect back to the concerns of Chapter 1—it is within those philosophically public investigations that we give a context-specific use to the concept of autobiographical privacy.

That Dostoevsky's description of the mind of the underground man is of a kind that—despite the initial appearances and the received image of this piece of psychological literature—repudiates the Cartesian image of the work becomes increasingly clear as the *Notes* progress. Having changed his self-descriptive terminology from 'man' to 'mouse', he minutely describes the mouse's process of self-lacerating recollection. Ensconced within its 'mousehole' after another perceived insult, hiding 'in its nasty, stinking, underground home our insulted, crushed, and ridiculed mouse promptly becomes absorbed in cold, malignant and, above all, everlasting spite'. This description of a mind positioned into—indeed given its mental content by—outward circumstances and events is followed by a portrait of its mental process: 'For forty years together it will remember its injury down to the smallest, most ignominious details, and every time will add, of itself, details still more ignominious, spitefully teasing and tormenting itself with its own imagination. It will itself be ashamed of its own imaginings, but yet it will recall all, it will go over and over every detail' (p. 60). This mind, however atypical or idiosyncratic in its specific self-recriminations, is one that rehearses the minutiae of its public encounters (and in this particular case its perceived humiliations); the more one looks to Dostoevsky's text with the philosophical question that gives such a reading its 'light', the less the Cartesian legacy seems to have any hold here—the details of these self-investigative memory processes continually loosen the grip of that picture.

The steady absorption into the underground man's mind of the cold, malignant spite of which he writes, and which he cultivates over many years of self-flagellating memory, is, as we have seen to a certain extent, public. How might that claim now be more fully articulated? I said above that the knowledge the underground man is depicted by Dostoevsky as having (knowledge of what we can intelligibly call his inner states) is not a matter of private, inward, *ostensive* definition. In the course of his unearthing of the long-buried misleading presuppositions

about linguistic meaning that made the idea of a private language seem so intuitively plausible (when it is in truth disguised incoherence), Wittgenstein wrote his way into the much-discussed example of the beetle in the box. That example has its power oppositionally: it can cause us to rethink radically the notion of a language made private by virtue of the inviolable metaphysical privacy of the inner mental objects that allegedly serve as the referents to which the words—more specifically the psychological nouns if we are thinking of self-knowledge concerning mental states, and psychological verbs if we are thinking of self-knowledge concerning mental actions—refer and that possess meaning by virtue of that referential relation. This picture of the word-meanings of our mental vocabulary is quickly unsettled by Wittgenstein's observations, and this picture and its disentangling give still more 'light' to the philosophical reading of Dostoevsky's *Notes* as we are pursuing it in relation to the clarification of the nature of autobiographical or self-revelatory language.

The underground man begins his fourth section with ' "Ha, ha, ha! You will be finding enjoyment in toothache next," you cry, with a laugh. "Well? Even in toothache there is enjoyment," I answer. I had toothache for a whole month and I know there is' (p. 62). He then goes on to a typology of moans: candid moans, malignant moans, enjoyed moans, moans expressing the aimlessness of one's pain, the moans of an educated man of the nineteenth century, moans with a nasty flourish, and so forth. Does the underground man know the meaning of the word 'toothache', or more generically the word 'pain', by reference to a private inner object (represented as Wittgenstein's 'beetle' in what follows)? Does he know the meaning of each of the distinctive moan-types by reference to an inner sensation? Wittgenstein writes:

Now someone tells me that *he* knows what pain is only from his own case!—Suppose everyone had a box with something in it: we call it a 'beetle.' No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at *his* beetle.—Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing.—But suppose the word 'beetle' had a use in these people's language?—If so it would not be used as the name of a thing. The thing in the box has no place in the language-game at all; not even as a *something*: for the box might even be empty.—No, one can "divide through" by the thing in the box; it cancels out, whatever it is. (*Philosophical Investigations* §293)

This looks back to the 'de-psychologizing' remarks considered above and again traces back to those early remarks in the *Tractatus*: It is the

use in the language that gives meaning, not an inner event in which outward sign is attached to its corresponding inner-object referent. The box might possess a chameleon-object, constantly changing; each person's box might possess something different; a person's box might be empty. The word 'pain' does not get its meaning as a private name; the word may still mean (if it has a public use) while the alleged meaning-determinant is inconsistent, different among individuals, or even absent. Dostoevsky goes to considerable length—far too great to be presently recounted in the line-by-line detail it deserves—to *situate* the language of the underground man's pain into a public community of discursive practices. He describes the larger 'audience' for the moans, the man's family, the responses of his hearers, the various, and here again layered, intentions with which he performs his moans (where here again he discovers some of his layered intentions in the act of writing about them), and his own ironically detached responses to himself (thus showing that his own pain-language, in the significant sense, is public unto himself). But can the moan-types themselves be knowable as items of individuated internal ostension? Consider Wittgenstein's observation in *Philosophical Investigations* §268:

Why can't my right hand give my left hand money?—My right hand can put it into my left hand. My right hand can write a deed of gift and my left hand a receipt.—But the further practical consequences would not be those of a gift. When the left hand has taken the money from the right, etc., we shall ask: 'Well, and what of it?' And the same could be asked if a person had given himself a private definition of a word; I mean, if he has said the word to himself and at the same time has directed his attention to a sensation.

Even though these remarks can shake the hold of the picture of private ostension as the source of meaning, there nevertheless often persists the idea that the underground man may create his own 'moan-language' through such inner processes. Against such conceptual persistence, we need to press our knowledge of why any such explanation given in terms of inner processes would prove, in the end, empty. Suppose he writes, just as does Wittgenstein's famous sensation-diarist, an 'S' (with a sub-number for each variety) every time he has a specific toothache type. Wittgenstein writes: 'I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign "S" and write this sign in a calendar for every day on which I have the sensation.—I will remark first of all that a definition of the sign cannot be formulated.'

Although rather cryptic, this remark in fact says a good deal: The definition cannot be formulated because the definition would, by proceeding in our public language, transgress at the very earliest point the limits of the private language allegedly being formulated in the act of inner ostensive definition. The very *idea* of a definition of the sign is already positioned into our public language—indeed it is only available there. Thus, presuming an understanding of the concept 'definition' is to presume far too much, *if* the language is to proceed from the inside out, from the first-person sensation as the originating determinant of the meaning of the sign. (Of course, the concepts 'meaning' and 'sign' similarly transgress in being similarly public.) Wittgenstein continues: 'But still I can give myself a kind of ostensive definition.—How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be!' So the underground man cannot point to his sensation in the ordinary sense (he could point to his cheek or his jaw in that sense, but such ordinary acts of pointing again already presume a vast background, what Wittgenstein refers to as the 'scaffolding' of our public language). Is the underground man left, then, with an inward version of outward pointing? Can that give his 'S' its meaning? Could it do so with anything like the stability and reliability of our public language?

In response to his imagined interlocutor, who insists that the private speaker can impress upon himself 'the connection between the sign and the sensation', Wittgenstein writes: '—But 'I impress it on myself' can only mean: this process brings it about that I remember the connection *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about "right".'

The underground man does, as Dostoevsky shows, speak with force, authority, and the hyper-self-consciousness we have already seen him display repeatedly about his toothaches and his catalogue of moans. And to stand back, Dostoevsky as author is showing (from the vantage point of his Christian faith) in the case of the underground man precisely what he predicts as the human wreckage that will result from a godless, Westernized, un-Russian mode of being—in short, a horrific specimen of moral disaster. But as Dostoevsky also shows, the underground man voices his complaints *within*, and in no intelligible way prior to, the public stream of life into which his (and our) words have what Wittgenstein referred

to early in his work as functions, later as uses, and still later as practices. In a telling phrase the underground man speaks of his life as 'that hell of unsatisfied desires turned inward' (p. 60), but those desires, as we saw above, are situated relationally in a public life and are coherent because of that ineradicable positioning. It is after that, and not prior to it (as the Cartesian picture of isolated selfhood and the corresponding picture of private language would have it), that the desires are *turned* inward. At one point Wittgenstein writes: 'You learned the *concept* "pain" when you learned language' (*Philosophical Investigations* §384). If we try to get behind, beneath, or before language in characterizing what the underground man might do in creating his private moan-language by leaving the concept of pain out of the account, referring only to the sign 'S' in his diary and the brute sensation *and nothing more*, we need to be reminded again that the very concept of a linkage, an association between the sign 'S' and the sensation is itself a concept in, and not prior to, our shared linguistic practices. And if the idea of a nevertheless possible private language persists even beyond this late point—as perhaps the very last line of defense for the picture of private objects identified through Cartesian introspection that would serve as the meaning-giving referents of our psychological terms—we need to be reminded that the most fundamental term in any account developed along private-language lines is the *word* 'sensation'. And that, as a word of *ours*, is too already public. In §261 Wittgenstein thus writes: 'What reason have we for calling "S" the sign for a *sensation*? For "sensation" is a word of our common language, not of one intelligible to me alone.' Reinforcing the fact—a fact that is rather uncomfortable in some fashionable post-Cartesian quarters quick to celebrate what is there called 'interiority'—of the inescapably public nature of linguistic meaning against the recurrent image of mental privacy and the attempt to revivify it by getting behind or beneath even the concept 'sensation', Wittgenstein adds: '—And it would not help either to say that it need not be a *sensation*; that when he writes 'S,' he has *something*—and that is all that can be said. "Has" and "something" also belong to our common language.'

Every attempt to get to a position just before language brings language with it. Any such attempt—logically similar to the questions 'What was happening just before the beginning of time?' and 'What is just on the other side of the boundary of space?'—leaves behind the very conditions that render the question coherent. And so Wittgenstein adds next: 'So in the end when one is doing philosophy one gets to the point where one would like just to emit an inarticulate sound.' But we would

grasp at this strategy in the hope of giving a linguistic or verbal sign its inward referent by emitting inarticulate yet meaningful sounds that are not yet linguistic. The futility of this is apparent, and Wittgenstein adds: '—But such a sound is an expression only as it occurs in a particular language-game, which should now be described.' If meaningful, that sound is already inside of, and not prior to, a language-game.

Dostoevsky's full and fine-grained description of the underground man's groans is just such a description, and what it shows, indeed, is what Wittgenstein in the passages considered above says. Dostoevsky's *Notes* is a kind of writing that does cast a good deal of light on the nature of self-awareness, of self-scrutiny, and of (if in this case more than a little tortured) autobiographical attentiveness. Wittgenstein's writings work in their own underground as well, digging beneath the often unwitting presuppositions and foundational question-formulations that can powerfully shape our subsequent thinking. They both demonstrate that the inner life, given a thorough description by Dostoevsky and a thorough rethinking by Wittgenstein, is not what we may all-too-easily think it to be in accordance with a profoundly entrenched Cartesian legacy. It is, rather, what philosophical-literary investigations show it to be once description, in Wittgenstein's sense, has supplanted explanation, and once literature has been given its distinctive 'light' by philosophy. Dostoevsky's *Notes from Underground*, among other things, shows—in this respect like Wittgenstein's *Philosophical Investigations*—that, with regard to the constitution of the self, its reflexive mental experiences, and its projects of self-awareness, language is anything but an afterthought.

As we have seen, Wittgenstein said at one point that language is itself the vehicle of thought. And it is emerging that language, for parallel reasons, is what one might then call the vehicle of selfhood. The linguistic dimension of the self is, as we are increasingly coming to see, necessary (and not merely an *ex post facto* contingency) to its constitution. But then how, if we free ourselves of the picture of inner ostension and the correlated false criterion of inward correspondence as the means with which to verify a self-narrative, should we better understand the very question of true self-narration?

5

The Question of True Self-Interpretation

MANY find that the intuitive plausibility of the doctrine that has been aptly termed interpretive multiplism,[1] that is, the belief that there can be more than one ideally admissible interpretation of a cultural entity, tends to expand when we are talking about others and to diminish when we are talking about ourselves. The reasons for this asymmetry in our pre-analytical intuitions are complex, and extend into the philosophy of mind, the philosophy of language, and—most fundamentally—our vexing question concerning the nature of selfhood and the proper description of autobiographical consciousness. Although it is impossible to cover all these areas at once, in this chapter I will continue attempting to show some of the significance of Wittgenstein's writings on mind and language as they pertain to what we might call the self's autobiographical sense or awareness, and in particular to reconsider, in the light of Wittgenstein's remarks, not only the puzzle of true self-interpretation, but also the presuppositions, of a metaphysical–linguistic kind, that are too often embedded within that puzzle's very formulation. The task at hand will thus entail a close reading of those remarks of Wittgenstein's most pertinent to this particular topic—and the remarks that here again are too little considered both within Wittgenstein scholarship and beyond—that concern not just the nature of linguistic meaning and especially the nature of, as speakers, our knowledge of that meaning, but more precisely the nature of *retrospective* meaning, i.e. the very distinct nature of our knowledge, not just of what we mean, but of what we—from a position of hindsight (or, more accurately, a position of the future vis-à-vis that now past language)—*meant*. And the task at hand will entail a look into some actual practices of self-interpretation (it is by now clear that it was Wittgenstein who was most concerned among philosophers of language to turn to our practices to counterbalance

---

[1] Michael Krausz, *Rightness and Reasons: Interpretation in Cultural Practices* (Ithaca, NY: Cornell University Press, 1993), see esp. ch. 2.

the impulse to oversimplify and thus to theorize overgenerally); thus, in this chapter we return to the author of the most influential, and most philosophically sustained, piece of self-interpretation. Indeed, now passages of Augustine's *Confessions*[2] will prove as conceptually intricate, if in a rather different way, as Wittgenstein's: in some cases they will show the misleading power of the presuppositions regarding meaning in autobiographical understanding; in others they will show a mind beginning to break free from its own misleading metaphysical self-concept. But first things first: what can we say at present about the sources of the intuitive belief that interpretive singularism is the most plausible position with regard to self-knowledge?

First-person privileged access is a metaphysical picture of the self that has been, deservedly, subjected to a good deal of critical scrutiny. The philosophical picture of the mind, introspecting upon inwardly and directly observable contents to which it alone has direct access, is, as we have now seen from a number of vantage points, the problematic legacy of Cartesianism. And its immediate analogue in the philosophy of language will at this stage of our discussion not be difficult to identify succinctly: We, as speakers, have inwardly and directly observable access to the contents, indeed—as we saw this pernicious picture of language unfold—access to the meaning of our utterances. We also know that this dualistic picture of the mind and of language, *as explicitly stated*, is now expressly defended by very few; the overt enfilades of Ryle,[3] and the undercover work of Wittgenstein in meetings with his inner interlocutor,[4] changed—one hopes for ever—the reception of this generic picture of the self and that self's relation to its language. But, as we

---

[2] Trans. R. S. Pine-Coffin (Harmondsworth: Penguin, 1961).

[3] Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949).

[4] For an insightful study of Wittgenstein's employment of various voices within *Philosophical Investigations*, see Jane Heal, 'Wittgenstein and Dialogue', in T. Smiley (ed.), *Philosophical Dialogues: Plato, Hume, Wittgenstein* (Oxford: Oxford University Press, British Academy, 1995). An interesting question arises concerning Wittgenstein's use of (internal) dialogue in most of his later writings, namely, is that dialogical presentation only a contingent matter, i.e. merely the way Wittgenstein found himself most able to present his thoughts that might well have been put differently or more directly, or is the dialogical presentation *internal* to the content, and thus necessary to what those philosophical writings show and to how, in short, they work (as a 'machine to think with')? I for one tend strongly toward the latter view, but to argue the point directly would take us fairly far from present purposes (although I do think that, once the change in the way of seeing an issue that Wittgenstein's writings encourage has happened, one can look back and see how the dialogical progression of thought has been instrumental to that change—an issue I will reconnect with in Ch. 7, Sect. 3).

have also seen, its influence, or what we might call its residue, continues to be detectable;[5] my point here is that it is discernible in the intuitively plausible asymmetry between our self-interpretive singularism, our first-person belief in a single and ideally correct answer to any question of self-knowledge (here particularly *linguistic* self-knowledge, that is, knowledge of what we meant—and, equally importantly, of what we did not mean—on a given past occasion) and the multiplism we may accept with regard to the interpretations of others. It is, simply put, if in a still newer guise, the diehard idea of privileged access at work (as it so often is in philosophical problems of selfhood) behind the scenes, shaping the dualistic conception of introspection as we saw it from various points of view in the first three chapters. And here, in the present context of the question of true self-interpretation, we can see that this diehard idea exerts its influence (for us now obviously) in its presence but also, perhaps less obviously but no less significantly, in its absence, its denial. How so, precisely?

When facing the problem of true self-interpretation, three options readily present themselves to any reasonably categorically tidy mind. First, we might argue that there is, in any question concerning what we did and did not mean, a determinate interior mental event and that the truth of the matter just is a direct correspondence to this. And—crucially—it is here thought that we can know the determinate mental event of having meant one thing and not another (or a number of others, as multiplism as applied to linguistic meaning would have it) via direct introspection: we have this in mind when we utter the interpreted language, and we—with unique access—accurately recall having had it in mind when we later give the single correct interpretation of our original meaning. This, as the first categorical option, shows the influence of the dualist-introspectionist picture of the mind through its robust presence: the picture is intact, and one is affirming direct and privileged access.

Second, we might argue that, while it is true that there is or was a determinate mental event that constituted the true meaning of our questioned utterance, in fact we do not have direct introspective access to

---

[5] This claim is made by many of those working inside the Wittgensteinian tradition but also by those outside it. Recall, here again, the epigraph to this book by Donald Davidson ('Knowing One's Own Mind', in Quassim Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994)): 'There is a picture of the mind which has become so ingrained in our philosophical tradition that it is almost impossible to escape its influence even when its worst faults are recognized and repudiated . . . ' (p. 61).

it (as Freudians of a linguistic persuasion might argue), so we must face (and overcome through depth-analysis for the Freudians) a distinctive variety of skepticism: we know that there is or was such a determinate interior event constituting the meaning, but we also know that we cannot (at least directly or without a protracted and unguaranteed effort at self-interpretation) know it. This shows the influence of the Cartesian-introspectionist picture of the mind and language through a muted presence: the dualistic picture is intact, but one is denying direct privileged access.

The third option is not, of course, the absurd claim that we can know a determinate inner mental event that is not there, but rather the claim that the inner determinate event is not there and thus that there is no possibility of our knowing it. Here, as an old-fashioned reductive behaviorist might put the matter, there is no possibility of truth-as-correspondence-to-inner-facts precisely because there are no inner facts to which our interpretation might correspond. It is in this case that the dualist-introspectionist picture shows its influence through its absence: the *underlying structure* of the picture is intact, and it is only reinforced in the denial of the inner determinate meaning-event. The robust introspectionist, the skeptical introspectionist, and the behaviorist all share the structure, the fundamental, presupposed, structure of the question of self-interpretation. But a fourth option, a thoroughgoing rejection of the very conceptual structure of this question, might now seem an interesting possibility; from that perspective we might indeed be able to contemplate our autobiographically interpretive practices without the prismatic distortion of underlying dualistic theory. But this cannot be stated directly or succinctly: How would we formulate, with only the minimal tools before us, the question of true self-interpretation in a way free of the dualist-introspectionist picture that, at this stage of the discussion, we know so well? And how could we describe the significance of looking both to the details of retrospective meaning and to autobiographical practices *before* the fact? Perhaps there is a kind of true self-interpretation not envisioned by, nor envisionable within, these foregoing categories, and that may be one way of briefly characterizing what a close look into Wittgenstein's relevant remarks allows. If self-interpretive singularism has a lingering plausibility, this plausibility depends on our belief that, after all and regardless of what else is said, (1) we mean determinate things when we speak and (2) we must in some way know what it is that we mean. If those two elements are true, then—so we are strongly inclined (for reasons I've hinted at)

to believe—there is a *right* answer to any question concerning what we did and did not mean. But how should that particular form of rightness be characterized? With that question, along with our accrued backlog of considerations in mind, let us turn, in the next section, to Wittgenstein on meaning, then to Augustine on self-interpretation, and finally in this chapter to a reconsideration of the presumption embedded within the very structure of the question of true self-interpretation.

1. MEANING IN RETROSPECT

Guessing at how a word functions, Wittgenstein has said and shown, will not yield valuable philosophical results. The necessary task, as we have seen, is to '*look* at its use and learn from that' (*Philosophical Investigations* §340).[6] I have already insisted that a vast repository of usage of philosophically troublesome, and indeed sometimes troubling, concepts such as the self, indeed what he called the mysterious 'I', is to be found in literature (particularly of an autobiographical or semi-autobiographical kind); we might note that the traditional categorization of some literature as *philosophical* literature[7] implicitly acknowledges this fact. But Wittgenstein's next remark, immediately following the foregoing assertion, concerns *prejudice*, specifically the prejudice that stands in the way of taking his philosophical advice, of looking—with the right background, interests, focus, and frame of mind. And this prejudice, or more accurately these conspiring prejudices, are, as he said, 'not *stupid*', and the philosophical difficulty, as we have now seen on different parts of the philosophical landscape, lies in removing them. Wittgenstein, as we saw in the preceding chapter, made this remark in the context of his investigation into the question 'What is thinking?', but he might have placed it with equal aptness in his discussion, at the very end of part I of *Philosophical Investigations* (but also pursued throughout his other writings),[8] of what it is—and what it is not—to

[6] *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958).
[7] See e.g. the studies of *Middlemarch, Anna Karenina, The Brothers Karamazov*, and *Remembrance of Things Past*, in Peter Jones, *Philosophy and the Novel* (Oxford: Clarendon Press, 1975).
[8] It is, instructively, impossible to list *all* of the passages devoted to the analysis of meaning throughout Wittgenstein's writings, precisely because the multidimensional nature of the topic—extending not only into explicitly linguistic investigations but

mean something. For it is precisely the deeply seated belief, or again the pre-analytical intuition, that the mental phenomenon of meaning something not only underwrites self-interpretive singularism, but that it also requires, or indeed proves, the prior existence of an inner point of consciousness, an inner dualist–Cartesian self that is the private place within which the act, state, or process of meaning something occurs.

Wittgenstein's investigation into this phenomenon in at least one very clear sense runs exactly parallel to the foregoing investigation of thinking itself: meaning something is, on examination, found not to be at all what we expected when coming to the subject with certain philosophical presuppositions, and the subject does not reduce to a single uniform mental act, process, or state. He finds, here too, various phenomena, not a phenomenon, and—de-psychologizing this subject as well[9]—they are not *mental* in the way we might have expected. And this result, I want to suggest, pertains quite directly to our understanding of the nature of the self and its first-person autobiographical consciousness (although Wittgenstein did not draw this out explicitly either, and in this way too it runs parallel to the discussion of thinking), and removes one of the central prejudices against seeing the concept of, well, the self-aware self perspicuously, which is, naturally, a precondition for correctly understanding the very question of self-interpretation. But there remains, of course, the possibility that the belief in the inner-self-as-necessary-location-of-meaning-something is *not* a prejudice, but instead is accurate; as we follow Wittgenstein's discussion it will become clear that, along with his project of the de-psychologizing of our understanding of what it is to mean something, he retains a respect for first-person authority with regard to meaning something. In short, we find Wittgenstein, on this particular score, taking a 'middle way' between (or, perhaps better, showing a radically different alternative to) Cartesianism and its polemical partner, behaviorism; Wittgenstein's discussion, here again not reducible to or accurately presentable as a unified theory, escapes the misleading influences of the former without reducing the subject to the latter.

also into privacy, intention, rule-following, mental states, images, aspect-perception, memory, and so forth through an indeterminate list—gives it a highly mobile boundary. Thus, what is and is not relevant to the study, or the elucidation, of meaning within Wittgenstein's writings is determined (appropriately and as an illustration of its own larger message) by the context of the specific inquiry.

[9] For a fine guide to the Fregean work that stands behind Wittgenstein's de-psychologizing projects, see Anthony Kenny, *Frege* (London: Penguin, 1995).

The first thought that might well strike us when reflecting on the experience of meaning something is, as indicated briefly above, that we can easily *remember* having meant something, and this memory gives us the content upon which to introspect. Again, it is by now no secret that even this starting point is already a heavily freighted—a dualistically freighted—way of framing the issue; Wittgenstein thus begins his close scrutiny at precisely this point. In *Philosophical Investigations* §661 he asks, in remembering having meant *him*, if we are thereby 'remembering a process or state?' This is enough at least to unsettle the initial presumption that meaning something is a distinct mental process individuated within consciousness, and implicitly challenging this individuation, he follows it with the longer-pause-giving question 'When did it begin, what was its course, etc.?' Is, indeed, meaning something the *kind* of thing that *has* a course? If a process, it seems that it should. But there must be, or so we are strongly inclined to think, a process or mental act that constitutes our meaning something; what else could constitute this? Against this intuitive sense (where this intuitive sense is itself caused by an unwitting subscription to an underlying philosophical picture) Wittgenstein asks us, in §675, to imagine asking or being asked the question 'what was going on in you when you uttered the words . . . ?' and adds, pointedly, that the 'answer to this is not: "I was meaning . . . "'! A statement about what one meant by a word, he tells us shortly (§676), is not equivalent to a statement about 'an affection of the mind'. It takes little reflection to see that Wittgenstein's points here are sound: we do not answer questions concerning what we meant in saying *x*, or what was going on in our heads when we said *x*, by 'reporting' on inner processes—which in turn we find we are not anyway able to describe in terms of their beginnings, courses, and so forth.

But are these points compelling in any larger sense, beyond what initially appears to be their extremely limited scope? It seems not; that is, they are certainly not sufficient to dislodge the presumption, intuition, or belief that the Cartesian inner self is nevertheless behind, or required by, or presupposed by, or a precondition for, meaning something, precisely because we might well continue to hold that, despite how we would or would not answer such questions, we still mentally picture the person or thing we mean and that such a mental envisagement is precisely what 'meaning something' amounts to, is constituted by. But Wittgenstein has a ready answer here: 'If I say "I meant *him*" very likely a picture comes to my mind . . . but the picture is only like an illustration

to a story.' And from the mental picture *alone*, or from being told of the existence of such a mental picture, it would, one can quite readily see, prove mostly 'impossible to conclude anything at all', for only when one already knows the story that the picture illustrates, only when one is in a position to apprehend 'the significance of the picture', would one be able to posit the picture as the content of that which was meant—which is of course the wrong way round. Such a mental picture should *constitute* the act of meaning something, not *follow* it, if the explanation is to have force. Still, this issue too seems preliminary; even if the mental act of producing mental pictures is much more problematic than we might initially have thought in giving content to the phrase 'meaning something', does this really uproot the conception of the self behind it?

One might well insist that the matter of meaning something is in fact reducible to its essence and that this essence, most fundamentally, is a matter of *stipulation*. 'Meaning something' is, one might argue, less mysterious than it is beginning to look; 'by *this* I meant *that*' is as far as we need to go, and this ensures a correct—indeed *the* correct—answer to any question concerning what we really meant. But Wittgenstein worked extensively on the ordering and reordering of the entries that make up *Philosophical Investigations*, and it seems clear that this labor was often undertaken in order to capture, and to anticipate, the natural moves, the natural sequential unfoldings, of philosophical thought. It is thus not surprising that we find his well-known 'abracadabra' section next (*Philosophical Investigations* §665):

Imagine someone pointing to his cheek with an expression of pain and saying 'abracadabra!'—We ask 'What do you mean?' And he answers 'I meant toothache.'—You at once think to yourself: How can one 'mean toothache' by that word? What did it *mean* to *mean* pain by that word? And yet, in a different context, you would have asserted that the mental activity of *meaning* such-and-such was just what was most important in using language.

But—can't I say 'By "abracadabra" I mean toothache?' Of course, I can; but this is a definition; not a description of what goes on in one when I utter the word.

This section tracks a number of important shifts or movements of thought on this matter. First, it gives voice to the strong impulse to posit *stipulation* as the essence of meaning; the speaker shows this by saying and then briefly explaining his having said 'abracadabra'. Second, the strong competing intuition *against* stipulation as the content of meaning something is recorded next; when we are shown in an actual example what we might endorse in the abstract, that is, pure stipulation,

we react with a feeling of deep implausibility: *that* couldn't possibly go very far in explaining what it is *actually*, or indeed *ordinarily*, to mean something. Third, one wonders if the stipulation thesis, as shown to us here, even makes sense: what does it *mean* to mean that? Significantly, here we find yet another argumentative thread that is powerful but left unexpressed, or only implicit, by Wittgenstein: 'Meaning something' is precisely what is in question, and we might try—successfully or otherwise, as we have just seen in the second stage of this section—to explain it by stipulation. But then *that* stipulation, as we see, goes unexplained ('what does it *mean* to *mean* pain by that word?'), and thus we have at least the hovering threat, if not the grounded reality, of an infinite-regress problem in regard to stipulation. Yet, fourth, we find a reconsideration from the vantage point of the concrete case of what we may have posited in the abstract; 'in a different context' we would indeed have asserted the mental act of meaning something as the essence of the matter. In fact, at this fourth stage of movement, one seems to uncover again the seemingly inextinguishable plausibility of the stipulation thesis: can't we now just say we mean 'toothache' by 'abracadabra'?

And here, as the fifth and most important step, we see the *coup de grâce* to the stipulation thesis as an explanation, not of a momentary *encoding* of meaning, that is, determining momentarily to mean one word by another, but of the countless ordinary cases of meaning something. We *can* mean 'toothache' by 'abracadabra', but this is a definition, and not, most significantly, a description of the inner mental act, state, or process that occurs when we utter a word and mean something by it. Advancing mystery is the one thing that *is* clear at this stage: in regard to a mental event that allegedly constitutes 'meaning something', the matter is becoming increasingly mysterious. That which initially seemed obvious now looks implausible, and any thesis enjoying newly revived plausibility fails on closer inspection to provide the explanation we need. The truth of the claim we saw at the outset, i.e. that a statement 'I meant this by that word' is fundamentally different in *kind* from a statement 'about an affection of the mind', is becoming ever more evident. Still, Wittgenstein has not yet given us enough to relinquish the presumption, which, despite what has been said, at least *in the abstract* seems unavoidable in giving an account of what it is to mean something, namely, that there is some kind of mental object to which we give our inward attention when meaning something in particular.

2. THE PAIN AND THE PIANO

It must be with this sense of obviousness—a sense that is about to be unsettled along with the preceding cases—that Wittgenstein initiates the next part of the discussion by asking us to imagine that we are in pain while simultaneously hearing a piano nearby being tuned (*Philosophical Investigations* §666). In such a case, he reminds us of the obvious fact that it makes a great difference if, in saying 'it'll soon stop', we mean the pain or the tuning. But he brings in this contrast in order to focus the question more finely: 'but what does this difference consist in?' He admits—and here the admission suggests that the philosophical presumption we naturally bring to such a case can be at least in part right—that in many cases a directing of the attention corresponds to meaning one thing or the other; interestingly, he likens this to a case of nonperniciously construed introspection, where a particular way of shutting one's eyes might be called 'looking into one's self'. But the argument, brought out in the following section, is that, while there may (just as there may not) be correspondence to a particular directing of the attention, such a correspondence does not *in itself* prove either necessary or sufficient for meaning something. First, one can perfectly well *mean* pain in saying 'It'll get better soon' and yet be faking it; that the pain is simulated does not preclude the meaning of pain in saying that. One may say (reducing to the vanishing point the 'object' of pain to which one is allegedly directing one's attention, and which thus allegedly determines the content of 'meaning something' in this case) 'It's stopped now' and yet still perfectly well mean pain. One means pain, yet one is not, indeed one *could* not, be 'concentrating his attention on any pain'. And there is the parallel point in the honest case; if one genuinely has pain and then says 'It has stopped now', one means something without the inner referent.[10] Moreover (and Wittgenstein adds this a bit later, in §674), does one use the locus of attention and the strength of that attention as joint criteria for what we meant? Does one say, consistent with this thesis, 'I didn't really mean my pain just now; my mind wasn't on it enough for that'? Or do we ask ourselves

---

[10] This way of putting the matter, i.e. positing an 'inner referent', opens the way into the private-language argument and is, as stated, a conceptually dangerous formulation in that it could reinforce the very inner–outer dualistic picture that is being challenged by Wittgenstein's reflections.

what we meant by the words 'It'll stop soon' *because* our attention was divided between the pain and the piano? Wittgenstein admits that there can be a corresponding focusing of the attention, but that is never going to provide the ineliminable and definitive essence of 'meaning something.'

Wittgenstein offers, not surprisingly, a partial diagnosis of the condition of believing that *some* kind of pointing, in this case through directed attention, is necessary for meaning something; the diagnosis is familiar. While believing, or clinging at some conscious or other level to the intuition that we brought to the surface in preceding chapters, that ostensive definition is the essential element of meaning (Wittgenstein does not articulate this in the present discussion but it seems indisputably implicit at this point), one might observe that it is possible to 'refer to an object when speaking by pointing to it' (*Philosophical Investigations* §669). In such a case pointing is simply and unproblematically a part of the language-game.[11] But upon this base of incorporating verbal, behavioral, and intentional aspects, we then construct the analogy to the inner case, that is, 'And now it seems to us as if one spoke *of* a sensation by directing one's attention to it.' But, to encapsulate what follows, the analogy from the outer to the inner case is hardly sound. Again, one *can* direct one's looking or listening to a particular thing, and this *can* constitute meaning when the particular move of focusing attention in this way as a form of meaning something is legitimized within the circumscribed language-game. But this is *in*essential; in telephoning someone to say 'This table is too small' we may indeed point to the table as we speak. But the meaning is hardly *dependent* upon that; moreover, the hearer can obviously understand without witnessing either the physical pointing or any imagined inward directing of the speaker's attention as mental pointing. The analogy is not only unsound; it is here again misleading, in its power to establish conceptual expectations that need not in truth be fulfilled. The point is not, of course, that we do not in fact mean the piano in the one case or the pain in the other (that would evidently and disorientingly defy the obvious), but rather that we need not explicate either case of determinate meaning *in the dualistic terms* that seem to reassert themselves here so naturally.

---

[11] My discussion here rests on the much fuller treatment of the concept of the language-game as it is developed in Wittgenstein's later thought that I offer in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), 9–44.

Wittgenstein considers the cases in which we ask if someone really meant what was said, for example, 'When you were swearing just now, did you really mean it?' (§667), and this, he implicitly suggests, often does not concern 'meaning something' centrally in the way that it appears to do. Thus it may be asking, rather, 'Were you really angry?' Here again he invokes the innocent (i.e. non-Cartesian) conception of introspection as we saw it in Chapter 3 above, saying that the answer given to such cases is often a result of *this* kind of introspective reconsideration. And such answers fall along a wide spectrum: the examples offered include 'I didn't mean it very seriously', 'I meant it half-jokingly', and so forth. Similarly, and instructively (since the capacity to mean something is not literally halved in such cases; we still mean, if in a less unitary, or a less committed, way), we can say 'I was half thinking of him when I said that'.

That discussion leads directly to another vitally important section of *Philosophical Investigations*, in which again a layered or multistaged sequence of thought unfolds (§678).

What does this act of meaning (the pain, or the piano-tuning) consist in? No answer comes—for the answers which at first sight suggest themselves are of no use.—'And yet at the time I *meant* the one thing and not the other.' Yes,—now you have only repeated with emphasis something which no one has contradicted anyway.

Again, he has not disorientingly contradicted, or defied, the obvious and incontrovertible fact that we do mean one thing and not the other. So what, to put it bluntly, exactly *is* going on here? First, the fundamental impulse to locate the essential mental act of meaning something is again given voice. This impulse is *very* hard to quiet, not only for the reasons thus far discussed but also because the impulse, or the insistence, seems to protect the conceptually linked notion of the inner self from being reduced out of existence by a thoroughgoing behavioristic account; it protects against the sense of loss that any such reductive or eliminative account brings in its wake. And, because the phrase 'meaning something' is after all widely used, and indeed a universal experience, *must* there not be something determinate, unitary, and readily identifiable in which this meaning something consists? Do we not, to reframe the matter in recursive terms, mean something when we use the phrase 'mean something'? (Recall Wittgenstein's preceding question, 'What does it *mean* to *mean* this?') But, second, now that we have followed Wittgenstein's argument to this point, he observes

a new condition or reflective state, a new movement of thought: 'No answer comes.'

Well, Wittgenstein is dead wrong about that: answers do come—or rather they have. The various construals of meaning something we have seen thus far include (1) an easily remembered process or state, (2) a process that follows a course and upon which we can report, that is, an 'affection of the mind', (3) a mental picture constituting the something we mean, (4) an act of stipulation, (5) a focused directing of inward attention on the inward referent upon which we concentrate, (6) an act of 'pointing' analogous to pointing, and subsuming all the previous, (7) the essence, the necessary and sufficient condition required by, the words 'meaning something'—and they all fail; even though they all quite naturally suggest themselves in this problem area, they 'are of no use'. Yet there again comes a third, contrasting movement of thought, manifesting the irrepressible sense that there simply must be such an answer and that it is implicitly voiced within the explicit claim, 'And yet at the time [of speaking] I meant the one thing and not the other.' This is indeed the (incontrovertible) case, but the matter is not dropped here by an author most concerned to examine human practices, to 'look and see', precisely because, while he may want to 'leave the world alone', he most assuredly does not want to leave our *thought*, our thinking about that world, alone.[12]

And the problem (as intimated too briefly above), as we see in the fourth stage of this section, is not with the simple truth that we meant one thing and not another, but with our *construal* of that truth. In saying 'Yes,—now you have only repeated with emphasis' a claim that not a single element of what you have said thus far repudiates, we see that it is possible to retain the simple and undeniable truth of our meaning something while jettisoning all seven of the just-listed misconstruals or attempts to reduce the phenomenon in question to a formula or mentalistic essence. What is thus striking about the fourth stage of *Philosophical Investigations* §678 is that we are left with a sense that meaning something, as a human practice that occurs inside given language-games, is going to be clarified, or indeed *is* being clarified, through a layered process of confusion-removal. This section, like so many before it that we have examined in previous chapters, offers instructions in what *not* to think, but it also at this stage seems

---

[12] For a searching discussion of this issue in Wittgenstein's philosophy, see Jonathan Lear, 'On Leaving the World Alone', *Journal of Philosophy*, 79 (July 1982), 382–403.

to promise a positive aspect—a clear view of the practice as it lies before us. And with this it reinforces the sense that there may be no answer, *of the kind for which we have been looking*, to the question 'What does this act of meaning consist in?' And if there is no such answer *of this kind*, then the conceptual foundations undergirding our intuitive belief in self-interpretive singularism are seriously destabilized. Indeed, in this case of meaning something, just as is the case with the question 'What is thinking?', there may be only a different *kind* of answer, an answer—or a distinctive conceptual satisfaction—that is, again, the result of having obtained a perspicuous overview of the matter. And that is precisely the kind of satisfaction available from philosophical–literary and philosophical–artistic interpretation. But to the point at hand, we want to ask how exactly Wittgenstein follows up this fourth stage: how does he further isolate and illuminate the phenomenon of meaning something, from which he has already carefully removed veils of misconstrual?

The next topic that we encounter is that of the *connection* between the 'meaning' and the 'something', i.e. the mental connection that seemingly must link the mind of the speaker to the object, or thing, meant—the speaker's mind to the piano or to the pain. Here too, as one can now predict, we will see a seemingly inevitable philosophical presumption quickly unraveled. Noting that we do not ask of a speaker who has been cursing someone if the speaker is sure he cursed '*him*, that the connexion with him was established' (§681), Wittgenstein then gives voice to what he will shortly identify as the mistake of proceeding from this observation to the belief that the connection is thus 'very easy to establish, if one can be so sure of it!' But that, he suggests, is to presume too much. In the case of the piano and the pain, in saying that one was thinking of the piano, one is not thereby committed to saying that one thus had to *observe* that such a connection existed. And indeed—although here Wittgenstein's discussion is quite incomplete and can leave one with the mistaken impression that he is claiming that this might well *always* be the case—one can make the connection retroactively, a connection that did not at the time of speaking exist. Wittgenstein wants *not* to choose between (1) saying that such retrospective answers with regard to what one meant in a past utterance observe a pre-existing connection or (2) that such retrospective answers create a connection that did not theretofore exist. He says, simply (§682): 'Can't I say *both*? If what he said was true, didn't the connexion exist—and is he not for all that making one which did not

exist?' This, it must be said even by someone largely sympathetic with Wittgenstein's philosophical undertakings, seems strange, equivocating, self-contradictory, clearly erroneous, or any combination of these four serious failings.

The problem, as becomes increasingly clear through these final sections of *Philosophical Investigations*, is with the way in which the question is framed, particularly with its smuggled presumptions. *Given* that the idea is in place of a connection being a mental linkage that somehow reaches out from mind to object, Wittgenstein wants to say 'both'—although he might with greater help to his reader have said, more directly, 'neither'. What he does in the following passages is thus wholly to reconstrue the very concept of the connection. One strategy for making this reconstrual convincing (and he is up against the dead weight of deeply entrenched Cartesian or mentalistic presuppositions) is to show that we often take such retrospective answers with regard to past meaning *not* as reports on mental linkages but as hypotheticals concerning what *would* have been said had particular further questions been asked at the time of the initial utterance. 'They say, for example, that I *should have* given a particular answer then, if I had been asked' (§684). But, even though this 'does say something about the past', it is 'only conditional'—thus weakening (or beginning to shift the great weight of) the sense that the mental connection must be there initially. This point is amplified in the following examples, particularly where it is observed that where we give someone an arithmetic rule for the expansion of a series, it will be correct to say in response to a question, for example, 'Yes, I meant you to continue the series beyond 100', and yet the speaker, the giver of the rule, may have thought *no such thing*. Thus, highly significantly for our understanding of autobiographical language, the criterion for the truth of a claim concerning what speaker X meant will not invariably refer to the events in the mind of the speaker.[13] Some connection exists—he did after all mean something

---

[13] It is perhaps worth noting that the application of this point—that is, that the criterion for the truth of a retrospective-meaning assertion will not invariably be provided by events in the speaker's mind—to the arts of self-revelation, e.g. autobiography, autobiographical fiction, and self-portraiture, instructively blurs the line between historical fidelity or accuracy and revisionism in self-description. This may suggest that there stretches a continuum of intermediate cases between the much-debated polar extremes of *ex post facto* objective accuracy and (relativistic) revisionism and that, if we are sufficiently attentive to the nuances of retrospective meaning, we will see that we are not forced to choose between these polarized or oppositional positions. On the value of considering such intermediate cases, see *Philosophical Investigations* §122.

and not nothing or not just anything—yet it cannot be of the kind envisaged in the philosophical presumption behind the framing of the question. 'Is it correct for someone to say: "When I gave you this rule I meant you to . . . in this case"? Even if he did not think of this case at all as he gave the rule?' (§692). Allowing no false doubt to seep into the investigation, Wittgenstein answers his own question quickly and unambiguously: 'Of course it is correct.' And then he adds, in a phrase, the heart of the matter (and this is the reason behind his apparently strange equivocation above): 'For "to mean it" did not mean: to think of it.' Thus, again with a special significance for our understanding of the workings of autobiographical and self-descriptive language, there may be no particular and determinate mental act of thinking that gives the singular correct answer to what we did and did not mean by a given utterance. The conceptual foundations for self-interpretive singularism have not just been destabilized, they have collapsed.

There is a sense in which such retrospective answers regarding initial meaning can be after-the-fact reinforcements of pre-existing connections, and another sense in which they are, even though correct, created. Thus, Wittgenstein's reply above, 'Can't I say *both*?', is after all reasonable, if seriously misleading at that particular stage of the discussion. But it is, naturally, the broader context, the particularization of the language-game within which the utterance has force—in short, the *point* of the utterance—to which we must look for clarification of, and answers to, such questions. And indeed *meaning* will reveal itself not to be invariably at center stage throughout a consideration of numerous cases: 'Instead of "I meant him" one can, of course, sometimes say "I thought of him"; sometimes even "Yes, we were speaking of him".' And suggesting that the intuitive sense of necessity of the (mysterious) mental link, the seemingly essential connection, uniting the 'meaning' and the 'something' will diminish through such an inquiry into cases, Wittgenstein adds: 'Ask yourself what "speaking of him" consists in' (§687). That is, look to the nuances of the circumscribed games in which we use these words without philosophical presuppositions—precisely cases of the kind provided in literary, and specifically autobiographical, works (a prime exemplar being Augustine's *Confessions*).

But does all or any of this mean that such connections do not exist, that they are mythical? This, again, simply could not be the case: we do mean something and not nothing. The issue is indeed, as we can perhaps now see more clearly, one of removing layered misconstruals: 'Certainly such a connexion exists, only not as you imagine it: namely

by means of a mental *mechanism*.' And then, parenthetically reminding us again of the danger and power of misleading analogies, Wittgenstein adds: '(One compares "meaning him" with "aiming at him")' (§689). This position does not reduce to what is now widely discussed under the heading of externalism,[14] precisely because (what we will call within the oppositional structure internal) first-person knowledge on the part of the speaker is preserved. Yet it is certainly not purely any variety of post-Cartesian internalism either; 'meaning something' is, on this view, de-psychologized. One way to attempt to state the point generally, if only as a corrective—although this too possesses a strong power to mislead in that it allows the removal of first-person authority and to too great a degree approximates externalism—would be to say that the connections are in the language, or in the language-game, not in the mind of the speaker. But that is to put forward a far more succinct definitional thesis than Wittgenstein articulates in these final sections of *Philosophical Investigations*. What he does say, significantly—and in keeping with the claim I am making that literary and artistic investigations are indispensable sources for the kind of understanding Wittgenstein strongly suggests we need—is that '—All this points to a wider context' (§686). And the discerning examination of such contexts will prevent (or dispel after the fact) the philosophical confusions, impulses, and smuggled presumptions he has been exposing and diagnosing. Thus, there may be a fourth option as outlined in the introduction to this chapter above, one that rises from the ashes of the thoroughgoing rejection of the very conceptual structure that the robust introspectionist, the skeptical introspectionist, and the behaviorist all explicitly share. Thus, while the fourth option, as we shall see, in a sense rises from the ashes, it does *not* assume a form that is similar to that which it replaces. If, indeed (as we shall see below) the fourth option were a *theory*, succinctly expressed, it would

---

[14] Broadly stated, I refer to a position that manifests itself in the philosophy of language in the view that the meaning of what is said is dependent upon features, facts, or circumstances external to the mind of the speaker, and in the philosophy of mind that the phenomenological content of an experience similarly depends on features, facts, or circumstances external to the mind of the experiencing subject. (A related position arises in epistemology, where a person may be said to know a given thing without that thing being presently in, or in some cases available to, the mind of the knower; this would lead into the related examination of Wittgenstein's response to Freudian theory. In this connection, see Frank Cioffi, 'Wittgenstein's Freud', in Peter Winch (ed.), *Studies in the Philosophy of Wittgenstein* (London: Routledge & Kegan Paul, 1969), and Jacques Bouveresse, *Wittgenstein Reads Freud* (Princeton: Princeton University Press, 1995).)

most assuredly be anti-Wittgensteinian in character, and it would, as a theory of retrospective self-knowledge, surely elide the context-specific detail that is indispensably prerequisite to the attainment of an overview of relevant particularities. Without these assembled particularities, the fourth option would have no content other than a general (and thus in this way self-contradictory) gesturing to pragmatics. With them, the never completed awareness of contextualized particularities constitutes precisely the 'assemblage of reminders' of which Wittgenstein wrote, and the therapeutic value of them, of which he also wrote, is evident.[15] With this in mind let us look to—of all things—the telling particularity of the price of butter.

Butter can obviously rise in price, but it would be an extreme misconstrual to say that this is an *activity* of butter (or an inward process or inner state). 'Meaning something' is a phrase whose grammar can be much more easily, and far less obviously, misconstrued. The 'album' of the *Philosophical Investigations* was, again, carefully assembled. All of the foregoing—the entire discussion that collectively points not to a hidden linkage, connection, process, state, or act in the mind of the speaker but to a wider context—is neatly compressed into the final section (§693):

---

[15] A number of writers have paid special attention to the remarks Wittgenstein wrote that suggest a deep affinity between the difficulties and conceptual dangers inherent in philosophical investigation on the one hand and the achievement of self-knowledge on the other. Thus, the therapeutic value to which I refer would have both philosophical and personal senses. See e.g. James Conant, 'Putting Two and Two Together: Kierkegaard, Wittgenstein, and the Point of View for their Work as Authors', in Timothy Tessin and Mario von der Ruhr (eds), *Philosophy and the Grammar of Religious Belief* (London: Macmillan; New York: St Martin's Press, 1995), where it is argued that, within Wittgenstein's later work, 'the etiology of philosophical confusions is as complicated—and as difficult to survey—as our lives and our language. So the procedure of uncovering our individual confusions must remain a piecemeal one . . .' (p. 303). For a contrasting view (which houses a number of Wittgenstein's remarks of immediate relevance to present concerns, regardless of this debate concerning the similarity or difference between philosophical work and life-interpretive work), see D. Z. Phillips, *Philosophy's Cool Place* (Ithaca, NY: Cornell University Press, 1999), 46: 'But in saying "You cannot write anything about yourself that is more truthful than you yourself are," "Nothing is so difficult as not deceiving oneself," "If anyone is *unwilling* to descend into himself . . . he will remain superficial in his writing," and "Working on philosophy is really more like working on oneself," Wittgenstein is referring to difficulties in *doing philosophy*, difficulties in giving the problems the kind of attention philosophy asks of us.' It is in the problems—the *philosophical* problems—of the nature of the self, and particularly self-interpretation, that the two positions converge: gaining conceptual clarity about the (philosophical) concept of self-knowledge is part of the attainment of (personal) self-knowledge.

'When I teach someone the formation of the series . . . I surely mean him to write . . . at the hundredth place.'—Quite right; you mean it. And evidently without necessarily even thinking of it. This shews you how different the grammar of the verb 'to mean' is from that of 'to think.' And nothing is more wrongheaded than calling meaning a mental activity! Unless, that is, one is setting out to produce confusion. (It would also be possible to speak of an activity of butter when it rises in price, and if no problems are produced by this it is harmless.)

But we must return to the fundamental question raised earlier and, of course, central to any attempt to understand the self's past thought and meaning in the larger context of autobiographical or biographical self-interpretive inquiry: Does this investigation into the problem of 'meaning something', like the parallel investigation into the question 'what is thinking?', give us clear and final reason to abandon with full closure the broadly Cartesian conception of the self and that self's autobiographical consciousness with which we have now seen Wittgenstein interact in numerous ways? The answer is probably a hesitant No; hesitant, because these reflections, in de-psychologizing 'meaning something' and showing the construals to which Wittgenstein is opposed ultimately to be 'wrong-headed', do in one sense argue against the broadly Cartesian conception of the self, and this seems to suggest a Yes. And No, because it does remarkably *preserve* rather than obliterate first-person authority with regard to retrospective meaning, if of a kind very much unlike what we might have initially envisaged when coming to the problem of self-interpretation with the dualist-introspectionist conceptual structure, or philosophical picture, in mind. Again, Wittgenstein is charting a middle way—although we must bear in mind that his *via media* does not synthesize the polar opposites: it rejects their mischaracterizations and more accurately tunnels under, rather than strides between, them. And so a final important question also wants answering: Does anything intrinsic to the phenomenon—or phenomena—of meaning something necessitate the *preservation* of any Cartesian or post-Cartesian conception of the self? The answer to that, quite against one's natural first impressions and intuitions—the ones with which we began this inquiry into the substance of Wittgenstein's final sections of *Philosophical Investigations*—is, remarkably, No.

Richard Moran has put forward a very helpful and lucid clarification of the distinctive—indeed I would say unique (in a sense we will discuss in the next chapter)—authority we have in the first-person case concerning the content of self-knowledge (which I will read in the light of the question concerning our distinctive authority regarding

the meaning of self-descriptive utterances).[16] Moran rightly emphasizes *agency* rather than the passivity of the traditional inner-spectatorial picture, or the picture drawn from simple external object-perception but turned inward. He writes: 'A person is credited with first-person authority when we take the question of what he *does* believe to be settled by his decision as to what he *is to* believe' (p. 134). And Moran reminds us that a comprehensive understanding of self-knowledge should include a place for explaining 'how it is that a person can speak about his own mind, without appealing to evidence about himself, where that independence of evidence contributes to rather than detracts from the authority of what he says' (p. 135).

It is in the course of this discussion that Moran focuses on the phenomenon, as we metaphorically describe it, of 'stepping back' from the ongoing flow of our engaged mental lives. Moran links this concept to that of the exercise of freedom, and, as he points out, this conjunction may well initially seem a rather odd one: what does the assuming of a position of reflective distance have to do with an exercise of freedom or free will? But he nicely shows (through an insightful discussion of Sartre, the recounting of which would take us too far afield at present) that once the grip of the inner-spectatorial picture is loosened so that we can see the power of agency *in the very act of self-reflection*, then the convergence will strike us as far more intuitively fitting than we may initially (again, in the grip of Cartesian presumptions) have thought. Indeed he characterizes what is special about mental life in just this respect: we do not invariably sit back and reflect upon pre-existent mental objects, but we rather *make up* our minds in and through the act of active reflection. (This is the active sense of self-reflection to which we will return shortly in Chapter 6 in connection with Iris Murdoch.) Moran sees the emphasis on the active nature of this kind of self-reflection also clearly represented in the words of philosophers since Sartre, and these various formulations help to specify the philosophical content of the emphasis. He finds Christine Korsgaard writing, 'For our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question,'[17] and Thomas Nagel nicely inflects her point as well. Nagel writes,

---

[16] Richard Moran, *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001).

[17] Christine Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 93, quoted in Moran, *Authority and Estrangement*, 142.

The new data provided by reflection always face us, in other words, with a new decision. . . . the reflective self cannot be a mere bystander because it is not someone else; it is the very person who may have begun with a certain unreflective perception, or desire or intention, but who is now in possession of additional information of a special, self-conscious kind. Whatever the person now concludes, or chooses, or does, even if it is exactly what he was about to do anyway, will either have or lack the endorsement of the reflective view. . . . Given . . . that he is now self-conscious, anything he does will imply endorsement, permission, or disapproval from the reflective standpoint.[18]

With these observations in mind—both of which again emphasize the active power of autobiographical scrutiny, if not by that name—Moran is in a position to see that, as he puts it, the 'metaphor of "stepping back" from our current mental activities is a richer and more complex image than it may seem on the face of it' (p. 143). Given the 'new data' to which Nagel refers, and the judgmental power that the 'distancing' that Korsgaard describes generates, we find ourselves able, as Moran says, to, for example, resist the impulse to believe something (and for us let's make that of a self-descriptive kind) and thus resist the natural flow of action that would follow upon (or, I would add, actually enact on the level of praxis) the belief. And when we do so, or, as Moran puts it, when we suspend the force of some such impulse in an effective sense, then in that case our 'stance toward it makes a difference to what happens' (p. 144).

'This sort of reflection on oneself', Moran rightly insists, 'is not a matter of simply appraising one's state of mind' (p. 145); that would be, again, a too passive way to put the matter. What we should believe—and again I would, for our present context, add the phrase 'about ourselves in the partly creative, partly constrained acts of autobiographical self-engagement' (to which I will return in Chapter 6, Section 3)—is a question, Moran writes, 'that is answered by a decision or commitment of oneself rather than a discovery or some antecedent truth about oneself' (p. 145). And it is thus in this distinctive recursive sense that this process is an 'assertion of one's freedom' (p. 145). Just what it is to occupy this distinctive self-reflective and yet self-constitutive position, or what I have examined in previous chapters as the 'autobiographical situation', Moran succinctly captures in this way: 'To take oneself to be in a position to ask and answer this sort of question about one's belief or intention is to take oneself to be in a position to *make* something true in one's answering it'

---

[18] Quoted in Moran, *Authority and Estrangement*, 142–3.

(p. 146). Sartre, as he points out, in a memorable phrase claimed that 'there is no inertia in consciousness' (p. 151), and Nagel wrote, as we have just seen in his redescription of Korsgaard's point, that the 'reflective self cannot be a mere bystander'. The special—again, I want to insist, unique in the sense to be explained in Chapter 6—relation we have to our own thought gives refined content to the phrase 'first-third-person asymmetry', as Moran's discussion shows.

   The active, or what we might here call 'mind-making', character of this type of reflection Moran records in a memorable phrase of his own: 'This expresses a relation to one's state of mind that is exclusively first-personal and not shared by the best telepaths of our philosophical imagination' (p. 151). These imagined telepaths would, indeed, however acute as observers, always be spectators upon mental content that had a bounded identity independently *of them*—even if that content is relationally open inside the mind and memory of the first person (as we will investigate the matter in Chapter 7). Their relation to their own thought is *fundamentally* different from their relation to ours (as we have seen, from a different vantage point, in our discussion of Cavell in Chapter 2, Section 3, above). Like us, although they can step back in the actively engaged or 'mind-making' way Moran has described, they cannot wholly and detachedly step back from their own mental content in the way they are always already 'back-stepped' from ours. But because Moran's discussion, however helpful, proceeds in isolation from the kinds of detailed examples that would, on close inspection, show what he in more general terms wants to say as a higher-altitude account, we need to return, once again, to the particular case. So, with the additional cognitive stock provided by the discussions of Wittgenstein's underground maneuvers, of retrospective meaning, and of the issues covered in this section, we turn back, indeed, to an actively 'back-stepping' Augustine.

3. AUGUSTINE IN RETROSPECT

In his recounting of the famous episode in which he steals pears from a pear tree with a gang of juvenile acquaintances, we find the older Augustine reflecting—as we so often do throughout the *Confessions*—on the misdeeds of his youth. In this recounting, we find Augustine not directly stating what he was thinking at the time of his earlier misdeeds but—as we also see him often doing throughout the

Confessions—*speculating* on what he may or may not have had in mind. And, characteristically, he makes a distinctively significant theological-epistemological assertion: 'No one can tell me the truth of it except my God' (p. 52). Unable to subscribe, in this episode, to the view of robust introspectionism and perform the requisite act of retrospective introspection, he puzzles over his intentions: 'It is true that if the pears which I stole had been to my taste, and if I had wanted to get them for myself, I might have committed the crime on my own if I had needed to do no more than that to win myself the pleasure. I should have had no need to kindle my glowing desire by rubbing shoulders with a gang of accomplices.' And then he concludes, on the basis of this hypothetical concerning what he would have done if his desire for the fruit had been the central motivation: 'But as it was not the fruit that gave me pleasure, I must have got it from the crime itself, from the thrill of having partners in sin' (p. 52). This 'must' is obviously significant for the problem of true self-interpretation. Augustine is indeed reflecting on his own past as from the vantage point of another, indeed speculating on the self—on what his earlier words and deeds meant—as would (in the terms introduced above) the multiplist. And a moment later, he asks himself how he might explain his own 'mood' at the time of his thieving: 'How can I account for it? *Who knows his own frailties?*' (p. 52). This autobiographical skepticism is, then, clearly not indicative of robust introspectionism corresponding to the determinate mental-state conception of meaning, but neither is it, clearly, any kind of interior-denying behaviorism—he is not looking at his own behavior solely, as from the vantage point of another person, but rather is pursuing a grasp of his (hidden or presently obscured) intentions. But his voiced skepticism, 'Can anyone unravel this twisted tangle of knots?' (p. 52), does, as we have seen, have for him an answer: 'No one can tell me the truth of it except my God.' Here his philosophical thinking is clear: he believes, like the muted or skeptical introspectionist, that there *is* a single determinate fact of the case concerning what he meant or intended in committing the deed, but that it is hidden. His autobiographical writing would thus, in Moran's sense, be spectatorial and passive. Indeed, a bit later he articulates this position in connection with the Platonic distinction between moral appearance and reality, observing that 'the appearance of what we do is often different from the intention with which we do it', suggesting precisely that he is in the grips of the skeptical-introspectionist picture of the mind. And the manifestations of this picture are evident throughout: later he says, 'My heart lies before you, Oh my God. Look deep within.

See these memories of mine . . . ' (p. 77). Although God is here cast in the role of psychoanalyst (critics of psychoanalysis may say just the reverse), the underlying conceptual picture at work is identical to that outlined above in the introduction to this chapter.

Yet, intriguingly, he adds—and here we may see a mind beginning to free itself of its own misleading self-image—'and the circumstances at the time may not be clear' (p. 67), suggesting, in a brief brilliant flash of this text, that—in a fashion anticipatory of Wittgenstein's observations above—it may be circumstantial, contextual matters that determine meaning as much as, or more than, an (alleged) process of thought. Moreover, in searching for that determinate singular thought, we may be misdirecting our attention if we want to grasp what is significant for the determination of meaning as the core of self-interpretation. It may, in short, be a far more active process than initially envisaged. And as such, this passage points to the fourth option mentioned at the outset, the opening possibility of a thoroughgoing rejection of the conceptual structure, a rejection of the architecture (to which we will return in the conclusion to this chapter below) that would lead us to satisfy only the impulse to theorize and, in encapsulating a new position succinctly that merely replaces the one before it, answer to the demands of an underlying domineering philosophical picture of the mind or the self. A deeper, far less superficial conceptual satisfaction, precisely the kind of which Wittgenstein wrote in his various passages on the therapeutic nature of philosophy, would prove accessible only subsequent to the removal of distorting conceptual models, indeed only to a mind reflexively set free of its severely prismatic self-image.[19]

There are numerous further flashes of insight, in which, distinctively, Augustine sees, and says, that it is a particular model or conception of the mind that has been misleading him in his self-examination, making certain theses seem unavoidable or obviously true that in fact are merely epistemic mirages. He, for example, now comprehends in retrospect that he 'thought of evil not simply as some vague substance but as an actual bodily substance, and this was because I could not conceive of mind except as rarified body somehow diffused in space' (p. 105). And later again, he speaks of his wits having been blunted by a misleading

---

[19] The title of the fine study noted previously of Wittgenstein's development, David Pears's book *The False Prison* (Oxford: Clarendon Press, 1987), nicely encapsulates the kind of intellectual incarceration to which I here refer; in the matter of self-interpretive retrospective reflection the conceptual imprisonment can be unwittingly self-imposed.

presumption—not unlike Wittgenstein's analysis of the 'not-stupid prejudice' above—that precluded clear thinking:

My wits were so blunt and I was so completely unable even to see clearly into my own mind, that I thought that whatever had no dimensions in space must be absolutely nothing at all . . . For my mind ranged in imagination over shapes and forms such as are familiar to the eye, and I did not realize that the power of thought, by which I formed these images, was itself something quite different from them. (p. 134)

Yet this is layered in its significance: on one level it is a mind, within the course of a sustained effort of philosophical self-interpretation, again struggling to free itself from its own false self-image in a way consistent with (and again anticipatory of) Wittgenstein's analysis of retrospective meaning. On another level he does still—and this is what makes it a *struggle*—write of his inability to see clearly into his own mind, which revivifies the dualist-introspectionist picture at issue.

There are, of course, contexts in which one may ask oneself if it is not true that one was doing a certain thing at a certain time for a reason that was unknown or (perhaps more interestingly and like Wittgenstein's example of 'half thinking of him' above) half-known to oneself at the time; the very question of true self-interpretation often arises in just such cases. Augustine naturally provides a number of these. In one place he examines cases of delays in closure, gratification, or consummation of events, determines the varying depths of the delayed gratifications, and suggests a simple possible correlation: the more protracted the one, the deeper the other. Involuntary cases of delay show the principle: the victorious general experiences the greater triumph after a lengthier and more dangerous battle; sailors experience more profound joy at their survival after a terrifying storm in which their fear of impending death was, in a causally linked way, just as profound. And, somewhat strangely, we are, he observed, happier at the partial health of a friend who, at first extremely ill, is now in partial recovery than we were at the full health of the friend prior to the illness (p. 162). In such involuntary cases, no one speculates as to whether the delay, itself causally determinative of the depths of the subsequent joy, triumph, happiness, and so forth, was in some sense deliberate. But in cases in which persons prepare for their eating and drinking with (unnecessitated) hunger and thirst, or where 'drunkards eat salty things to make their throats dry and painful so that they may enjoy the pleasure of quenching their thirst' (p. 162), or where persons 'who are engaged to be married . . . delay

the wedding for fear that the betrothed have not "suffered the trials of a long courtship" ' (p. 162), we might well have an open question concerning the motivation. And the questions of this kind, asked of one's self, indeed strongly suggest both that the robust-introspectionist model cannot accommodate such cases (and such cases are central to genuine self-interpretation) and that the behaviorist model would eliminate out of existence the very possibility of a half- or unknown meaning to our actions. Is then the muted or skeptical-introspectionist model supported here? No, precisely because the cases of the kind in question preserve the possibility not only of there having been a hidden intention (which 'half-known' cases *would* suggest) but also of after-the-fact explanations arising not from the determinate thought of the speaker or actor but rather—again as Wittgenstein's analysis leads us to see—from the context, the very circumstances which, as Augustine says, may not initially be clear.

Many further cases show this, not only in Augustine's autobiographical practice but also in the episodes of others that he relates. There is a famous episode, specifically that in which Augustine's mother, Monica, is transformed by an angry word. Her servant-girl, in the context of the quarrel, calls Monica a 'drunkard': as it happens, Monica took daily sips from the wine-barrel, initially only a drop but gradually increasing to a semi-inebriating quantity. Monica herself had not, during the course of this advancing disease, as Augustine describes it, been aware of the progression. With the one 'harsh word of rebuke', and—significantly—intended as 'a most bitter insult'—the servant-girl, with words cutting like a 'surgeon's knife', corrected (in this case, by inculcating self-awareness) Monica in a single stroke. Any full account of what transpired would include the epiphany-like suddenness of self-awareness that the rebuke occasioned; this would indeed be one aspect of the meaning of the linguistic action. Yet it was not intended, *in terms of a determinate mental event*, as a corrective, but rather, as Augustine precisely specifies, 'to provoke her . . . not to correct her' (p. 194). We cannot know how the servant-girl would have replied to the question 'But even so, was it not in the larger context clearly within the scope of your idea, your remark, that it should or could bring about a new self-awareness on the part of Monica, and that this would reform her?' But in very many similar cases in our ordinary experience we can know: the answer is Yes. One might indeed say that in one way such a meaning was clearly not intended, but in another way, and at the same time, it was; it is likely that doubled answers of this kind indicate that the

model or conceptual picture we are presupposing in the formulation of the question is either misleading or insufficiently subtle to capture the nuances of our self-descriptive practices. And in the larger context of Augustine's autobiographical practice, these models and conceptual pictures embedded within questions of self-interpretation fare no better, in the end, than they did within the context of Wittgenstein's investigations into retrospective meaning.

Let us stand back and take a broader view of the ground covered in this chapter. We began by observing a presumption that one very often, if not invariably, encounters in those questions in which the subject takes itself as its own object of investigation: that is, questions of self-interpretation. That presumption, variously manifested in contrasting positions, for example, that of the robust introspectionist and that of the skeptical introspectionist, was, simply stated, that in any such question there will be a determinate and fixed mental event of meaning that originally determined the absolute and singularly correct answer to any question concerning what we meant in a given utterance or action (or a pattern of action over time, a subject to which we will return in the subsequent chapters). And there was, I suggested, a conceptual template or picture of the self beneath this presumption that the mental-event-denying behaviorist also shares, that is, the fundamentally dualistic ontological separation of outward verbal or gestural behavior from the hidden inward mental actions or content that gives the meaning to those outward, physical actions.

But this presumption, however intuitively plausible it may seem at the outset of any inquiry into the nature of self-interpretation, did not survive the scrutiny given it within Wittgenstein's investigation into the initially seemingly curious nature of retrospective meaning. Indeed, we now, given those reflections and observations on the facts of our practices, have reason to believe that the presumption is misplaced and that, as Wittgenstein puts it, the logic, or the philosophical 'grammar', of meaning *is* very different from the logic, or grammar, of thinking. Similarly but more deeply, the dualist-introspectionist conceptual picture upon which this presumption is founded and which undergirds the very formulation of the question of singularism in true self-interpretation is similarly unearthed (although its *full* excavation, analysis, and removal could not be completed here or indeed in any one chapter; this is a larger project that weaves itself throughout this book, appearing in differing ways from different points of view—or on different sectors of the landscape). And we saw, if only in brief, how the emphasis on the active

power of self-reflection as seen in Moran and others can shake the hold of the inner-spectatorial picture that is part and parcel of singularism.

Our turn to practices, to Augustine, showed two things: first, we saw cases in which a remarkably powerful philosophical mind was laboring under the self-misapprehension that the presumption enforces; second, we saw that powerful mind beginning, in places, to break free of that presumption. The false presumption, what we might in shorthand call 'mental-event singularism', in truth survived neither Wittgenstein's philosophy nor Augustine's autobiography. But does this mean that we should enthusiastically embrace a parallel 'mental-event multiplism', that is, the view (indeed hinted at in a few places within Augustine's text) that the question of the meaning of an utterance or action will have multiple true answers because the utterance or the action is an outward manifestation of multiple inward mental events? The answer is a resounding No: that would be to remain a captive, to stay within the misleading conceptual template or picture from which reflections such as those reconsidered here might free us.

There is once again a real sense in which Wittgenstein's observations, since they are significant for self-interpretation, tell us with ever-increasing precision, or with increasing conceptual magnification, what *not* to think. Should we not, at this fairly late stage, demand another, better, conceptual picture to replace the one that these considerations would remove? The answer here too is a resounding No—although this is perhaps less immediately clear. For the impulse to theorize is not easily diminished. The robust introspectionist has a bold and clear thesis to advance: simply look inward, transparently and immediately, at the meaning-content of the utterance, and report it accordingly. The skeptical introspectionist similarly has something, with equal boldness and only slightly less concision, to say: look inward, but with an awareness that the mental meaning-content, although wholly determinative of the significance of the overt utterance in question, may be initially obscured from our inner view, our introspective gaze. These twin positions (they share the same conceptual parentage in the foundations of Cartesian dualism) posit, indeed take as a given, that the determinative mental meaning-content is the *kind* of thing for which we should be looking in any question of retrospective verbal self-interpretation, and that, once found, the singular truth will be unproblematically in front of us. And the behaviorist, something of a younger sibling (with the same conceptual parentage, but deriving its identity from its direct and forceful opposition to the older pair), also has a perhaps even more

bold, and now antithetical, thesis to advance, with bracing concision. Asking for a new, concise, fourth conceptual model or picture to replace the ones that the foregoing considerations have removed—one that would display an equal measure of boldness and concision—is to ask to satisfy a philosophical desire we would do better to quell. It is to satisfy an impulse, of long and distinguished standing in philosophy, to preserve the edifice of an explanatory structure by replacing any removed element with a newly chiseled element that takes the same place. But there is an alternative, of, again, a radical kind, shown in the collection of strategies Wittgenstein developed within his middle and late philosophical writings (he referred to 'a new method' that had been found) to the incremental restoration of conceptual architecture. Wittgenstein also wrote of razing to the ground such structures and removing from our conceptual landscape the impediments to a clear view of our practices, in our present case, of the self-descriptive practices that would, if assembled into a perspicuous overview or a conceptual mosaic of particularities, dissolve philosophical puzzlement.[20] This *kind* of philosophical solution—indeed dissolution—is *very* unlike a new, sharply cut theory to advance in place of those which have been supplanted. And one finds such a position—one of a very different kind that pragmatically turns to the particularities of our practices within actual contexts of self-investigation—satisfying only *after* the investigation, which in this case involves the close reading of Augustine but might also include any other of a vast number of philosophically relevant autobiographies, memoirs, and any other form of self-interpretive writing containing retrospective meaning-determination.[21] Fittingly, the proof

[20] One might usefully reconsider the methods employed by John Wisdom in his once widely discussed *Other Minds* (Oxford: Blackwell, 1952), work that is still of considerable value in diagnosing the motivations to picture the mind in a way still very much in evidence in contemporary philosophy. In connection with these methods, see also R. W. Newell, *Objectivity, Empiricism, and Truth* (London: Routledge & Kegan Paul, 1986), esp. ch. 5: 'Reason and Particular Cases: John Wisdom'. A more recent study, showing both an awareness of the great value of the Wisdomian variety of patience I am endorsing and the relation between particularity and clarity, is Paul Johnston, *Wittgenstein: Rethinking the Inner* (London: Routledge, 1993).

[21] There are many recent volumes of and about autobiography and memoir, and no sign of the slowing of material relevant to the attainment of an overview of our multiform self-reflective practices. See, for a few examples (here again), Jill Ker Conway, *When Memory Speaks: Reflections on Autobiography* (New York: Alfred A. Knopf, 1998); Leigh Gilmore, *The Limits of Autobiography: Trauma and Testimony* (Ithaca, NY: Cornell University Press, 2001); and Paul John Eakin, *How Our Lives Become Stories: Making Selves* (Ithaca, NY: Cornell University Press, 1999). For a collection of papers from

of any such claim will be, indeed, in the pragmatics. Augustine realizes in retrospect that his own beliefs—in his case theological, but the point applies across the board to any self-interpretive question concerning love, affection, commitment, life-choices, and so forth—were put to the test *in extremis*, at the death of his mother (p. 200). His autobiographical report on what he did and did not believe was verified not in reference to an inner mental event but in reference to the pragmatic, circumstantially situated and personally engaged belief that manifested itself within and throughout the emergent patterns of his actions, words, hopes, fears, aspiration, regrets, and many other irreducibly human events and experiences. And in some cases, the truth—as in this case—is singular: his belief, as he expressed it most succinctly, was *real*. In other cases, the truth is multiple, in the sense that the meaning, the significance, of a given utterance or action will have multiple trajectories with regard to emergent patterns of actions or linkages to other utterances and actions. And in some cases, only seemingly paradoxically, the single truth is that one is genuinely divided or ambivalent,²² or an utterance or action is genuinely ambiguous. But what all these answers have in common is that they are—as Wittgenstein says and Augustine shows—true *not* by virtue of a verified correspondence to a determinate meaning-giving mental event, but rather, true on the level of pragmatic *use*. And we can only see the philosophical significance of those practices clearly

various disciplines on the relation between self-narration and selfhood, see Ulric Neisser and Robyn Fivush (eds), *The Remembering Self: Construction and Accuracy in the Self-Narrative* (Cambridge: Cambridge University Press, 1994). For a collection exploring various ethical considerations, see Paul John Eakin (ed.), *The Ethics of Life Writing* (Ithaca, NY: Cornell University Press, 2004). For a recent study showing the need for, and providing, a significant expansion of the personal-identity problem, see Marya Schechtman, *The Constitution of Selves* (Ithaca, NY: Cornell University Press, 1996).

²² At one point Augustine provides a perfectly lucid, and perfectly human, example of this genuinely divided ambivalence. He likens his thought at one stage of his life to 'the efforts of a man who tries to wake but cannot and sinks back into the depths of slumber. No one wants to sleep forever, for everyone rightly agrees that it is better to be awake. Yet a man often staves off the effort to rouse himself when his body is leaden with inertia. He is glad to settle down once more, although it is against his better judgment and is already time he were up and about' (*Confessions*, 165). The truth, of course, is that in practice *both* seemingly antithetical claims—he does and does not want to arise—are correct descriptions of such a person, and as a description of a person in a state like Augustine's at that particular moment are (seemingly paradoxically), to use the terminology in play presently, singly multiply true. See also Augustine's similarly intriguing descriptions of a mind that will not obey itself (pp. 171, 173); it only appears to be 'the strange phenomenon' Augustine says it is when in the grip of a misleading picture of the self.

if our minds are freed from distorting presumptions and misleading overgeneralized conceptual pictures.[23] In that sense, being shown what *not* to think is, remarkably, at the same time being shown what to think—without the imposition of another conceptual picture to replace the one we must labor to remove. Indeed, this too-little-discussed stretch of Wittgenstein's philosophy might successfully serve to remove a particularly pernicious conception of self-interpretation that in practice proves, somewhat ironically, only self-defeating. And as such, these considerations may in practice prove, fittingly, to be indispensable tools in the structuring of the mind's true image of itself.

In earlier chapters we considered the role images of selfhood play in shaping our thinking about autobiographical knowledge, where we similarly examined a number of issues directly pertinent to our understanding of the nature of, or in Wittgenstein's sense the 'grammar' of, introspective and self-revelatory utterances. And in the past two chapters we investigated a number of the ways in which we can construe the act of thinking itself, along with considerations of retrospective meaning. It is now time to turn directly to the distinctive character of our perception of persons, and to what this shows us about—as the content of our autobiographical awareness—our perception of ourselves.

[23] One might usefully compare here Nietzsche's characteristically extreme and re-orientingly insightful remarks on the mind's impulse to create false images of its own workings, what he calls 'the antecedentia of action', in *Twilight of the Idols* (Harmondsworth: Penguin, 1968), esp. 'The Four Great Errors', §§3–5, pp. 48–52. These passages offer, I would suggest, a rather blunt diagnosis of how we can eventuate in a 'haunted' condition in which we puzzle over the hidden real meaning of an earlier utterance or action that would (allegedly) provide the singular self-interpretive truth.

6

The Uniqueness of Person-Perception

W<small>E</small> now know that to say of Wittgenstein that he is an anti-Cartesian involves more than a bit of misdescription: that term correctly describes the behaviorist, who shares the fundamental metaphysical categories of the inner and the outer, the private and the public, with the Cartesian, but denies the reality of one category. It would be equally misleading to call Wittgenstein—without any further or more exacting specification of what is meant by the term—an anti-behaviorist: that term, conversely, correctly fits the Cartesian. Wittgenstein's position, or his overview of the language-games holding significance for our concept of selfhood and self-revelatory language, cuts beneath the metaphysical presuppositions of both of these polarized, antithetical theories of the self. We know that the Cartesian argues for, or is under the influence of, a picture of the self in which the ultimately private and inner point of consciousness, as the first given of human existence, is introspectively knowable instantaneously, transparently, and without mediation. This Cartesian picture, we also know, holds the greatest significance for our understanding, indeed our conceptual modeling, of person-perception: other-minds skepticism, and its logical extreme, solipsism, are its natural corollaries. On the Cartesian view, we *infer* from outward signs (signs that are on this picture only contingently associated with the hidden inward events and contents of private consciousness) that one or another inner state is present in the mind—indeed appearing on the private inner stage—of the person we perceive. It is under the influence of this picture that we find it plausible to utter metaphysical claims (in some cases false, in others incoherent) of the 'We can never really know the contents of another's mind' kind. Educated guesswork ('educated' because of our past familiarity with the observable behavior or external signs to which the hidden cognitive and affective content of consciousness is allegedly linked) would, on this picture, be the best we could do.

And we know that the behaviorist, placing an antithesis squarely against the Cartesian's thesis and thus sharing the identical unexamined presupposed metaphysical categories of inner and outer (where ontology itself is believed to keep the gulf between the two impassable), argues for, or under the influence of, a picture of the self in which the first given fact of human existence is external, physical, behavior. This too, clearly, holds great significance for our conceptual modeling of person-perception: humans are organisms that exhibit patterned responses to overt stimuli, and any speculation concerning the hidden interior is on any explanatory level excess baggage dragged aboard by the self's mythological, over-psychologized misconstrual of itself. On this picture any stated experience of an interior is epiphenomenal at best, pseudo-self-descriptive nonsense at worst. And it is under the influence of this picture that we find an exclusively externalized model of person-perception, i.e. one in which our knowledge of another person (if Cartesianism has the 'other mind' problem perhaps we have in behaviorism the 'other body' problem) is given its content wholly and only by the amalgamation of successively accumulated slices of overt behavior.

We have, in all of the preceding chapters, traced a good number of the multiform influences on our thinking that would lead us in one of these polarized directions or the other (or both, if we are large and contain multitudes), showing that these twin positions contain far more incoherence (on the side of Cartesianism) or misconstrual (on the side of behaviorism) than it would initially appear, and than it indeed did appear to many (before the private-language considerations and their many intimately related writings, e.g. on avowals). Wittgenstein, as we have seen now in a number of philosophical settings or locations on the larger landscape of self-understanding, observed that the mental and the physical fall into different language-games, and that the language-games of physical objects and of sense-impressions are different.[1] He wrote, as we saw at the close of the first section of Chapter 3, of a 'complicated relationship' between them, and he there proffered the warning: 'If you try to reduce the relations to a *simple* formula you go wrong.'

This inducement to accept the complexity of human phenomena without yielding to the impulse to simplify, to reduce, to theoretically straighten the crooked timber, fits well with everything related to selfhood and person-perception that we have seen Wittgenstein say,

[1] Ludwig Wittgenstein, *Philosophical Investigations*, 3rd edn, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958), II. v, p. 180.

e.g. that avowals, as a way of understanding first-person expressive utterances, relieve us of the burden of the false picture of self-expression as the self 'reading off' from inward facts that are only immediately observable by that self, or that self-expressions invariably involve an inner object and its designation. And this inducement fits well with many of the broader features of Wittgenstein's methods, his reminders, his variegated encouragements to change our way of seeing a philosophical problem, e.g. (1) his method of showing more capaciously with examples that which cannot be said succinctly; (2) his repeated reminders that words are deeds, that human actions (of the nonatomistic kind we will discuss in connection with Iris Murdoch later in this chapter and with Donald Davidson in the next) come first in any accounting of the person, the human subject, and that actions, like words, occur within a stream of life, an expanded contextual field; and (3) his encouragements to maintain a constant vigilance concerning the dangers of facile overgeneralization and doctrinal rearrangements of the investigated phenomena, always giving actual experience priority over the impulse to reorder multiplicities into unities. We might call this a fidelity to the nuances of lived experience. And most strikingly, when, in moving to another author whose writings hold a good deal of significance for our larger comprehension of autobiographical knowledge, we look to a number of genres of Goethe's extensive writings, including poetry, plays, novels, autobiographical writings spanning a number of volumes and periods of his life, and his critical writings on art and literature (which I will emphasize here), one finds not only a broad affinity, but indeed—if we look closely and with a measure of patience to the particularities in his texts—something approximating an enactment of Wittgenstein's conceptual clarification. To put it succinctly—albeit perhaps too succinctly—Goethe in a number of places and in a number of ways shows what Wittgenstein says. But let us begin our look at Goethe with the man himself.

1. THE CASE OF GOETHE

In a letter from one of his friends, F. H. Jacobi, it was stated of Goethe, with brevity and power: 'This man is autonomous from tip to toe.'[2]

---

[2] Walter Kaufmann, *Discovering the Mind: Goethe, Kant, Hegel* (New York: McGraw-Hill, 1980). In preparing this section I initially approached Kaufmann's chapter on

But any reader of Goethe's works will quickly learn that the autonomy referred to is anything but a philosophical or indeed metaphysical autonomy. In fact we know this from another letter describing Goethe, this one by Johann Christian Kestner, in which Goethe's contextualized, non-Cartesian autonomy is elucidated. Kestner describes Goethe: 'He is violent in all of his emotions but often has a great deal of self-control. His way of thinking is noble; free of prejudice, he acts as he feels without caring whether others like it, whether it is the fashion, whether the way one lives permits it. All compulsion is hateful to him.' And later Kestner emphasizes, 'he is not what one calls orthodox. But not from pride or caprice or to make an impression.' (Kestner was the very man who was within a year to marry Charlotte Buff, with whom Goethe was also passionately in love, Buff inspiring *The Sorrows of Young Werther*.)

Goethe's autonomy is sufficiently deep to allow not only a disregard for the orthodox, but also a disregard for the impression unorthodox thought, work, and behavior makes. And—still further like Wittgenstein, although my point concerns the anti-metaphysical nature of Goethe's autonomy, not the characterological similarities between the two—reminiscent of Wittgenstein's conversations on religion with Drury, Malcolm, Rhees, and a few others, Kestner adds the words 'About certain very important issues he speaks to few and does not like to disturb others in their calm ideas.'[3] But the autonomy Goethe exemplified was realized *within* a community of social actions, expectations, norms, and mores, and not separate from, or intelligibly *separable* from, that social context. This necessity of social context with regard to Goethe's personal autonomy is symptomatic of the fundamental truth concerning the self and self-expression in both words and actions that he showed if not always said, throughout his writing: 'In the beginning was the deed.'[4] And in Goethe's practice, just as in Wittgenstein's, well, not theory, but rather the *undermining* of theory, we find him creating and developing characters not through, as Walter Kaufmann observes,

---

Goethe for an overview of his works in relation to questions of the mind but expected little of direct significance linking Goethe and Wittgenstein. What I found was a goldmine of observations, quotations, passages from correspondence, etc., very much of which was directly relevant to the linkage between the two authors. So it can readily be seen how deeply indebted to that discussion I am here—yet Kaufmann does not mention Wittgenstein in the context of his writings on Goethe.

  [3] Ibid. 14.
  [4] For an insightful Wittgensteinian discussion of this phrase, see Peter Winch, 'Im Amfang war die Tat', in *Trying to Make Sense* (Oxford: Blackwell, 1987).

the enumeration of qualities but through the depiction of actions.[5] Nor is a self in Goethe's writing ever described, or given a pre-contextual essence, in advance of the actions, the deeds—again behavioral (in the broadened sense) *and* verbal—which constitute it.

In his 1810 *Theory of Colors* Goethe wrote:

We really try in vain to express the essence of a thing. We become aware of effects, and a complete history of these effects would seem to comprehend the essence of the thing. We exert ourselves in vain to describe the character of a human being; but assemble his actions, his deeds, and a picture of his character will confront us.[6]

Kaufmann remarked: 'To know Faust's mind, we observe what he says and does; and to create Faust's mind, the poet invents speeches and deeds for him. But that means that we can dispense with the concept of mind as an entity'—and Kaufmann makes clear that he means by that a Cartesian dualism of mind and body, where in this context too (recall the epigraph from Davidson) the mind is construed as a private point of consciousness inwardly knowable only to and by itself via pre-contextual or pre-relational introspection. Kaufmann puts the point in positive terms as well, i.e. telling us how to think, as well as how not to think, of the concept 'mind': 'Mind [for Goethe] becomes an inclusive term for feeling and intelligence, reason and emotion, perception and will, thought and unconscious.'[7] The idea of the mind becomes, as he suggests, 'a kind of shorthand'—which is perhaps not the worst way of summarizing a broadly Wittgensteinian view of the matter. Similarly, Kaufmann states that he, like Goethe, is concerned with 'the whole mind', i.e. all the extensive landscapes of inquiry that are misleadingly made to appear unitary and essentialistic by the single word 'mind'. Kaufmann, with similar resonance to the later Wittgenstein but again without mentioning him, also pointedly observes: 'the question whether *showing* something is inherently inferior to proving it or deducing it from pure concepts is part of Goethe's legacy'.[8] To take one quick example that itself could be extended into a full study, Faust's repeated self-deceptions and willful misdescriptions of his situation have continually to encounter the too-clever analyst in Mephistopheles; Goethe *shows* the epistemic fabric of self-deception, rather than attempting to unify the admittedly curious but multiform

---

[5] Kaufmann, *Discovering the Mind*, 23. [6] Ibid.
[7] Ibid. 24. [8] Ibid. 25, my italics.

mental phenomena of self-deception into the 'S knows not-P and yet believes P' form. (It should be mentioned at this point, incidentally, that it is also Mephistopheles who says, in giving epistemological advice to a student, 'Yes, stick to the words at any rate; | There never was a surer gate | Into the temple Certainty'[9].)

Wittgenstein has said that it is only in the stream of life that words have meaning,[10] and Goethe both knows, and shows, this easily forgotten truth. Here again the reader of Goethe's works will recognize his repeated observances of this fundamental truth, in varying genres and on different scales. On the largest scale, he shows it in the design of the first *Bildungsroman*, *Wilhelm Meister's Apprenticeship*, in which the knowledge of the *development* of a self constitutes the very substance of our knowledge of that self. On a smaller scale, he frequently emphasizes the process, and not the product, of the self-development he experienced in his poetry, and on a still smaller scale he shows a sustained concern for the stream of life, of action, of changing circumstances, and the context in which the artworks originated in his art and literary criticism. In a letter to his friend the composer Zeltner, Goethe puts the point compactly (and in doing so intimates one of the deep parallels between understanding persons and understanding works of art): 'Works of nature and art one does not get to know when they are finished; one must catch them in their genesis to comprehend them to some extent.'[11] Meaning, for Goethe as for Wittgenstein, is inseparable from context, and contexts are not static; any successful analysis, or overview, of a philosophically or conceptually troublesome concept must first acknowledge that the deeds (and of course the words, as a subcategory of those deeds) in the beginning that constitute the self, that speak (as we saw in Chapter 5) for the self (as we saw in Chapter 3), are not performed or said in static, invariant, unchanging contexts. They are done, and said, *in* the stream of life (as we have now seen in our discussions of Augustine, Kierkegaard, Dostoevsky, Cavell, and other autobiographical writers in Chapter 1).[12]

---

   [9] Kaufmann, *Discovering the Mind*, 34.
   [10] I offer a discussion of the aesthetic relevance of Wittgenstein's pointed remark in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), *passim*; see esp. 'Literary Interpretation and Philosophical Investigation', 169–78.   [11] Kaufmann, *Discovering the Mind*, 32.
   [12] In connection with the necessity of understanding both works of art and persons within the expanded context of a developmental history—in the stream of life—consider Goethe's own words on his earlier *Faust* from the perspective of many years later: 'It

There are many further linkages. The Wittgensteinian conception of avowals was developed in order to do battle with the bewitchment of our intelligence by means of language, specifically in that the elucidation of the nature of avowals unearths and removes the picture of objects-and-designations as it applies to the inner life. That assures that we no longer pre-reflectively assume that self-expressive (and, with Goethe's achievement in view we now see more clearly, also self-constitutive) utterances are the verbal results of 'reading off' from inner pre-linguistic facts. Again in personal communication, this time in a conversation with the scientist Jacob Soret (the translator of Goethe's *Metamorphosis of Plants* into French), Goethe said: 'I have never affected anything in my poetry. What I did not live and what did not well up from inside me I did not express in poetry. Love poems I wrote only when I loved.'[13] This may at first glance appear dualistic, Cartesian, and thought and said in accordance with the pre-avowal picture of inner-object-and-designation. But it appears this way only at *first* glance: what Goethe certainly does not say, and implicitly argues against—but by showing in a broadly Wittgensteinian fashion, not deducing—is that the experience that serves as the criterion for the truth of the poetry is both identifiable and experienceable prior to, and ontologically in isolation from, the language that expresses it. And indeed that language, in Goethe's case, is genuine—it is motivated by inward experience; but here, precisely as Wittgenstein suggests is necessary to the intelligibility of the term, we want to put the word 'inward' to *use*, without Cartesian presuppositions (as we saw initially in Chapter 1 and a number of times since). Goethe, like Wittgenstein, is thus no behaviorist in disguise; he is not in the business of *reducing* the inward to the outer. On the contrary, he *shows* us how to rethink the Cartesian–behaviorist categories. Far from separating the self-constituting introspective utterance from its experiential content, Goethe in fact shows the nondualistic avowal as

represents a permanent record of the development of a mind, tormented by everything which tortures all human beings, stirred by the same things that trouble us all, engulfed by what we all abhor, and delighting in the things we all desire. The author has long since put these states of mind behind him . . .'. Goethe said, shortly before that, that in looking at the book, he is 'reminded of that period in my life when the work was conceived' ('Faust (1828)', in Goethe, *The Collected Works*, iii: *Essays on Art and Literature*, ed. John Gearey (Princeton: Princeton University Press, 1994), 185–6). In short, his own perspective on his own greatest work is autobiographical, seen in the contextualized terms of mental development, and is thus anything but static or invariant: he catches even his own work in its genesis to comprehend it.

[13] Kaufmann, *Discovering the Mind*, 32.

it functions in self-revelatory practice. In the lines for the epigraph of his investigation of the hopelessness of an old man's romantic love for a young woman in his 'Marienbad Elegy' (lines taken from his own earlier *Torquato Tasso*), we find the unforgettable words 'A god gave me to utter what I suffer.' These are words that neither describe nor designate an inner emotional object existent prior to and separable from the utterance, nor describe outward behavior in isolation from what we can intelligibly, nondualistically, call the inner life.

There is one point at which Goethe says what seems initially a strong contender for the most un-Wittgensteinian thing sayable. He writes (again in the *Theory of Colors*[14]): 'every attentive look into the world involves theorizing'. But if we look closely at what he says next, it emerges that 'anti-theorizing' might be closer to the mark. For he adds, 'But to do this consciously'—and Wittgenstein was the most concerned author to state the influences of unwittingly held conceptual pictures and strong forces on our thought quietly emanating from down below in the intuitive substrate—'with self-knowledge, freedom, and, to use a daring word, irony'—and Wittgenstein was similarly most concerned to write and think with self-knowledge and freedom from the tyranny of misleading grammatical appearances as well as to maintain, if not exactly an *ironic* stance toward the investigation, at least a constant mindfulness that we may at any turn take conceptual confusion for profundity—'that skill is needed if the abstraction we are afraid of'—and who more than Wittgenstein has made us afraid of abstractions?—'is to be harmless and the experienced result for which we hope is to be vital and useful'—and Wittgenstein, finally, is no stranger to the values of pragmatic vitality and usefulness for our ideas. Goethe is, despite his remark about theorizing (again, the one that subverts itself by gradually morphing itself as the sentence progresses into a remark about anti-theorizing), clearly opposed to the 'false hypothesis' that 'fortifies itself', for when it does so, 'when it is accepted universally and becomes a kind of creed that nobody may doubt, that nobody may investigate, that is the disaster of which centuries suffer'.[15] And he goes on to lament—just as indeed Wittgenstein did throughout his writings—the pernicious power of what he calls an 'ossified doctrine', an unwittingly held doctrine that exerts a powerful shaping influence on our thought and that, owing to the seeming naturalness of the corollary conceptions it engenders, is never unearthed and subjected to investigative scrutiny. (The doctrine

---

[14] Quoted in *Discovering the Mind*, 38.   [15] Ibid. 45–6.

of the Cartesian self and its corollary skepticism concerning our ability to know the contents of another's mind constitutes a perfect example.) Goethe, like Wittgenstein, resists the falsifications of a unifying theory in favor of a many-colored investigation into our practices. Reality does not reduce to a single color. Thus, Faust says: 'In many-hued reflection we have life.'[16]

If we turn to Goethe's critical writings on the arts and literature we find, first, many further strands linking his critical practices to the views of Wittgenstein; second, and far more importantly, we find words and critical deeds that again *show* what Wittgenstein's philosophy says. Under the first heading, i.e. the strand linking the two, we find in Goethe's essay 'On German Architecture'[17] a clearly expressed doubt concerning judgmental principles and generalities. He writes: 'Principles are even more damaging to the genius than examples' (if we could delete the word 'even' and add to 'examples' the phrase 'which are of the greatest value' we would have a full convergence of opinion between the two authors). Of his visit to Strasbourg Cathedral, he writes that when he 'first came to visit the cathedral, [his] head was filled with general notions of good taste',[18] but then goes on to articulate finely the particularities which justify the contextualized critical appreciation and judgment of the building. Revealingly, Goethe also describes the experience of the sublime on encountering the façade the first time, and in doing so makes a familiar appeal to the unsayability of the experience: 'But what unexpected emotions seized me when I finally stood before the edifice! My soul was suffused with a feeling of immense grandeur'—but then he returns to the particularities giving rise to precisely that feeling in precisely that context—'which, because it consisted of thousands of harmonizing details, I was able to savor and enjoy.' And while a variant of the Kantian conception of the sublime is present (Goethe adds, 'but by no means understand and explain', and a bit later, 'It is hard for the mind of man when his brother's work is so sublime that he can only bow his head and worship'), he goes on to *describe* the features of the cathedral, and the organic sense of necessity generated by the interrelations of its

---

[16] Ibid. 54. It is clear that, like Wittgenstein, Goethe *worked* his way, against the pull of a unifying essentialism, to an embracing of particularity and multiplicity. In his 'Postscript' (Goethe, *Collected Works*, iii. 233), the editor writes, 'he was forced in his botanical research in Italy to abandon his hope of finding a single plant from which all others derived, his *Urpflanze*. He similarly stopped short in his aesthetic philosophy in believing in a form of beauty from which all others should or could derive.'

[17] *Collected Works*, iii. 4.   [18] Ibid. 5.

parts, its 'harmonious proportions'.[19] The descriptions of those qualities taken together render the sublime experience intelligible, and indeed, if in a slightly paradoxical way (in that the content of the sublime is said to extend beyond the bounds of the imagination), readily imaginable. In short, that critical detail gives the very explanation that Goethe—in the grip of the philosophical picture of the sublime—explicitly says is impossible, and in doing so brings the architectural experience similarly within the bounds of understanding, which he also explicitly placed the experience of the sublime quite beyond. Like many passages of philosophical writing in which examples and cases are always close to hand (e.g. Plato, Augustine, Hutcheson, Kierkegaard, Nietzsche), Goethe's examples do more compelling philosophical work than do his explicit pronouncements, and in this respect too display a deep affinity with Wittgenstein's later philosophical methods.

In his extremely engaging essay 'On the *Laocoön* Group'[20] we find exactly one such case in which the fully described example performs the valuable philosophical service of showing the nature of person-perception. Here action, thought, reaction, expression, and emotional state are all non-inferentially perceivable not *via*—which is the mistaken behavioral conception of person-perception—but *in* behavior or, if that term now has Cartesian-antithesis connotations clinging to it and thus derives part of its meaning oppositionally and hence shares the conceptual foundations of that which it opposes, the depicted action itself. And, as we shall see, the example also shows a context in which an inferential perception related to observed behavior *does* indeed occur—but in a way wholly different from the philosophically misguided conception of inferential perception that behaviorism would force upon us.

Goethe explains the posture and the perceived movement of the father in the group in terms of reaction: 'his sudden pain from the bite is the primary cause of his movement'.[21] The snake, wounding the father in a particularly sensitive spot above and just behind the hip, produces the immediate reaction emanating from a source of pain that is, as unmediated person-perception, instantly known to the viewer. Goethe describes the behavior: 'the body strains in the opposite direction, the abdomen is drawn in, the shoulder is forced down, the chest thrust out, and the head is inclined toward the afflicted side'. And Goethe here too shows how we see not an isolated, static moment, but movement

---

[19] *Collected Works*, iii. 6.   [20] Ibid. 15.   [21] Ibid. 18.

within the passing time of the context. 'The fettered feet and wrestling arms indicate the situation or action immediately preceding', and it is this continuum of depicted time in the sculpture that makes the rare combination of 'struggle and flight, activity and passivity, resistance and surrender' possible for the sculptor. Indeed Goethe emphasizes that the particularities of context determine (perhaps uniquely) the aesthetic possibility; that combination, not possible in a static moment, would also 'not be possible under any other circumstances'. Goethe implicitly underscores the natural human knowledge of the viewer that obviates the need for inferential perception or hypothetical behavioral deduction of an '*If* the body strains, the abdomen is drawn in, the shoulder is forced down, the chest thrust out, and so forth, *then* the figure is in pain' kind, noting that we cannot imagine the bite being inflicted anywhere else, for then all the gestures would be different or emanate from a different central point of pain. 'We cannot conceive [of Laocoön's gestures] being more appropriate than they are here,' and the complex facial expression is similarly comprehensible precisely within that context. But again, the example does still more work, it shows more, than Goethe perhaps realizes. He laments the poor restoration of the structure, executed in such a way that the restored head of the snake does not correctly locate the actual bite. And here we see the inference that does occur: from the imagined center of the collection of gestures depicted the viewer can, and indeed does, exactly locate the identification of the bite. But if we look to that spot in the poor restoration Goethe saw, we find not the snake's fangs, but rather the protruding remains of both jaws against the father's body at just that point, with the snake's restored head elsewhere. The rather striking contrast thus emerges, i.e. that while we can and do infer the point of the bite, no similar inference—contra the central tenet of behaviorism which thus is here revealed as falsifying the character of the third-person perception of (what we might here call) biographical consciousness—occurs in, or is prerequisite to, the viewer's natural human perception of the pain and the combination of states previously listed. In recognizing the condition of the father, we have, not an analysis of behavioral evidence yielding a hypothesis, but rather precisely what Wittgenstein identified as an irreducible attitude toward a soul. And it is in that very context that Goethe adds the words 'Far be it from me to dispute the unity of mind and body.'[22]

---

[22] Ibid. 20.

Thus, like Wittgenstein, who memorably said that the best picture of the soul is the human body, Goethe sees the mind, the self, not *through* or *behind* the body, but rather *in* the action of the person. He sees, as do we, 'anxiety, fear, terror and paternal love rush through this man's veins, grip his heart and furrow his forehead', and indeed he sees that in this sculptural work 'mental and physical suffering are portrayed to perfection'. But he is using those categories, as Wittgenstein would have us do, intelligibly *in context*. The fangs' wound is painful physically, and the fear and terror for Laocoön's sons is painful emotionally or mentally, but those terms are both *used* in a way decidedly independent of the Cartesianism–behaviorism metaphysical dichotomy; the mental anguish is not 'behind', or hypothesized only indirectly and 'through', the physical behavior. It, like the pain of the bite, is discernible in the action of the character in a way inconsistent with—indeed incompatible with—the dualistic dichotomy. Like the word 'inference' above, we can find a necessary occasion for the uses of 'physical' and 'mental', but, also like the real inference above, the real uses repudiate the metaphysical picture they may seem, at an uncontextualized glance, to support.

2. THE MIND SHOWN: LEONARDO, REMBRANDT, AND MIMETIC ACTORS

The words 'mental' and 'physical' thus have meaning *as they function* in (sometimes separate, sometimes complicatedly interconnecting) language-games. Wittgenstein said, as we have seen just above and, at an earlier stage of the discussion, at the close of the first section of Chapter 3, that there is a 'complicated relationship' between them, and we will invariably 'go wrong' if we try to reduce that relationship to a '*simple* formula'. Still another place we see the irreducible complexity of which Wittgenstein speaks is in Goethe's writings on Leonardo. Wittgenstein's example, discussed in Chapter 3, of the doctor and the nurse who reports of the patient 'He is groaning', suggests that the mental and the physical do not in practice correspond to the categorically neat metaphysics of dualism, i.e. these words of the nurse give a sense of the whole person—the patient—in a way that is not reducible to the neat categories of the mental or the physical. Similarly Goethe describes Leonardo's mind as we see it *in* his work. He writes: 'The many gifts bestowed on him by nature were concentrated mainly in his

visual perception.'[23] A gift for extraordinarily acute visual perception is indeed a fact about a person that we might think of as definitively *mental*; it is about the inner conscious experience of that person. Yet if we look to Goethe's subsequent remarks, it is not only that the ontological divide is crossed repeatedly in both directions, but moreover that—as Wittgenstein suggested—the 'simple formula' of a neat ontological separation is just too simple for the facts. The truth here is irreducibly complex, and to fail to acknowledge that complexity, or indeed to cultivate an ever-more-nuanced awareness of it, is to blind ourselves to what is most original and most radical in Wittgenstein's later work. Goethe continues: 'This is why, despite his many other talents, he was at his greatest as a painter . . . And since the clarity of man's visual perception is part of the domain of the mind, it follows that our artist was endowed with clarity of apprehension and intelligence to the highest degree.' The clarity of Leonardo's visual perception is not an inner and hidden fact knowable only unto himself through Cartesian introspection. It is a *publicly knowable* fact that we see in his work, his painting, in a manner that belies any general pronouncement or ossified doctrine concerning the epistemic limits on our knowledge of another's mind that are enforced by ontology. And Goethe articulates the reasons we see Leonardo's intelligence as well as the corresponding strength of his visual perception in his work: we see no indulgence of impulses, no arbitrary or random brush strokes, but rather only exquisite planning, deep reflection, a profound mastery of proportion, and related aesthetic virtues that, like these, are clearly visually discernible in the work, knowable by minds other than Leonardo's, and yet are indisputably qualities of *mind*.

Goethe similarly praises Leonardo's *Last Supper* for showing 'the whole gamut of emotions, from the most restrained demeanor to the most passionate outbursts'. Goethe's words once again deliver philosophical significance,[24] specifically in his remarking on the very

---

[23] Collected Works, iii. 53; this and the next two quotations.

[24] In his essay 'Ancient versus Modern' (ibid. 90–3), Goethe quotes a young critic, Karl Ernst Schubarth, who writes of his preference for Shakespeare over Goethe on the grounds that Shakespeare's 'unselfconscious . . . intuit[ion]' compares favorably with Goethe's 'arguing, pondering, hairsplitting, analyzing, and overemphasizing'. Schubarth is right about Goethe's argumentative and analytical labors, and I do not mean to suggest that Goethe's *words themselves* do all the (philosophical) work; he is, broadly speaking, a philosophical author at work on the clarification and elucidation of ideas and the arguing of one position over another. It is in this respect of interest to note what Schubarth says further of Goethe's work: 'when I read Goethe . . . from the very beginning I have to

great amount of observation, on the very great deal of visual scrutiny of human expressive action, of human behavior—nondualistically construed—that must have been required during the sixteen years Leonardo worked on this painting. The enormously intelligent depiction of facial and gestural expressivity is taken *from nature*: it is not taken from an inner model,[25] an inner emotional state private to the mind of Leonardo that dictated the external facial and bodily forms that mirror its private inner contours. Similarly, in Goethe's piece (with a title particularly promising for present purposes) 'Rembrandt the Thinker',[26] we find a quotation from Giuseppe Longhi, the engraver and art historian, who writes:

I cannot pass over in silence [indeed here we *can* speak] the etching of the Samaritan [in Rembrandt's *The Good Samaritan*], in which Rembrandt drew the old man in the doorway in the attitude characteristic of people who have a tendency to tremble. And now, as all the man's memories surge to the surface, he actually appears to tremble, an effect no other painter has ever been able to achieve in his art.

Rembrandt's achievement is described perfectly, to which anyone who has seen the work will attest. We see—and only in a narrow sense is this impossible—the tremble, as we see thoughts and memories surge to the surface of the old man's mind. Longhi, as Goethe fathoms in quoting and commenting on Longhi, has captured the aesthetic fact of the case, a fact too complicated for any simple formula concerning the relation between the mental and the physical to accommodate. One would deny the aesthetic fact, i.e. what Longhi, Goethe, and we do in truth see in Rembrandt's etching, only in the grip of a simplifying mental-versus-physical dichotomy, by making a last-ditch appeal to what we *really* see, arguing that we in brute fact only see physical lines on the drawing's surface and we infer the rest.

struggle with opposing views and overcome them and be on guard, lest I take for plain truth what is meant to be rejected as totally erroneous'. The identical words might well describe Wittgenstein's later philosophy, and indeed a number of the misinterpretations in the first years of scholarship on *Philosophical Investigations* were mistaken owing to the conflation of Wittgenstein's interlocutor with Wittgenstein's authentic voice. I offer a much fuller discussion of this latter issue in 'Wittgenstein's Voice: Reading, Self-Understanding, and the Genre of *Philosophical Investigations*', *Poetics Today*, 28/3 (Fall 2007), 499–526.

[25] I present a more detailed examination of this false, picture-driven model of artistic creativity in *Art as Language: Wittgenstein, Meaning and Aesthetic Theory* (Ithaca, NY: Cornell University Press, 1995), ch. 4: 'Artistic Intention and Mental Image', and ch. 5: 'Against Creation as Translation'.   [26] *Collected Works*, iii. 66.

That would constitute the exact aesthetic analogue to the misleading behaviorism we have seen Wittgenstein battle against and undermine, and—as would be evident to anyone looking directly at the drawing—it would radically misconstrue the nature of our perception of the work across, beneath, over, and apart from the strict and simple metaphysical divide between the mental and the physical. Goethe, like Wittgenstein, is interested in what we have seen him call the 'whole mind', the full range of what Wittgenstein described as the extensive and various language-games that give mental predicates and descriptions a home. And like Wittgenstein, Goethe possesses too capacious a mind—as we see it in his writings—and too encompassing a vision, or indeed an overview of this landscape,[27] to find a reductive 'formula' anything but anemic if not bloodless when brought up against real examples.[28]

---

[27] In *Meaning and Interpretation*, ch. 1: 'Language-Games and Artistic Styles', and ch. 2: 'Forms of Life and Artistic Practices', I attempt to provide an overview, along with its aesthetic ramifications, of what Wittgenstein means by two concepts central to his philosophy, 'language-game' and 'form of life'. These concepts, taken together and rightly understood, I believe show why Wittgenstein turned to geographical and cartographical metaphors in describing (indeed) the territory, the coverage, or the ground of philosophical work and the encompassing and perspicuous overview (*Übersicht*) of it; this metaphor stands in striking contrast, in opposition, to that of an incisive, microscopic analysis of otherwise concealed essence. This metaphorical usage obviously reflects the radical change in the one generation of philosophical work from Russell to Wittgenstein.

[28] Such real examples of the genuine and humanly situated perception of personal, characterological, and moral conditions and qualities that defy the mental–physical dichotomy and that are noninferentially perceivable in face and posture are readily and frequently offered by Goethe. See e.g. his closing paragraph of 'Rembrandt the Thinker': 'As for the faces of the six figures, we see nothing at all of the Samaritan's and only very little of the profile of the page holding the horse. The servant, encumbered with his human load, has a resentful and strained expression, and his mouth is closed. The unfortunate injured man registers utter helplessness. The old man's mien, which characterizes him as a decent person, good-natured and trustworthy, stands in sharp contrast to the taciturn and determined face of our robber chief in the corner' (*Collected Works*, iii. 68). In the face of—or indeed in our unmediated human perception of and attitude toward the faces portrayed in—such examples, any attempt to separate ontologically the language-games of the physical and the mental, the physiognomic and the moral, and to stipulate a simple and succinctly expressible relation between them pales into implausibility. In short, Goethe's facial descriptions are real; they are not descriptions of bodily movements upon which we base an inference to inner quality. And at one point Goethe speaks, tellingly in respect of present concerns, of the inscriptions of experience on the face: 'Since the suffering we have endured and the actions we have performed leave an indelible imprint on our face, it is not surprising if every work or achievement which may result from our struggling bears the same imprint. To the attentive observer it reveals a human being . . . ' ('Ancient versus Modern', 91).

There is, it is true, one remarkable passage in Goethe's essay 'On Acting'[29] in which he seems to think and write in a way deeply concordant with, not the complicated practices of the 'whole' mind, but rather the simple Cartesian dichotomy. In stating the practices in which an actor should engage in order best to develop mimetic technique, Goethe instructs as follows: 'stand before a mirror and speak the passage to be declaimed softly, or better just think the words. The advantage of this method is that one is not carried away by declamation, but rather can easily notice any wrong movement which does not reflect what is thought, or softly spoken.' Does this not reaffirm precisely the dualism of inner and outer, of inner cognitive or emotive content and outer expression of that content in speech or gesture? Does it not reinstate the very picture of human selfhood, of autobiographical consciousness, that has been in dispute since the epigraphs of this book? As we shall see, it does not—but it gets worse before it gets better. Goethe continues: 'The actor can then also select attractive and appropriate gestures and lend impact to the entire mimetic action through movement that is artistically analogous to the meaning of the words.' That might well seem substantially more than enough to give a defender of the Wittgensteinian position or collection of therapeutic methods pause. It speaks of a movement, chosen and deliberately performed, that is chosen because it is judged, before its physical realization, to be the correct mimetic depiction, the correct outward manifestation, of the inner thought. It speaks of behavioral movement that stands as the outward analogue to the meaning (and presumably inward meaning at that) of the words.

But, first, this is *acting*, and as such is perhaps the exception that proves the rule. Language itself is not a code; a code, as a deliberate association of meanings with words—where those are categorically distinct entities—is in truth parasitic on language. Similarly, acting is not after all normal or ordinary behavior; it is a deliberate association of thought or meaning with chosen behavior (or at least it may be this in the early stages of rehearsal, before, significantly, the character played by the actor develops). And acting is similarly parasitic on real behavior, real action, thought, and speech.[30]

[29] *Collected Works*, iii. 221–2.
[30] I am very much abbreviating some of the philosophically informative distinctions and relations between pretending and doing. See e.g. J. L. Austin, 'Pretending', in *Philosophical Papers*, ed. J. O. Urmson and G. J. Warnock, 3rd edn (Oxford: Clarendon Press, 1979).

Second, we have not seen here the claim that the cognitive or emotive content, yet to be shown in behavior, is itself of a Cartesian kind, i.e. that it is in the first instance pre-verbal and only contingently articulate. On the contrary, it is the *words* that Goethe says need to find their natural (only seemingly so, since it is acted) behavioral counterpart. And to get the right counterpart, as Goethe indicates, the actor needs to replicate mimetically, not pure inward psychological content, but rather genuine human action itself. Like the painter painting *not* from an inner introspected model, but rather from life, so the actor, in Goethe's account, similarly imitates not pure and metaphysically private thought, but rather contextualized, intelligible, and noninferentially grasped meaningful action (or again, human behavior in the nonpernicious sense of the term). That irreducible truth, again showing that the words of Goethe are only seemingly or superficially supportive of the dualistic picture, is found in his next words. And on closer investigation, as before, these words do more philosophical work than at first appears. He continues: 'But it must be presumed that the actor has first carefully studied the character and the situation of the person to be represented, and that he has gone over the material thoroughly in his mind. For without this preparation he will not be capable of either declaiming or moving correctly.' If the actor has *carefully studied the character and the situation of the person to be represented*, then he has imagined, to put it generally, the form of life[31] in which the character meaningfully behaves (in the largest and most encompassing sense of that term). The actor will imagine, not the metaphysically solitary thoughts of another mind that will, contingently, produce introspective utterances emanating from a pre-verbal dark interior, but rather the language-games into which the character enters and within which he, in a natural language, expresses himself and reveals his mind.

Finally, when the actor goes over the material thoroughly *in his mind*, what he is going over is itself, again, the public, knowable, readily discernible meaningful *action* on the part of the character—action that will correspond no better to simple, formulaic dualistic categories than did the earlier examples. Without such preparation the actor will not be able to behave as a whole human being. And without this capacity, the actor will not be able to be perceived in the most distinctive manner—indeed in precisely the nonreductive manner of

---

[31] For an investigation into what Wittgenstein meant (by looking closely into his own uses of the term) by this term of art, see the chapters referred to in n. 27 above.

person-perception that Wittgenstein has articulated[32] and that is unique among all human perceptual experience—that is, to be perceived with an *Einstellung zur Seele*,[33] an attitude toward a soul.

Goethe, like Wittgenstein, is a profoundly serious thinker about the mind, about the nature of the self, and about the distinctive nature of our perception of each other. Like Wittgenstein, Goethe's writings do not settle into traditional dualistic categories: he too is neither Cartesian nor behaviorist, nor does his position take its form as an antithesis to those twin options. Both Goethe and Wittgenstein, however dissimilar they may prove to be in other respects, deeply comprehend the irreplaceable value of particularity for understanding and of context for intelligibility. Wittgenstein's extensive philosophical investigations heighten our sensitivity to intricacy and nuance in a distinctive way that continually reorients our way of looking at conceptual difficulties and that continually frees us of the theoretical pictures and conceptual models that exert undue—and prismatically distorting—influence on our thought. Goethe's extensive aesthetic investigations—if we know where and how to look—show us, throughout numerous genres, a strikingly similar achievement. But to describe more fully and accurately the distinctive nature of person-perception as it operates in seeing both others and ourselves anew, we need to turn to Wittgenstein's writings on aspect-perception, or 'seeing-as'.

3. IRIS MURDOCH, THE 'UNFROZEN PAST', AND SEEING IN A NEW LIGHT

In 1957 Iris Murdoch wrote, with perhaps too much concision, 'Man is a creature who makes pictures of himself, and then comes to resemble the picture.'[34] And in an earlier diary entry of 14 June 1952 she had written: 'There is a lot which I don't put into the diary, because it would be too discreditable—and maybe even more painful.' Characteristically,

---

[32] I offer discussions of precisely the nonreductive nature of person-perception in *Meaning and Interpretation*, ch. 4, sect. ii: 'Against Reductionism', and in 'Leporello's Question: *Don Giovanni* as a Tragedy of the Unexamined Life', *Philosophy and Literature*, 29/1 (Apr. 2005), 180–99.

[33] See Peter Winch, 'Eine Einstellung zur Seele', in *Trying to Make Sense*.

[34] Murdoch's diary entries are quoted, contextualized, and insightfully discussed in Peter J. Conradi, *Iris Murdoch: A Life* (New York: W. W. Norton, 2001); this quotation p. 272.

she quickly turns to reflect upon, to refine, and to qualify what she has just written, adding parenthetically: 'At least—no major item omitted but certain angles altered—and painful incidents omitted.'[35] Still before that, shortly after having listened to a lecture entitled 'The Past' by Elizabeth Anscombe in October 1947, Murdoch, as her biographer Peter Conradi reports, was reflecting on 'what she might feel if presented with documentary evidence—for example, journals—about her forgotten past'. She then writes a passage that both gives voice to a kind of guarded skepticism concerning self-knowledge and makes a strong claim concerning the *active* nature of our involvement with our past. She writes: 'Suppose I were given evidence about what I thought at the time. My diaries, etc. I think I would not accept that evidence. I'd still feel I didn't know what my past really was.' Describing these diary pages—pages within a diary attempting to advance our understanding of the degree to which a diary is credible, authoritative, and revelatory—Conradi says: 'Over many pages of reflection, she reaches towards a distinction between a "frozen" and an "unfrozen" past. So long as one lives, one's relationship with the past *should* keep shifting.'[36] Indeed, in those pages Murdoch also put this point compactly and with emphasis: '*re-thinking one's past is a constant responsibility*'.

With these lines Murdoch has brought a number of interlocking themes into play, and each of these themes articulates a fundamental problem of autobiographical consciousness. How do we go about the life-defining project of making a 'picture' of ourselves which we then come to resemble? How does the selectivity displayed in choosing what to include and what to leave out—or semi-wittingly avoid—in diary-writing determine both the content and the outlines of that picture? How might we characterize, and give more exacting expression to, the

[35] Ibid. 274.
[36] Ibid. 275. A reading of her diary, as we have it in Conradi's biography, shows that this was for Murdoch anything but a matter for abstract speculation, detached from life. To take only one example, upon learning of the wholly unexpected separation of two of her married friends, she writes: 'I had thought of them as so indissolubly connected & somehow of that part of my history concerning them as so completely ended. . . . An extraordinary sense of time rolling backward. . . . In some way, the parting of those two *reopens my own past.* It is as if they, together closed a door for me, ended a certain piece of my history, & closed the book. Now that they are parting that force is no longer exerted.' Later she wrote to one of the separated parties, 'I'm very glad the future contains you,' and it is clear that for Murdoch that particular future will be heavily inflected by a retrospectively revised past. Ibid. 430–1.

process—a process that is clearly centrally significant for any project of self-investigation—of altering the 'angle' of a life's 'major items'? Skepticism is never just whimsically adopted (not, that is, with human seriousness anyway), it is *motivated*. So precisely what, we want to ask, motivates the skeptical stance toward the evidence of which she writes such that, even with the evidence plainly before us, we find it plausible to claim that we still would not know 'what our past really was'? And perhaps most importantly, what does it mean to relate to one's own past as 'unfrozen', to have one's relationship with one's own past keep shifting? How can we more exactly articulate what it means, in Murdoch's words, to 're-think' the past?

In facing such questions we should here again realize that we do not usually face them alone, but rather with a host of presuppositions. To recapitulate briefly, we have seen these to include: (1) philosophical conceptions or, in Wittgenstein's special sense, pictures, of selfhood; (2) analogies that both lead and mislead; (3) grammatical similarities between cases of different kinds that lead us to take them as more similar than they are; and (4) epistemological intuitions that import into the proceedings from the outset expectations about precisely which conditions must be satisfied in order to arrive at a true proposition with self-revelatory content. We have seen all of these in action as we worked through the preceding chapters, but in this section I will focus primarily upon the last: it is all too easy, when reflecting in a preliminary way about the kind of autobiographical understanding that comes from a grasp of one's past, to conceive of the problem as a polarized epistemological dichotomy. On the first pole (one that has, incidentally, become extremely fashionable throughout the humanities of late), we picture ourselves projecting onto the past the content that we claim to perceive in it, and thus succumb (perhaps unwittingly) to a not unfamiliar variety of self-deception. Here, the narrative self is indeed a narrative construction, and, rather like Hume on causation, what we get ourselves to believe we perceive in our past is of our present retrospective making. Alternatively, on the second pole, we picture ourselves as accurately and nonprismatically perceiving what is in the past in and of itself, where the true self-revelatory proposition is one that is verified through correspondence between present utterance and past fact. Here, the narrative self is one that is constituted not by present active retrospection but rather by the passive, factually constrained accurate memory of those past episodes of one's life. The very idea of 'getting it right', the content of autobiographical verisimilitude, would

seem, at first glance, to reduce to precisely this model of objective reportage. Yet Murdoch emphatically asserts that the past, properly understood, should be 'unfrozen', and that one has no less than a moral obligation to 're-think'.[37] Moreover, she says this without ever so much as giving a hint that one should thus embrace any variety of subjectivist constructionism, that the self-concept that results from an active engagement with one's past is entirely fluid, indeterminate prior to any particular narrative construction, and contingent upon that construction for its created and not discovered sense of stability over time. Perhaps, reminiscent of the polarized dichotomy we considered in Chapter 5, it is the large-scale dichotomy between perception on the one pole and projectivism on the other itself that is false, perhaps the truth of the matter is far too intricate for this intuition-supported dichotomy to begin to accommodate.

We speak, of course, of our understanding of our past, and of the pictures of selfhood drawn and supported by that understanding, in ocular terms. We speak of how we see a situation, of seeing it differently, of tenaciously or intolerantly seeing a circumstance in only one way and of one's being unable to see it in any other way, and so forth. We similarly have developed a subtle vocabulary concerning how we see ourselves. Although we have been warned of the dangers of unexamined ocular metaphors in epistemology,[38] these linguistic practices are enough to suggest that a close scrutiny of the subtleties of visual perception may prove helpful to the questions articulated by Murdoch. Indeed, such scrutiny may show precisely how we can sustain an active or unfrozen engagement with our past and how we see it, while still not forfeiting all hope of satisfying our quite fundamental human

---

[37] It is clear that she feels this as a moral imperative throughout her life. At a time of great happiness shortly after her marriage, we find lines stating that life now has 'such a quality of simplicity, warmth, and joy', and that it is now strange to read her earlier much more difficult, and occasionally troubled, diary entries. But even here she gently chides herself for not sufficiently now engaging in 'deep consideration of the consequences of my past actions'; this is evidently a long-ingrained commitment to an active engagement with the past that she clearly feels to be indispensable to a life worth living. And in reference to those earlier difficult entries, she further reinforces this *active* autobiographical sense in quoting Virgil, to the effect that 'the day may dawn when this plight will be sweet to remember'. But Conradi shows that the entry is more complex still (and still more illustrative of what she means by an active rethinking of the past): the lines concerning 'simplicity, warmth, and joy' were written 'in a later hand and ink' and, as Conradi goes on to say, were thus 'a truth grasped retrospectively' (pp. 400–1).

[38] See Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton: Princeton University Press, 1979).

desire for self-descriptive accuracy and autobiographical rightness of a kind that avoids an epistemic descent into an 'anything goes' narrative constructionism. And it is in §xi of part II of Wittgenstein's *Philosophical Investigations* (pp. 193–229) that we find a most sustained inquiry into visual experience that, as we shall see, proves deeply relevant to these issues.

Approximately one-third of the way through that section, Wittgenstein arrives at the question, which he puts in the voice of his picture-driven interlocutor, 'Is it a *genuine* visual experience?' (p. 204). This follows the discussion of an aspectival shift in seeing that a diagram can be seen as two hexagons, where earlier it was seen as a different pattern. The philosophically recidivistic interlocutor—giving expression to an impulse we can all quite naturally feel at a juncture like this where an aspect has just shifted—emphasizes 'genuine' because he construes the real visual experience as one side of a philosophical dichotomy where accurate perception precludes subjectivized projection. If genuine, so one might think here, the visual experience will be authenticated by what the object seen possesses in and of itself: the criterion for the genuine will here be construed as the replication on an ocular level of the intrinsic properties of the object seen. (The direct analogy to the autobiographical case—to the 'second pole' above—is clear enough.) But Wittgenstein does not accept this simple dichotomized picture, of course: he replies: 'The question is: in what *sense* is it one [a genuine visual experience]?' (p. 204, my italics). He has at this point already discussed the general contrast between two kinds of case: (1) seeing that an object is an *x*, and (2) seeing that object as an *x* not because of what we see but because of what we know. Here, for the interlocutor, the former would constitute a case of genuine seeing, the latter inauthentic because too mind-dependent. Wittgenstein's reply, in the form of the question 'in what *sense* is it one?', invites the interlocutor to expand significantly the frame of contemplation, asking him to consider cases in which we would have occasion to mark a contrast between the genuine and its opposite. (Incidentally, only to hint at the circumstantial complexity that this dichotomy fails to acknowledge, its opposite might be called 'nongenuine', 'inauthentic', 'false', 'untrue', 'inaccurate', 'prismatic', 'distorted', and many other things—no two of which, as any reasonably trained Austinian ear would quickly detect, come to the same thing in context.)

Wittgenstein turns shortly to one case in which we might describe the act of seeing in question more as one of knowing than of seeing,

where 'someone treats the picture as a working drawing, *reads* it like a blueprint' (p. 204). Here it makes sense to speak of an architect or engineer seeing much more in a blueprint than one not similarly trained would see. If one were to insist that the trained eye here sees much more than is *really* in the line drawing and thus that they 'see in' (as a variant of 'seeing-as') more than is there, we could quite readily understand the contrast in play (or, if not the *contrast* between seeing and thinking, then certainly the emphasis more on the one than on the other). But this would not constitute, precisely, a contrast (or shift of emphasis) between *genuine* seeing and one of its opposites: the architect and engineer are able to see what the drawing signifies, what it, in a manner of speaking, *implies*, without transgressing the bounds of the genuine. (Indeed, no question of the *genuine*, precisely speaking, has discernibly arisen.) The dichotomy between perception and projection is, in short, already destabilized: neither intrinsic-property-reflecting brute perception nor freely imaginative projection seem to capture what is significant about this blueprint case. And does the self-interpreting autobiographer, looking back at the broad outlines—in some ways the 'blueprint'—of his or her life and coming to appreciate the interrelatedness of the important events in life (of the kind of which Murdoch wrote) by connecting those 'dots' with a narrative thread,[39] clearly engage in one or the other? As Wittgenstein's discussion, even at this early stage, shows, these polarized categories are far too crude: the facts of the case are considerably more intricate. Wittgenstein, a bit earlier in this section, wrote: 'The concept of "seeing" makes a tangled impression.' And, resisting the impulse to falsify by over-straightening (and in a manner reminiscent of Aristotle in suggesting that we only look for a degree of categorical neatness consistent with the nature of the field being investigated), Wittgenstein adds: 'Well, it is tangled' (p. 200).

So can the criteria for an act of genuine seeing be stated in general terms, and would not any such criterion for genuine seeing generate, *mutatis mutandis*, a criterion for genuine (or true) self-'seeing', i.e.

---

[39] This metaphor has been developed at length in relation to the problem of explaining how the contours and trajectory of a whole life can make what we can sensibly refer to as *sense*; for two particularly insightful studies, see Richard Wollheim, *The Thread of Life* (Cambridge, Mass.: Harvard University Press, 1984), and Richard Freadman, *Threads of Life* (Chicago: University of Chicago Press, 2001). I offer a discussion of the nature and value of such narrative threads in 'Narrative Catharsis', in John Gibson, Wolfgang Huemer, and Luca Pocci (eds), *A Sense of the World: Essays on Fiction, Narrative, and Knowledge* (London: Routledge, 2007).

autobiographical self-understanding? In seeing 'that an animal in a picture is transfixed by an arrow' (p. 203), Wittgenstein asks, if the picture is a silhouette, do we *see* the arrow or do we 'merely *know* that these two bits are supposed to represent part of an arrow?' Importantly for the question concerning the possibility of stating general criteria derived from a general model of seeing, Wittgenstein considers both of the following emphatic responses to the silhouette case: ' "But this isn't seeing!"—"But this is seeing!" ' These are both, we can see at a glance, rational and defensible responses that we imagine arising in different contexts, i.e. where the particular point, or conversational goal, of the response fits into a larger pattern of locutionary interaction—a language-game.[40] And it is within such circumscribed language-games that the particularized and context-sensitive criteria for seeing emerge as salient. 'It must', Wittgenstein writes, 'be possible to give both remarks a conceptual justification' (p. 203); and so it is. And he asks again, here responding in turn to the second, positive response to the silhouette question above, '*In what sense* is it seeing?' (p. 203), the point being that the determinate sense will be given by—and only by—the context of the question and its response. If we can grasp, on the level of detail in both the blueprint and the silhouette cases, that the responses might intelligibly, rationally, and defensibly go either way, the prospects for a *general* criterion for genuine seeing are diminishing rapidly. What, at this point, shall we say of the analogy to the self-interpretive situation?

Seeing the intrinsic properties or features of an object, and then giving those properties their bluntly factual corresponding descriptions, constitutes the 'second-pole' model that carries its own (again, too crude) way of construing autobiographical truth: we look back at the past experience and give it its bluntly factual corresponding description. But if, in both the blueprint and the silhouette cases, we can fully comprehend *both* 'But this isn't seeing!' and 'But this is seeing!', then it is clear that a general, overarching criterion for genuine seeing is perhaps something we could *stipulate*, but not something we could *discover* within the fabric of our experience. Consider parallel cases, for example Nabokov[41] as autobiographer in place of the architect or

---

[40] Here again I rely on my much fuller account of this fundamental Wittgensteinian concept and its significance for aesthetic contexts in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), 9–44.

[41] Consider in this light Nabokov's observation in his autobiographical *Speak, Memory* (New York: Random House, 1989), ch. 8: 'the supreme achievement of memory . . . is the

engineer, reflecting on the resonances of his formative early Russian experience as they are sounded throughout his adult life in the United States, or a person engaged in self-investigation who *begins* to see (we might metaphorically say that what he is coming to see is seen at this stage only in silhouette) that an early separation trauma significantly fueled his later unacknowledged desire to recapitulate repeatedly that separation experience but to do so volitionally, i.e. to take the role of agent, rather than victim, of separation. These would be cases in which the defensibility of the self-interpretation would not derive from a generalized, overarching criterion for the true autobiographical proposition. If a confidante of the person in this last example says (critically) 'But *that* isn't seeing!' or (in congratulation for hard-won and successful self-investigation) 'Now *that* is seeing!', neither reply is given because the intrinsic properties of an isolated life-event are given accurate description or not. But it is at this precise juncture vitally important to see that this lack of a generalized, case-transcending criterion does not drive us to the skeptical extremes of post-modern narrative constructionism,[42] i.e. back to the 'first-pole' picture. Neither of the two replies is given either (1) because of an application to the present case of a generic case-transcending criterion, or (2) because of a contingent espousing of an arbitrary life-construction. A reply, rather, is given because of a capacious grasp of the life of which the event in question is one significant part and where the rest of that life is known in sufficient detail to see linkages—linkages reported *within* that life's narrative (as we saw in Chapter 5 above)—that give that life its teleology, its sense. Such a life would not, indeed could not, be understood in Murdoch's sense as a sequence of 'frozen' episodes hermetically sealed unto themselves. And this is true precisely in the way that acts of aspect-perception, and their rightness or wrongness, cannot be described in a brute manner hermetically sealed from the contexts within which they occur.[43] That is a false model of objectivity.

---

masterly use it makes of innate harmonies when gathering to its fold the suspended and wandering tonalities of the past'. I discuss this passage and its philosophical significance in 'Davidson, Self-Knowledge, and Autobiographical Writing', *Philosophy and Literature*, 26/2 (2002), 354–68.

[42] See e.g. Richard Rorty's chapter 'The Contingency of Selfhood', in *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989).

[43] For example, in an entry from 1954 Murdoch writes: 'the "who am I to be jealous?" aspect doesn't stop me being in great pain' (Conradi, *Iris Murdoch*, 379). Whether she is right or wrong, or in different ways both right and wrong, to doubt her right to feel

David Pears offers a helpful discussion of Wittgenstein on the self in both his early and later philosophy that holds considerable significance (beyond what Pears expressly covers) for the present issue of aspect-perception in its connection to self-description.[44] Pears observes that, in Wittgenstein's early thought, the limits of language do not demarcate the realm of the sayable from the realm of the possible-but-not-yet-said, but rather the limit beyond which there is 'only a void or vacuum in which logic cannot breathe'. Pears then, encapsulating the analogy from the linguistic to the visual that Wittgenstein had articulated in the *Tractatus*, adds: 'just as there is nothing like the contents of the visual field beyond its limits' (p. 105). Referring to the element of truth that Wittgenstein at that early point found in solipsism, Pears writes:

He commends it [the linguistic solipsist's claim] as a dramatization of a profound truth about the limit of language. The truth is that beyond its limit there is not a whole range of candidates for the status of real possibilities waiting to be adopted and included in the list of actual senses of sentences in our language. Beyond the limit there is only a void into which we may extend our language by constructing new patterns of speech. We do this by adopting new rules of inference after they have been proved to follow from existing rules, or simply by adopting new definitions of old words. We do not extend language to fit what we discover beyond its limit, because there is nothing to be discovered in a void. We can shift the limit further out into the void only by constructing extensions to existing language, and never by producing innovations to fit what we discover. There are no discoveries to be made in a void. (p. 106)

The parallel of this linguistic point for the present discussion is telling: we might, as have seen, picture new aspect-perception as antecedent to new self-description in such a way that the new aspect is believed to capture facts of the self that are in a sense waiting, just beyond the reach or boundary of our present self-description, for articulation. There, again, the criterion for the truth of the newly dawned aspect just would be that pre-perceived fact waiting in the 'void'. The picture Pears has articulated

---

jealousy as an aspect of the experience of emotional pain is a question that could never be answered in a 'hermetic' way, i.e. in any way but a full knowledge of the extended and indeterminately bounded context in which the jealousy arises. This also shows, incidentally, that just as in visual experience (where not all perception can intelligibly be described as aspect-perception—we do not see a fork as a fork), not everything in 'self-seeing' can be characterized as an aspect: the 'who am I to be jealous?' element is (as she indeed says) an aspect; the pain, by contrast, is not.

[44] David Pears, *Paradox and Platitude in Wittgenstein's Philosophy* (Oxford: Clarendon Press, 2006), esp. ch. 5: 'Ego'.

from Wittgenstein's philosophy can, as a new picture to supplant the old, therapeutically shake the hold, or loosen the grip, of that old and entrenched correspondence-criterion picture. Indeed, it may be that, very much like this alternative picture of the growth of language as its boundaries are extended into what was once not 'a whole range of candidates for the status of real possibilities waiting to be adopted', but rather the vacuum in which logic (and, for us, self-descriptive language) formerly could not 'breathe', language, on this picture, extends itself in the bootstrapping fashion of (as Pears puts it) 'adopting new rules of inference'. So there *will* be constraints on that growth, on extensions of our language, on that new picture, and those extensions will be rational, justifiable, and indeed sense-making, but *not* (as would be consistent with the old picture) exclusively by linguistically 'mapping', or giving voice to, the undiscovered territory lying beyond the earlier limit, by giving voice to the pre-existing unnamed 'possibilities'. The growth of language is, on this new picture (I should say that Pears himself, working another vein,[45] does not discuss this in terms of the therapeutic value of picture-supplanting, or the ushering in of conceptual freedom from, an earlier embedded one), a more creative matter—but it is not for that creativity unconstrained, capricious, or wholly a matter of linguistic constructionism. The new language extends, but it does not invariably *extend to fit*, as Pears puts it. Pears thus says: 'Inventions are needed if we are going to succeed in extending language into this void, and

---

[45] Pears also provides a strikingly powerful reminder of a simple fact of our natural history that itself makes a contribution to loosening the grip of the Cartesian picture of the human being. He writes: 'Cartesian systematic doubt seems at first to leave us with a non-problematical "man within", a pure ego that synthesizes sensory input, and we do not worry too much about practical output. It is, of course, a mistake to concentrate on perception to the exclusion of action, and it encourages the idea that a person's mental life as a detached observer is self-sufficient. But, even if we take no account of major action, it remains true and important that an observer still has to move around in physical space in order to acquire and synthesize the different views that he needs to get of the same object' (ibid. 105). As we are seeing throughout the present volume, a person's mental life with regard to self-description and self-understanding is not that of a detached observer; indeed, we need to, as it were, 'move around' (another way of putting Murdoch's point) our pasts in order to acquire and synthesize the different views that we need to get of it. The similes of self-description that arise in recursive aspect-perception are, as self-directed views we acquire, perhaps more on the creative side; the synthesis of those that we then assemble over time are by contrast perhaps more constrained (as, in Wittgenstein's phrase, 'aspects that go together', as we shall see below). Incidentally, it is precisely this striking fact of our natural histories, i.e. that we need to occupy multiple points of view in physical space in order to acquire and synthesize the different views we take of an object, that is recorded in the early analytical cubism of Picasso.

though there will be constraints on the inventions, they will not be the simple constraints of applying old words in the old ways to new material acting as a cue' (p. 106). (We saw this point from another point of view in connection with Moran's discussion of the active power of self-description in the previous chapter.) The distinctive varieties of aspect-perception we are considering, i.e. recursive ones, can (as we can see more clearly having brought in a new picture) be creative in precisely this sense, i.e. they need not *follow* (or track, or map, or give voice to) the facts that allegedly pre-date them and that serve as their criteria for factuality. These newly perceived recursive aspects, these new similes of self-description, follow from, rationally extend from, the things we have seen and said before as they push into the void (which for us, again, is the formerly un-self-described). And so they are not merely exercises in a constructivist or anti-realist free-for-all. But nor need they be realist *in the sense of correspondence to prior facts*. (This connects directly to the discussion of the determination of retrospective meaning in the previous chapter.) We might opt for one familiar way of speaking about it and say that such newly articulated aspects create the reality they describe—but this places too great an emphasis on the creative dimension of aspect-dawning self-description to the exclusion of the constraints within which it moves. Or we might opt for another familiar way of speaking here and say that the creativity lies only in the *language*, but that the reality that language describes is necessarily prior to, and independent from, such post-dictive operations—but that (in addition to reinstating the old picture) places too great an emphasis on, and mischaracterizes, the constraints to the exclusion of the creativity. Indeed the entrenched realist–anti-realist polemic as applied to these phenomena is too blunt to capture the relevant nuances;[46] what a close study of Wittgenstein's remarks on aspect-perception gives us is itself a way of extending our existing patterns of speaking about such matters, ways that are to varying and case-specific degrees admixtures of the constrained and the creative.

---

[46] There is a direct parallel here, also an instructive one I believe, to the way we are inclined (given a host of preconceptions about and pictures of the Cartesian mind) to speak of the distinction between conscious and unconscious mental content; the truths of particular cases are far more nuanced than the categorically neat entrenched dichotomy could capture. For an astute discussion very much alive to the philosophical significance of this degree of particularity, see David H. Finkelstein, *Expression and the Inner* (Cambridge, Mass.: Harvard University Press, 2003), sects 5.4 and 5.5: 'Conscious or Unconscious', and 'Between Conscious and Unconscious'.

In examining the case of a triangle that can be seen variously (to which we might add: as the Great Pyramid, as the Pope's headgear, as having fallen over on its side, etc.), Wittgenstein writes: 'The aspects of the triangle: it is as if an *image* came into contact, and for a time remained in contact, with the visual impression' (p. 207). This reinstates—and reinvigorates, in that it gives voice to the impulse to fall back into the long-established grooves of thought—the bifurcation between perception and projection. From within these grooves it naturally seems impossible to get beyond the quandary over how we could conceivably adopt an 'unfrozen' view of the past with which we actively engage while maintaining any sense of self-descriptive verisimilitude. Wittgenstein's work in this section is, as we might now well expect, decidedly not to develop a theory from *within* these grooves, but rather to become enabled, through detailed examples, to think our way out of them. He writes in a parenthetical remark: 'In giving all these examples I am not aiming at some kind of completeness, some classification of psychological concepts. They are only meant to enable the reader to shift for himself when he encounters conceptual difficulties' (p. 206). And so he asks, 'Is being struck looking plus thinking?', and he answers, 'No. Many of our concepts *cross* here' (p. 211). He gives this stern answer because he has just traversed the following landscape: discussing a person who recounts having looked at a flower without being conscious of its color, and who then says, 'Then I suddenly *saw* it, and realized it was the one which . . .' (p. 211), Wittgenstein considers the response 'He looked at it without seeing it'. And he adds: 'There is such a thing. But what is the criterion for it?' Resisting the temptation to posit a single case-transcending answer, he adds: '—Well, there is a variety of cases here' (p. 211). So just as we can speak of looking at the flower without seeing it, we can speak of looking at an episode or event in one's life without at first, or for a long time, seeing it, and then at some point coming to see it. This is precisely analogous, in ocular terms, to what Murdoch meant in autobiographical terms, by 're-thinking'. The past is not changed and yet, only seemingly paradoxically, it is. Is being struck by the newly appreciated significance of an event in our past, like suddenly seeing the color of the flower and recognizing that it is the special one that . . ., a phenomenon we can helpfully describe in accordance with a simple additive template as looking plus thinking? No; many of our concepts cross here. But how *do* we then describe it?

The past is not changed by the active process given a name by Murdoch and given content, by analogy to sight, by Wittgenstein—and

yet, in a sense, it is.[47] Wittgenstein writes: 'The colour of the visual impression corresponds to the colour of the object (this blotting paper looks pink to me, and is pink)—the shape of the visual impression to the shape of the object (it looks rectangular to me, and is rectangular).' So far, this in and of itself might lead us to think that the simple perceptual side of the perception–projection dichotomy might suffice after all. But it is what Wittgenstein adds to this remark that complicates the picture, and that gives us an answer to the question asking in what precise terms we should describe autobiographical re-thinking. He writes next: '—but what I perceive in the dawning of an aspect is not a property of the object, but an internal relation between it and other objects' (p. 212). Coming to see an object as an *x* or a *y* is not reducible to, and not explicable in terms of, the perception of a property intrinsic to that hermetically sealed (i.e. non-relationally-embedded) object. But neither is it reducible to, or explicable in terms of, the projection of a (mind-dependent) property onto the object. And similarly, I would suggest that coming to see, first, that a given past action was an unidentified prototype for subsequent different—yet still in a heretofore undisclosed sense similar—actions, and, second, that that action was self-interestedness masquerading as altruism, is neither just perception nor projection. It is a dawning of an aspect, the seeing of an internal relation, first, between the prototype and its successors, and, second, between the series of actions and a range of examples involving more self-interest than altruism.

It is this process, I believe, that Murdoch had in mind as rethinking our way into an unfrozen past, in which the mind's role *is* active, but active in the way one hears that a musical passage is a variation of an earlier theme, not as we, with Leonardo, see landscapes in the myriad cracks of the plaster wall. Imagination is required, but with a distinctive kind of interpretive discipline: 'Doesn't it take imagination',

---

[47] This way of putting it has been helpfully described as the 'paradox' of aspect-perception that it was Wittgenstein's project in §xi to solve (or rather dissolve); see Stephen Mulhall, *On Being in the World: Wittgenstein and Heidegger on Seeing Aspects* (London: Routledge, 1990), 6–34, and Hans-Johann Glock, *A Wittgenstein Dictionary* (Oxford: Blackwell, 1996), 36–40. Glock succinctly encapsulates the dualism implicit in the Gestalt psychology to which Wittgenstein was in part here responding, a theory that would introduce two 'visual realities', one outward, and one inward (where the aspect is reified as a private mental entity). Much of the work undertaken in §xi is designed to show that no such private object could serve as the content of an aspect nor as a criterion for the correctness of the description of any visual experience. It would take a separate chapter to show this in detail, but I want to suggest here that both the content and the correctness of any autobiographical endeavor will be 'public' in the same way that the visual turns out to be in §xi.

Wittgenstein writes, 'to hear something as a variation on a particular theme? And yet one is perceiving something in so hearing it' (p. 213). The 'something' is not an intrinsic property the description of which is verified by brute correspondence. It is an internal relation that provides the expanded context within which the particularized criteria for rightness or wrongness (and the long continuum of possibilities between) emerge. If we then ask, 'But is it then at bottom really a case of projection?' or if we ask, 'Is it not then at bottom still just perception after all?,' this will be only a late-stage manifestation of the polarized dichotomy brought in with our pre-reflective intuitions about the matter. And as to one thing or another being 'at bottom', we should bear in mind that there is no reason to assume that one language-game should be prior to, or more fundamental than, another here.

It is here that we encounter another, related, epistemological intuition that comes into play and that can mislead in its own way. We may feel an inclination to insist that one language-game *is* prior to, more fundamental than, another in these cases, and insist in turn that this priority is of the first importance in distinguishing the real facts of the case from the less real interpretation of it. This insistence may concern either the objects of vision or, here again in direct parallel, the objects of autobiographical interpretation. Beginning this line of thought by noting that we do, after all, have to hear the melody before we recognize it as a variation, and that we have to remember the large life-incident of Murdoch's kind before we place it into a pattern of action, thought, and emotion that only emerges on reflection, we may then quickly—too quickly—say that the object of sight in and of itself, the melody unto itself, and the isolated life-event are fundamental, in a sense more real, and thus objective, whereas the aspect seen in the object—be it the status as variation in the melody or the prototypical power of the life-event—is not real in the same way and thus is subjective. And that generic distinction, at precisely this juncture, opens the way for skepticism concerning whether things of the latter kind are knowable or not, leading to a general pronouncement against the epistemological legitimacy of seen or heard aspects or self-interpreted aspects or internal relations. This line of thinking is mistaken, here again in being far too simple a picture to accommodate the 'tangle' here, far too crude to capture the intricacy.

This misguided line of thinking is, I think, not far from Wittgenstein's thoughts even at the very beginning of §xi. His first words in that section are: 'Two uses of the word "see"' (p. 193; and following quotations).

He then quickly contrasts (1) the case of responding to the question 'What do you see there?' (where there are drawings of two faces) with the words 'I see this', followed by a description, a drawing, or a copy) with (2) the case of responding with the answer 'I see a likeness between these two faces'. As Wittgenstein goes on to discuss the difference of category between the two cases, he is careful never to place the one as primary, as real, and the other as secondary, as merely subjective. Of the second category, he writes: 'I contemplate a face, and then suddenly notice its likeness to another.' And then, like the case we considered above when the seeming paradox emerges where the object seen both does and does not change, he adds: 'I *see* that it has not changed; and yet I see it differently. I call this experience "noticing an aspect".' Here, it is true, the categories 'seeing *this*' and 'noticing an aspect' are made distinct, but there is, throughout his analysis, never a point where they are described as ordered in a hierarchy of real 'knowability'.

This, as we can now see, is directly linked to his question concerning the 'genuine' visual experience above. His response to that interlocutor's query (voiced in perfect correspondence to the line of thought under consideration here) was not to accept the presumption concerning the fundamental or the objective (or the genuine) versus the secondary (or the false, the uncertain, etc.—the genuine's opposite), but rather to ask: In what *sense* is it a genuine visual experience? This reshifted our focus back to the relevant particularities of contexts within which the criteria for the genuine and its various contextual opposites emerge *in situ*. And so here, the safeguard against falling back into the grooves will be similar particularities: there may be a difference we want to mark between really having seen an object and not, or really having heard the variation-status of the melody or not, or really having seen that a given life-event laid down a template for a repetition-compulsion or not. But the essential point is that none of these more particularized cases correspond to the generic objective–subjective distinction, nor do they take their expected places on a hierarchy reaching from the unknowable to the knowable. It may emerge in a court of law that I did not directly see the bag of money in the hands of the accused but just saw a fleeting shadow on the subway wall that looked like the accused with a bag. It may emerge in a musical examination that, while I can recognize on the score that the melody is a variation on another, I cannot really hear it as such. And it may emerge that I am only pretending to accept my confidante's suggestion about the residual power of a given life-experience but that I do not really believe it. All

of these cases, again, mark contrasts between the genuine and not, but none of them, down on the ground of contextual detail—details that answer Wittgenstein's important question 'In what *sense* is it a genuine visual experience?'—correspond to the idea that the one category will be genuine, or objective, or skepticism-proof perception, while the other only concerns subjective or skepticism-inviting aspects. The one language-game is in fact *not*, despite our pre-investigative intuitions, primary to the other.

Wittgenstein does not, characteristically, directly repudiate the interlocutor's presumption, but rather, throughout a large number of examples and detailed considerations—indeed, by disentangling the tangle—provides the means to 'shift' when we need to, i.e. to break the twin molds of the generic objective–subjective and perceived–projected dichotomies. Contrasted to this underground method (of the kind we saw in Chapter 4, Section 3), we can now see that Murdoch's assertion was, for better or worse, more of a frontal assault on the interlocutor's presumption. By saying, with striking force in her diary entry of 17 October 1947, that even were she presented with 'documentary evidence—for example, journals—about her forgotten past', she would nevertheless 'feel I didn't know what my past really was', Murdoch stands the intuitive presupposition concerning the priority of the one language-game to the other on its head. Again, Wittgenstein does not do this, for the reason that meeting a philosophical thesis with its polemical antithesis does not grant the distinct variety of intellectual freedom, the 'shift' to what we might call our way of seeing,[48] that he is pursuing throughout his writings. But it is of interest for present purposes just to see that such an inversion *can* plausibly be made. In insisting that she still would not know what her life was like, what her past *really* was, Murdoch is giving priority not to the seeing of the faces before noticing the likeness, not to the hearing of the theme before noticing that it is a variation. She is giving priority, indeed, to just that part of the past that is unfrozen, the part that calls for an active rethinking. The internal relations that we either suddenly, or slowly, come to see, the aspects that dawn on the interrelations connecting our life events and that yield

[48] Here again, for a lucid and thorough study of Wittgenstein's writings emphasizing the central role of visual experience and its nuances (and the significance of these for philosophical method and the effecting of a change to our way of seeing a problem or whole problem-field), see Judith Genova, *Wittgenstein: A Way of Seeing* (London: Routledge, 1995). I further discuss the kind of work that allows the 'shift' mentioned here in connection with our thinking about first-person description in the next chapter.

coherence, that yield the narrative thread, are for Murdoch the primary parts of understanding a life. They are—to put it too briefly—the sense-making contents of autobiographical consciousness.

If our first reaction to Murdoch's striking claim is that it is willfully epistemologically perverse, wildly idiosyncratic, or obviously disingenuous and stated for dramatic, polemical flair—if indeed our first reaction is to reply to her claim about the insufficiency of documentary evidence with the question 'Well, what more could you want?'—perhaps a new aspect is waiting to dawn on that reaction and its motivations. For it becomes, for a close reader of Wittgenstein's §xi, increasingly clear that such a reaction is only the surface manifestation of a buried presupposition concerning the genuine versus its opposite, the real versus the merely fanciful, and, generically, the objective versus the subjective. What more Murdoch wants, of course, is not the written descriptions of the isolated episodes of a life, but (and here we can further specify the too-brief point made just above) rather the content—what we might in this context mark off as the genuine content—of the *sense* of a life. We have seen that Wittgenstein, throughout his multifarious writings on language and mind, rejected (or, better, undercut) the picture of human experience that both traditional empiricism and behaviorism share, i.e. that we subjectively construct the objects of the world out of objectively given raw data. That form of scientist reductionism seriously miscasts the nature of our perception. And that model, applied to the perception—the understanding—of a life's past would yield only a parallel miscasting of autobiographical reflection, precisely in its suggesting that the raw data of episodic experience is objectively given and the subsequent perception of the networks of internal relations connecting them is merely a matter of subjective projection.

Yet there remains a further problem. While these reflections may effect, or at least encourage, a 'shift' of thought of a much-needed liberating kind, still we want to know: If Murdoch demands a continual re*thinking*, does this not in its very terminology imply that the autobiographical activity she describes, and that is for her nothing less than a moral imperative, is in truth more *thought* than it is '*seeing*' in a new light a formative past experience? If so, does this not itself diminish the applicability of visual aspect-perception to aspects of selfhood? Wittgenstein voices this problem generically, twelve pages into §xi, with the question (asked after a number of examples of aspect-perception) 'Was it *seeing*, or was it a thought?' (p. 204). We already know, given

the preceding, that this question is too simple in its formulation. Wittgenstein, cueing our alertness to the presence of an underlying simplifying conceptual template or philosophical picture by voicing a phrase of the interlocutor's that is quickly and explicitly shown to be merely a grammatical manifestation of a picture-driven impulse, writes: ' "The echo of a thought in sight"—one would like to say' (p. 212). The echo of a thought would be, in this sense, a thought that gave the content to what was subsequently seen. But that dualistic way of putting the matter drives a wedge between what we will, in conformity to this bifurcated picture where intellection precedes and shapes sensation, now call (1) the intellectual content of the visual experience and (2) the purely sensory content of that experience. This is the bifurcation that Wittgenstein's remarks throughout this section intricately dissolve, and it is the bifurcation standing behind the question asking if the work of autobiographical rethinking is more thought than seeing (and thus if it is, after all and beneath everything else that has been said, still the equivalent in the realm of self-knowledge of Hume's conception of causation).

Wittgenstein shows that there is, apart from the particular cases of the kind considered above which on inspection do not correspond to the generalized bifurcation and in fact strongly argue against it, no sharp delineation between what we are led to call the intellectual content and the sensory content, between thinking and seeing, between mind and eye. One feels here an impulse to use the word 'suffused' as a way of reaching for a general formulation of the relation, i.e. the intellectual content suffuses the sensory data in an indissoluble union, but this too should, as Wittgenstein's inquiries here implicitly demonstrate, be resisted: even if better, the concept of suffusion enforces an implicit bifurcation at a prior state now gotten beyond. The word 'indissoluble' does the same, and working in concert they would lead us to picture the thought-suffused perception as a result of a prior assemblage of components of sensory and mental ontologies, thus repositioning, near the end of our inquiries, the very picture that empiricism and behaviorism share as discussed above. The grooves are deep.

If we, with Wittgenstein in §xi, shift our way of thinking out of these traditional grooves, we will come to see that the problem that reasserted itself just above upon realizing that Murdoch is after all calling for re*thinking* is not one that should be answered *generally*. It is not genuinely quieted by answering that aspect-perception is primarily cogitation and only subsequently sensory, nor the reverse, nor any other

ratio in general. It is true that we can arrange cases on a continuum ranging from, say, Leonardo's cracked-wall 'landscapes' on one extreme to the more disciplined, or restricted, case of the duck–rabbit (p. 194) oscillation in the middle (in that it still requires imagination to see one or the other—whichever one we do not start with), and on the other extreme the 'double cross' (p. 207) image (a white cross on a black background or vice versa) requiring what could be merely an optical switch without a (change of) concept (it is still a cross, black or white). And in that context of ordered cases we might sensibly speak of more thought in the seeing of landscapes than the seeing of the duck, and more thought in the seeing of the duck than in the seeing of the white cross. But here again, it emerges that this is not *at all* to say that this continuum corresponds to a continuum ranging from the objective to the subjective, the perceived to the projected, the fact to the fiction. And autobiographical rethinking is perfectly analogous to this: we *can* order cases on a continuum, but the criteria for the confirmation, or the hesitant acceptance, or the probability, or the plausibility, or the possibility, or the feared probability, or the minimally plausible, or the highly unlikely, and the wholly disconfirmed sets of connecting internal relations, of recontextualizing juxtapositions, and of rethought linkages between life-events, will appear *in* context and—exactly as we saw in the visual case—nowhere else. From such a detailed perspective—precisely the perspective Wittgenstein offers with regard to visual experience in §xi—the attempt to state generally the verification conditions for self-revelatory aspects that dawn in settings of self-investigation will indeed too closely resemble the attempt to repair a spider web with our fingers (*Philosophical Investigations* §106). What Wittgenstein called a perspicuous overview of the concept of aspect-perception is gained through a patient, case-by-case consideration of the polymorphous nature of reflective and imaginative seeing, and the more we know of these—as we do at the end of §xi—the less likely we are to yield to the impulse to reduce all of these to one paradigm and then to generalize from that.

Murdoch wrote, '*Re-thinking one's past is a constant responsibility*': it should be constant because of new light shed by the ongoing recon-textualization of our past deeds, words, and thoughts, because of the new or deepened ways of seeing ourselves brought in by sets of internal relations awakened by active retrospection. Wittgenstein's extensive and fundamentally important remarks on aspect-perception show, by ana-logy, how to answer the questions to which Murdoch's diary entries gave

rise at the outset of this section above. We make a life-defining picture of ourselves by awakening those sets of relations and connecting the 'dots'—the important life-experiences of which Murdoch wrote—with a narrative thread.[49] And that ongoing work-in-progress then becomes a picture we come to resemble, in that it determines which experiences are salient and which not, thus shaping, at least partially, our subsequent choices in response to the picture, the unfolding narrative. We selectively attend to a life's events accordingly, and we, in Murdoch's sense, can 'alter the angle' by controlling the internal relations, the life-structuring juxtapositions, the sets of associations awakened by one aspect or another. A heightened awareness of the deep analogy between visual aspects and the process of seeing ourselves can motivate for us, as I think it did for Murdoch, skepticism concerning the explanatory power of life-events *simpliciter*. And this shows us how to understand the past as 'unfrozen' and how we might more exactly articulate the process of Murdochian rethinking. Wittgenstein in the course of his investigation wrote: 'One *kind* of aspect might be called "aspects of organization".[50] When the aspect changes parts of the picture go together

[49] This is not to suggest that the line of demarcation separating important from unimportant life-events is invariably a clear one. What appears to be of great significance at first glance can recede in retrospect, just as what initially may appear insignificant may emerge, once connected to larger or longer life-themes or once set in striking contrast to one's situation in the present. To take one example (helpfully suggested to me by William Day and Victor Krebs), one may find oneself reflecting on *what at the present moment* one takes to be an utterly trivial or unimportant event in one's life, such as remembering that as a child the milk got delivered to your home, to the back door, and that sometimes you would see the milkman and greet him and converse amiably, and sometimes not, and find that it *then becomes* important in one's self-interpretation—say because you now find it to be a touchstone of sociability and uncomplicated human connectedness that you presently recognize you've lost with your neighbors, store clerks, etc., as if your heart has, gradually over many years, turned cold to the mass of humanity. Thus, what may initially appear trivially insignificant as an idle memory may be anything but that (and reverse cases, of retrospective diminutions of the significance of events, could easily be found as well).

[50] Here again, it helps to keep in mind that Wittgenstein is not working toward a reduction to a single comprehensive account of aspect-perception or seeing-as. On the contrary, he is continually adding layer after layer of complexity, of difference, of case-supported nuance. Brian McGuinness has written, importantly: 'The reader feels challenged by all Wittgenstein's writings, but it is an error to hope to reduce their message to a system. Better in the first place to feel their complexity, for in large part this complexity, the amount there is to be thought about in life—not excluding intellectual work—is their message. "I'll teach you differences", as Wittgenstein used to say (a quotation from *King Lear*)' ('The Lion Speaks, and We Don't Understand: Wittgenstein after 100 Years', in *Approaches to Wittgenstein: Collected Papers* (London: Routledge, 2002), 8). This fundamental point of interpretation, brought to bear on the

which before did not' (p. 208). Aspect-perception and autobiographical self-description are, it can gradually dawn on one, two parts of our picture that, indeed, go together.

The consideration of the nature, the character, of the perception of persons—particularly where this concerns the self's perception and description of its own 'unfrozen' past, has brought us to this point. From here, we want to ask what more we can say about memory in and of itself, and then how what has been developed as the therapeutic conception of philosophical progress might relate to autobiographical work.

present subject, would give us a sense of an ever-expanding overview rather than anything resembling an analysis yielding an account that can then be applied directly from the visual to the autobiographical case. This is a different conception of philosophical progress.

7

Rethinking Self-Interpretation

IF forced to capture in succinct form the difference between the writing of a biography and an autobiography, we might naturally appeal to the distinctive role of memory played in the latter. While it is true that a biography or memoir of a person we know, or knew, does depend on memory (as does Norman Malcolm's memoir of Wittgenstein,[1] for example), it is of course common to write biographically of a subject we did not know personally (as in Ray Monk's biography of Wittgenstein[2]). In the autobiographical case, however, it seems impossible to escape this fact: Even where the autobiographer relies upon what we will in this context call externals, e.g. letters, documents, photographs, diaries, calendars, journals, and countless other bits of data, in the case of the autobiographer—unlike the biographer—that data will have a memory-triggering function. That is to say, the autobiographer will use all the assembled materials not, as does the non-personally-acquainted biographer, to assemble a mosaic of what the subject must have done and might have experienced. Rather, the autobiographer will use those materials to stimulate, or revivify, memories of the events, actions, and experiences indicated by each bit of assembled data. And if the autobiographer finds himself or herself utterly at a loss to remember the event indicated by a given piece of memorabilia, e.g. a train ticket, we will doubt (assuming the memory of the autobiographer is otherwise fully intact) the veracity or accuracy of the evidence—in this case that the train journey was taken by the subject. (Perhaps he bought the ticket in his name but for the use of his daughter; perhaps he had to cancel the trip at the last minute; etc.)

Clearly, no such criterion of memory applies in the case of biographical writing of the non-personally-acquainted kind. And, while memory may

---

[1] Norman Malcolm, *Ludwig Wittgenstein: A Memoir (with a Biographical Sketch by G. H. von Wright)*, rev. edn (Oxford: Oxford University Press, 1984).

[2] Ray Monk, *Ludwig Wittgenstein: The Duty of Genius* (New York: Free Press, 1990).

well enter as a criterion in a biography of the personally acquainted kind, it will of course still not play the distinctive role that it does in the autobiographical case: Malcolm may remember what *he* said or did in connection with Wittgenstein, but the subject of the memoir is *Wittgenstein*, whose memories he of course will not have, as we say in such contexts, from the inside. Only the subject himself or herself can stand in that unique relation (to put it in brief and formulaic terms, a relation of identity between remembering and remembered self); memory functions as a criterion, as we think of it in accordance with our picture of the general distinction between biographical and autobiographical writing in a manner whose uniqueness is preserved by metaphysics. And the philosophical picture of how precisely this autobiographical memory would—according to that picture—work is hardly unfamiliar.

1. AUTOBIOGRAPHICAL MEMORY

It was perhaps Locke who most clearly enunciated the conception of the mind as a repository of ideas: here the experience impinges upon the mind (or makes its impression upon the *tabula rasa*). The sensory experience *itself* in life is of course then gone, and, as sealed in the past, unrepeatable, but its *image* is retrieved and brought before the inner gaze, before the mind's eye as it is focused upon the memory-images of its own past experience. And this, as we shall see below, also corresponds to a picture, or metaphysical model or conceptual template, of how recognition works: we match the present, new sensory impression with an image called up of the old impression; recognition just is, according to the picture, indeed *re*-cognition—recognizing is construed as an act of matching the present impression to its predecessor-image. Locke's successor Hume then saw a problem concerning just how we really distinguish the one from the other (i.e. the present data from the past image), since they match exactly (in a present successful case of recognition). His answer to this was provided, as we know, in terms of differences of force and vivacity—the sensations of the present are distinguishable from presently called-up images of past sensations because they are more vivid in the mind and make a more forceful impression. And Hume's delayed successor in turn, Russell, kept the underlying conceptual model intact while giving a slightly different answer regarding our capacity to mark the contrast between the present

impression and the past impression-image: the past impression-image, as brought into the mind as an *object* of consciousness, carries with it, or induces, a sense of familiarity that marks the contrast between the otherwise indistinguishable mental objects.

All of these specifications of the distinctive experience of memory (or its close cognate, recognition), as further elucidations of the subject's private act of remembering, are thus at the same time further specifications of the autobiographer's self-investigative project. But while familiar, they too, as picture-driven and overgeneralized ways of characterizing an involved and intricate set of human self-descriptive phenomena, seriously mislead.

Let us recall, one final time, this centrally important passage of Donald Davidson's, but this time given in full:

There is a picture of the mind which has become so ingrained in our philosophical tradition that it is almost impossible to escape its influence even when its worst faults are recognized and repudiated. In one crude, but familiar version, it goes like this: the mind is a theatre in which the conscious self watches a passing show (the shadows on the wall). The show consists of 'appearances', sense data, qualia, what is given in experience. What appear on the stage are not the ordinary objects of the world that the outer eye registers and that the heart loves, but their purported representatives. Whatever we know about the world outside depends on what we can glean from the inner clues.[3]

This picture of the mind and its contents, in multiple ways we have considered, is indeed ingrained in our philosophical tradition. It is a prerequisite to unearthing this picture and criticizing it in the light of day—in the light of our actual variegated practices as they proceed in the stream of life without subservience to the conceptual picture that is created in the attempt to unify them, to give them a general theoretical formulation—that we not only articulate the picture clearly (as Davidson has done). We also need, as Wittgenstein has shown, to understand a number of the moves of thought, the images employed, the pressures on our thinking, that taken together generate a false necessity concerning this way of thinking, this picture of the mind.

This picture powerfully reinforces the very idea of *objects* of consciousness; it is just these that would be inspected with the inward-directed act of perceptual scrutiny. And our image of memory—or, perhaps more accurately, our image of the rememberer (as we may see in our initial

---

[3] Donald Davidson, 'Knowing One's Own Mind', in Quassim Cassam (ed.), *Self-Knowledge* (Oxford: Oxford University Press, 1994), 61.

image of Augustine writing his *Confessions* (as examined in Chapter 4, Section 2, and Chapter 5, Section 3, above) or our image of Dostoevsky's Underground Man (as examined in Chapter 4, Section 3, above), hermetically sealed within a private world of self-reflection—is assembled around this fundamental idea of memory *object*: in remembering, whatever else may happen (e.g. triggered associations, awakened related memories), the essence of the phenomenon is that we retrieve, and bring into the theatre, the particular and determinate object of consciousness that is the memory. Davidson adds, a bit later in his discussion, that the 'solution in the case of mental states' is to 'get rid of the metaphor of objects before the mind'.[4] 'Of course,' he adds, 'people have beliefs, wishes, doubts, and so forth; but to allow this is not to suggest that beliefs, wishes, and doubts are *entities* in or before the mind, or that being in such states requires there to be corresponding mental objects.'

This last line is particularly significant for our present purposes: a natural but (as we shall increasingly see as this discussion proceeds) deeply misleading image of memory can influence our reflections on autobiographical writing in precisely the same way. People have beliefs, wishes, doubts—and of course memories—but in believing, wishing, doubting, or remembering, there need not be a mental *object* as the center of inward attention and that is necessary for the very cogency of the concept 'memory'.

The problem with this model or picture of mental content, for Davidson, is two-pronged. First, the dogma, as he calls it, 'that to have a thought is to have an object before the mind' runs afoul in this way (Putnam makes this argument[5]): an object, if at least partly constituted by external relations or relational properties, cannot reliably be held to correspond directly, in one-to-one fashion, to an object in the mind—a thought, as construed on this model—because the thinker of the thought, its possessor, may be wholly or partly ignorant of those identity-determining external relations or properties. Second, Davidson adds (to his encapsulation of Putnam's view) that we cannot then look to *another* kind of object—presumably one not partly constituted by external relational properties, because *that* object (in so far, I would add, as we can intelligibly conceive of it in the first place), in not being connected to the world, would not allow us to learn about the object in the world by inwardly contemplating the mental object. It,

---

[4] 'Knowing One's Own Mind', 62.
[5] Davidson reviews Putnam's argument, ibid. 63.

to further encapsulate the point, would be semantically detached, and thus not able to provide what we fundamentally want in bringing it, the thought-as-mental-object, before the mind's gaze in the first place. The deep significance of our modeling of autobiographical memory in just this way is perhaps apparent: if a remembered event is represented internally by such an allegedly corresponding thought-object, and if that remembered event is in part relationally constituted, then we cannot assume that the event does in fact correspond to the object before the remembering mind. And with this skeptical gap, we would forever be wondering if we, quite literally, knew what we were, if not talking, then thinking, about. If the memory-object is made wholly knowable by the remembering mind by severing its external relational-connectedness, then it loses what Davidson called, succinctly, its semantics. It would be, like Hume's sensations that are always unto themselves 'right' because they refer to nothing beyond themselves, a memory free-for-all, with no external constraint on memory rightness. (Here Wittgenstein might say, contrary to Hume, that thus here we can't talk about 'right'.)

Assessing the force of this last point, Davidson writes: 'The only object that would satisfy the twin requirements of being "before the mind" and also such that it determines what the content of a thought [is] must, like Hume's ideas and impressions, "be what it seems and seem what it is". There are no such objects, public or private, abstract or concrete.'[6] Taken together, these reflections give us growing reason to doubt that such mental objects, as alleged centers of inward-directed attention, are necessary in the way we initially pictured. Moreover, if they were to exist, either they would—if envisaged according to the underlying model or picture—fail to correspond to the external relations that in part individuate the remembered event, or they would lose their semantic link to the world, thus severing any independent criterion of rightness in memory.[7]

Davidson emphasizes the conceptual rewards of freeing ourselves from the grip of the picture of thoughts-as-mysterious-mental-objects. With this freedom, he suggests, we will (1) be able to see the natural linkages between the social nature of language, i.e. the external or relational determinants of thought and meaning, and (2) secure the

---

[6] Ibid.

[7] It emerged during the investigations into the much-discussed false-memory syndrome of the 1980s and 1990s that one psychotherapeutic camp explicitly (and ruinously) did just this, declaring it a methodological principle that every memory voiced by a patient is to be accepted, and reinforced, unquestioningly.

preservation of, rather than succumb to the behavioristic eradication of, first-person authority.[8] Again, the parallel point would be true of the picture of memories-as-mysterious-mental-objects: with the image of object-storage and object-retrieval (another instance of dangerously modeling the inner on the outer) removed, we would indeed have our view opened to a multiplicity of aspects of memory presented within autobiographical writing, subtle particularities discernible only in context that are otherwise crushed by the blunt force of the material-object model.

Wittgenstein labored both mightily and intricately against this falsifying, and conceptually blinding, model: he worked through the aforementioned pressures on our thinking propelling us toward the model in the first place. Wittgenstein, as we shall see, describes this picture with varying inflections in a number of philosophical contexts throughout his work. First, in a passage, the opening part of which is already familiar to us (*Philosophical Investigations* §604), he writes:

It is easy to have a false picture of the processes called 'recognizing'; as if recognizing always consisted in comparing two impressions with one another. It is as if I carried a picture of an object with me and used it to perform an identification of an object as the one represented by the picture. Our memory seems to us to be the agent of such a comparison, by preserving a picture of what has been seen before, or by allowing us to look into the past (as if down a spy-glass).[9]

Comparison, of course, requires that there be at least two objects or entities juxtaposed in such a way that we can discern the similarities and differences made evident through the back-and-forth scrutiny. Recognition, Wittgenstein also here claims, is too easily construed as comparison-conducted-inwardly, i.e. where, on the Lockean or Humean model, an impression is compared with an idea (or a sensation with a reflection). Wittgenstein describes the process of recognition as one akin to comparing a picture we carry with us with an object in our visual field; if comparative scrutiny yields a match, the recognition is successful. Then turning this inward, i.e. into a wholly mental act, he adds that our memory, taken as the agent of the comparison, is pictured as a preserver of images of what has been seen before. And, making the link to our present concern perfectly explicit, he adds that this is what

8 See Davidson's discussion in 'Knowing One's Own Mind', 63–4.
9 3rd edn, trans. G. E. M. Anscombe (New York: Macmillan, 1958).

seems to allow us to look into the past. This much is fairly familiar to us from the foregoing. But this is hardly the end of the issue.

Wittgenstein articulates a striking set of observations against this way of seeing memory and recognition in *The Blue and Brown Books*.[10] There, identifying this picture as primitive (the picture, as he encapsulates it within that discussion, of 'comparing the man we saw with a memory image in our mind and we find the two to agree'), he himself offers the fairly blunt observation that in most cases of recognition no such comparison takes place (by contrast with the more rare contexts in which it does take place, e.g. when we recognize a criminal by his photograph in the post office). The temptation to picture the matter in this way, he there indicates, is fueled by the simple and accessible empirical fact that there are memory-images. But the most powerful reminder of what we *actually* do comes next in his discussion: such an image, for example seeing in our mind's eye how the person looked when we last saw him a decade ago, comes into mind just *after* we have suddenly recognized him upon running into him in the street. The mental image, whatever role it may play, is not a precondition for recognition. Nor is it, *mutatis mutandis*, the precondition for memory. And in *Philosophical Investigations* §648 he very briefly describes a case, familiar enough, in which remembering clearly takes place but where mental images play no role whatsoever: ' "I no longer remember the words I used, but I remember my intention precisely; I meant my words to quiet him." What does my memory *shew* me; what does it bring before my mind? Suppose it did nothing but suggest those words to me!—and perhaps others which fill out the picture still more exactly.' It is the attention to particular cases that here again exposes the false sense of the obvious truth of the picture. And this kind of philosophical attention allows us to unearth and expose the presuppositions that would blind us to the intricate details of the rich weave of human memory-practices, a clear view of which is indispensable to our growing understanding of how autobiographical language actually works. In *Philosophical Investigations* §651 Wittgenstein presses the point further: ' "I remember that I should have been glad then to stay still longer."—What picture of this wish came before my mind? None at all.' This quite evidently reminds us that image-consultation is neither necessary nor sufficient as the essence

[10] (Oxford: Blackwell, 1958), 165. For a set of very helpful references leading to various but interrelated texts of Wittgenstein's, I am again indebted to Hans-Johann Glock, *A Wittgenstein Dictionary* (Oxford: Blackwell, 1996), 'Memory'.

of the concept 'memory'. And in the next line Wittgenstein both strengthens and extends the point to our memories of our feelings, in the case above of the precise emotional tone in which we experienced the wish to stay longer at the time: 'What I see in my memory allows no conclusion as to my feelings. And yet I remember quite clearly that they were there.' Indeed, the image of the remembered scene, even if it were there and functioning (which in this particular case it is not—his words were 'None at all'), would still not be the right tool with which to recover the emotional content of the wish—and yet one clearly can remember the emotional tone of such remembered wishes.

It would be a measure of the grip of the original picture if we were then, in light of cases of these kinds, to reduplicate the image-consultation model but in linguistic form, i.e. if we were then to say that the essence of remembering is to bring before the mind a visual image of the *words* we uttered at the time. Similarly, we might grasp at an image of our behavior at the time as the memory-content in our Cartesian theatres. But in *Remarks on the Philosophy of Psychology*,[11] volume i, §467, we find, in the context of an investigation into observing our own states of mind (with 'hoping' serving as the particular case), we find this: 'The word "observe" is badly applied here. I try to remember this and that.' And immediately following this (in §468), we get 'If someone remembers his hope, on the whole he is not therefore remembering his behaviour, nor even necessarily his thoughts. He says—he knows—that at that time he hoped.' This is underscored, in *Philosophical Investigations* §660, with the comparison between the grammar of an expression of the 'I was then going to say . . .' kind with the 'I could then have gone on' kind. In the former case, Wittgenstein remarks, we are remembering an intention, and in the second kind of case we are remembering having understood (in such a way that we then have more, or are able to say more, about the thing understood). The similarity of the grammars of these phrases is that, initially counterintuitively, we clearly can remember what we were going to say, but *without* thereby bringing to mind a mental image of the particular words or the phrase that we would have said. The phrase 'I could have gone on' suggests less determinacy of the linguistically pre-cognized kind; the remembered intention is a good deal more like this than we are, under the influence of the picture in question, inclined to think. And much retrospective self-understanding is precisely of this

kind, where we remember that at a given juncture we could well have said more than we did—we knew how to go on—and that that ability, unrealized, was a kind of outgrowth of our understanding at that time. The memory can indeed be determinate, but its content-determinacy is not given by the precision of an image, be it of a visual or a linguistic kind. Nor, bringing Wittgenstein's observations together, is the emotive content or tone of the remembered event given by an *image*.

2. THE 'DESCENT INTO OURSELVES'

At the outset of the preceding section we saw that, if there is an intuitive distinction between biographical writing and autobiographical writing, it is the evident reliance on first-person memory in the latter that makes the difference. But with both the Davidsonian and Wittgensteinian reflections behind us, how then should we think of the experiential content of a first-person memory? The answer—perhaps somewhat startling, given our intuitive expectations—Wittgenstein suggests, is that perhaps we should free ourselves of the very notion of *experiential content*, that is, as a uniform mental phenomenon that constitutes remembering, one that gives the word 'memory' its meaning. This suggestion is of course part and parcel of the challenge to the 'naming presumption' in earlier chapters, i.e. the presumption that psychological nouns and verbs get their meaning by naming mental objects or processes respectively.

We think of memory as showing us the past, and of the auto-biographer's memory as the instrument or device that shows him his own past; autobiographical truth will then be regarded as an accurate verbal description of what he sees therein. But Davidson suggested that the inner mysterious object would go the same way as the sense-data picture of experience, i.e. as that which shows us the outer world, as soon as we escape the pernicious dualistic model behind these views. Wittgenstein saw very much the same point: in *Zettel*[12] §663 he writes: 'But if memory shews us the past, how does it shew us that it is the past?' He replies to his own question, startlingly, as follows: 'It does *not* shew us the past. Any more than our senses shew us the present.' Our senses, conceived as data-collectors on the sense-data model, create the

---

[12] Ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967).

raw material for an inner picture of what we see before us. Similarly, the memory-experience would be the inner sense that collects the raw material for the inner picture of what we remember. Both, Wittgenstein and Davidson after him suggest, are deeply misled ways of construing our grasp of the present and our grasp of the past. In *Philosophical Investigations*, part II, §xiii, Wittgenstein makes this point in the most distilled form, reminding us that the very concept 'description' is out of place, and moreover deeply misleading when used in connection with the experience of memory. That concept, so used, *requires* that there be an object described.

Wittgenstein begins *Philosophical Investigations*, part II, §xiii (p. 231), with the sentence: 'When I say: "He was here half an hour ago"—that is, remembering it—this is not the description of a present experience.' That is, the act of remembering that he was here as manifest in our saying someone was here half an hour ago is not a description of the present inner content, the present visual image. But of course it is difficult to suppress the powerful sense that there simply must be such an image that constitutes the inner content of the remembering. And one reason for this insistent picture is that we *can* have what Wittgenstein calls 'accompaniments' to remembering—and these, to confuse the issue, may well be visual images. Thus, he continues the above discussion with: 'Memory-*experiences* are accompaniments of remembering.' That is to say, a vivid memory image may well spring to mind when we find ourselves reminiscing about a person, say a beloved grandparent, now gone. But the image is not *itself*, here again, either a necessary or sufficient condition for remembering; it comes as an adjunct of, or as an accompaniment to, the remembering. (To particularize the point, if this were the case, we would, in response to a sibling saying 'Remember how our grandfather always used to . . .', find ourselves unable to remember the grandfather *until* we called to mind the attendant visual image.)

It is also the case that many such images, as they may spring to mind in contexts of reminiscence, in truth are memory-images not of the persons themselves, but of photographs of those persons. And that is instructive for the easily observed reason that we here conflate the language-game of photograph-comparison (as in the criminal mugshot case above) and remembering. The former is always visual, the latter not necessarily so. Compressing the conceptual genealogy of this philosophical picture, Wittgenstein remarks a few lines later in the above discussion in *Philosophical Investigations* that we get the *idea* of a memory-content only because we assimilate 'two psychological concepts', and then,

quickly explaining that, adds: 'It is like assimilating two *games*. (Football has *goals*, tennis not.)' (p. 231). Comparing a person to a photograph, like comparing one photograph with another, is obviously and inescapably visual. Remembering—for all the various and divergent things that that can mean as we explore (by reading autobiographical writing in a manner mindful of these philosophical issues) the nuances of particular cases—need not take that form, and thus need not have visual content. And where it does, that content may well enter the scene only *ex post facto* with regard to the act of remembering—it may be a subsequent accompaniment.

Breaking the spell of the visual picture is also aided by thinking of cases ranging through the other four senses: If we begin an invitation to recollect with the line 'Remember when . . .', the impulse to unify, and to reduce all variations to a common essence, can lead us directly to the visual-image picture. But we should consider as well invitations beginning with 'Remember how those cherries from the tree in our backyard tasted', or 'Remember the feel of that sea-island cotton . . .', or 'Remember the aroma of . . .', or 'Remember the melancholy voice of the oboe in . . .'. In each case, we may also recall a visual image of the cherry tree, the garment, the kitchen, and the oboist, but the remembering of the four nonvisual experiences is not simultaneous with, nor identical with, nor reducible to, those images. It is also instructive to bear in mind that the word 'remembering' can itself impose, or strongly incline us toward, the visual picture, where the word 'reminiscing' does this to a perhaps lesser degree. Indeed, the idea of reminiscing seems to bring the verbal to the fore (in part because one usually reminisces socially, with another person; remembering may or may not have this social dimension).

The very concept of description, once in place (in place, as Wittgenstein has shown, where it should not be, i.e. where it is placed in correspondence with the picture of memory here under review rather than in accordance with the facts, the human practices and experiences that picture is supposed to capture succinctly), leads to an ordering whereby the alleged memory-experience is first and the description of that inner experience follows. In *Remarks on the Philosophy of Psychology*, volume i, §1131, Wittgenstein articulates a densely compressed three-stage argument yielding a conclusion of considerable power vis-à-vis our understanding of memory-fueled self-awareness. Announcing the subject with the word 'Memory', he gives, as the first stage, the example 'I still see us sitting at that table'. But then he asks: '—But have I really

the same visual picture—or one of those which I had then?' That is, if the view of memory presumed by Locke, Hume, and many others (including Augustine, as we shall see shortly) were correct, i.e. if the reflection were cut, as a seal, from the sensation, or if the idea were an only-less-forceful-and-vivacious version of the sensory impression, then the memory image would of necessity so correspond. But we do often picture past events in our imaginations, and as stimulated or occasioned by passing mentions of those events, in ways constructed or deriving from, but not necessarily in direct correspondence to, those past experiences. The striking and picture-unseating fact here is that language, in such cases, *precedes* the image. Then the second stage advances: 'Do I also certainly see the table and my friend from the same point of view as then, and so not see myself?' That is, the image called to mind in the course of our reminiscing conversation may well be an image that includes oneself in the scene—which evidently could not be the case were the empirical model correct. We do not see ourselves (except in mirrors, reflections, etc.) from our own distinctive perceptual vantage points, and it is of course only from those that we could have gotten the memory-originating sensations or impressions. To state the matter starkly, the linguistic is not (whatever else it may be) in a position of descriptive subservience to the visual. Then comes the third stage: '—My memory image is not evidence of that past situation; as a photograph would be, which, having been taken then, now bears witness to me that this is how it was then.' The photograph is indeed the central component of a separate language-game; it is not only an externalized version of what we carry and what we call up inwardly as the content of the memory. This is because it exhibits evidential weight about the past in a way and with specificity that the memory-image does not possess. Contra the empiricist picture, it is not the case that the memory-image, like a lightly contrasted black-and-white photograph in comparison to a full color photographic print, is a lesser version of that prior visual sensation. Here is the conclusion following the three stages: 'The memory image and the memory words are on the *same* level.'[13] The linguistic dimension of memory is not, in the sense this picture would dictate, a description, and it thus is not secondary to the mental object it allegedly describes.

The empiricist picture (again, a conceptual model that is hardly restricted to the empiricists alone) leads us astray with false analogies,

---

[13] A variant of this remark (suggestive of the remark's wider significance) is found in *Zettel* §650.

and it blocks our view of the multiform ways in which the words or phrases 'to remember', 'remembering', 'remembered', 'recollected', 'reminisced', and cognate terms are used. And a perspicuous overview of those uses is, as always (and as we shall see more fully in the subsequent section of this chapter), conceptually clarifying. These considerations show some of the thought—the analysis of the conceptual genealogies that lead to these powerfully influential and often submerged simplifying pictures of the acts of memory—that leads to Wittgenstein's exchange with his imagined interlocutor. That exchange was with the voice that does not remain vigilant about the grammatically concealed power of the misleading analogies we have heretofore excavated. Recall, now at this late stage, *Philosophical Investigations* §305, which indeed we are now in a position to see in a new light:

'But you surely cannot deny that, for example, in remembering, an inner process takes place.'—What gives the impression that we want to deny anything? When one says 'Still, an inner process does take place here'—one wants to go on: 'After all, you *see* it.' And it is this inner process that one means by the word 'remembering.'—The impression that we wanted to deny something arises from our setting our faces against the picture of the 'inner process.' What we deny is that the picture of the inner process gives us the correct idea of the use of the word 'to remember.' We say that this picture with its ramifications stands in the way of our seeing the use of the word as it is.

If memory is not, after all, dependent upon an 'inner process' as here pictured, i.e. if it is not given content by the memory-image, and thus the (phantom) process is not the act or the entity that provides the referent and hence the meaning of the words 'to remember', then the act of remembering as modeled according to this picture is not a source of knowledge *in the way we initially think*. We do not learn about the past by scrutinizing (and then, in autobiographical writing, describing) the inner process. The true nature of autobiographical reflection takes a different, far less unified course, and the meaning of 'to remember' is not reducible to a unitary inner process that provides the essence of the concept. We usually do not 'read off' the content of the memory from any inner entity, nor do we learn the concept of memory from inward-directed ostensive definition. Yet it would be wrong, as suggested in the preceding passage, to *deny* that remembering is a mental process (whatever, in varying particular contexts, might be meant by that phrase). Thus, in *Remarks on the Philosophy of Psychology*, volume i, §105, Wittgenstein writes, as an analogy to the case of hearing or thinking of a word in some determinate meaning: 'And is it like that

only with the experience of meaning? Isn't it so also with, e.g., that of remembering? If someone asks me what I have been doing in the last two hours, I answer him straight off and I don't read the answer off from an experience I am having. And yet one says that I *remembered*, and that this is a mental process.'[14]

These points, again, once grasped in an expanded context sufficient to encompass the pressures on thought, the intellectual genealogies that Wittgenstein assiduously traces that led to the illusory pictures they are meant to dispel, are of the first importance to our understanding of autobiographical writing. Our relation to our past is no more passive than is our relation to what we presently visually perceive: we are not the containers of memory-images that a true narrative would accurately describe. Rather, we are in a continual process of reconsideration (as discussed in Chapter 6, Section 3, in connection with Iris Murdoch), of reflective restructuring, and of repositioning the actions, events, occurrences, interactions, efforts, aspirations, achievements, intentions—in short, our words, deeds, and everything in between that, taken together, form the teleological trajectories, the narrative threads, of our selves. Such a developmental retrospective is never finally settled beyond the reach of rejuxtaposition with other related (and in some cases seemingly unrelated) life-events; such retrospective self-understanding is the result of an *active* labor of self-investigation, the content of which is dynamic, not static. And our relation to it is, in Murdoch's sense, 'unfrozen'. Memories, understood in this way, are not inert visual images filed into storage by time and date. They are remembered experiences of all composite kinds, and, like works of art and like human selves, they take on and cast off relational properties, networks of interconnections to other experiences both similar and different.

Some of these emerging self-narratives will carry deep conviction; others will seem plausible; others possible but doubtable; and still others implausible. And then some will be, as fanciful whole-cloth revisionism, rejected outright. But their placement on this epistemological continuum, if the considerations we have reviewed of Davidson's and of Wittgenstein's are right, will—highly instructively—*not* be determined by the extent to which these narratives function as verified descriptions

---

[14] An illuminating discussion (to which I am indebted) of this passage can be found in Joachim Schulte, *Experience and Expression: Wittgenstein's Philosophy of Psychology* (Oxford: Oxford University Press, 1993), 98 (see also the fine full chapter on memory, pp. 95–119).

of inner objects or memory-images. The concept of memory is, as we have seen, not reducible to that, but rather is shown in its use throughout the vast body of autobiographical writing that our tradition has generated. And, most fundamentally, the language of that active process is not, against the picture of memory articulated by the empiricists and as operative at various levels of awareness in many other philosophical and literary-critical places, secondary to genuine memory-content. In many cases, those verbal formulations, and recontextualizing reformulations, just are the content of memory. And in turn, this language is not—as we also may well initially picture the matter—secondary to an inner process that allegedly precedes it. Again, memory need not be a matter of 'reading off', and thus an act of memory—if we should call it that, i.e. name it in a manner suggesting a uniform inner experience—is not, in the manner we initially construe it to be, looking into the past. And the very phrase 'looking into the past', if taken as a rhetorical flourish that admittedly accentuates ocular experience and visual scrutiny, can be harmless if it is meant as a colorful way of alluding to the many and varying processes of retrospective self-investigation. If, conversely, it is taken to name a uniform process of mental-image retrieval and passive looking, then it can, in establishing deeply misleading analogies with only a few rhetorical moves, do far more harm than good. It prevents us from seeing—well, comprehending—the multiform employments of the word as it is used.

Before closing this penultimate section, there is one more strand to take up. I mentioned above that a more thorough way of dealing with entrenched philosophical pictures or simplifying conceptual models that obscure our view of our practices in the name of theoretical concision is, as Wittgenstein both recommends and undertakes throughout his writings himself, to work back through the genealogical lineage that makes such pictures, at a glance, seem obviously true and, as with the initial sense of the evident necessity of Davidson's combated inner objects, inescapable in providing any plausible account of the subject at hand. In the same way, Wittgenstein's combated 'inner process' initially seems equally inescapable. That the legacy of this way of thinking goes back to the classical empirical positions best articulated by Locke and Hume we have seen. But with a brief, close, and for us final look at Augustine's *Confessions*—along with our knowledge that Wittgenstein repeatedly read, and reflected upon, Augustine's great work of self-investigation—we can, I believe, see one of the early most exacting articulations of the positions unseated by the foregoing Davidsonian and Wittgensteinian observations.

Augustine describes memory as:

a spacious palace, a storehouse for countless images of all kinds which are conveyed to it by the senses. In it are stored away all the thought . . . and it also contains anything else that has been entrusted to it for safekeeping, until such time as these things are swallowed up and buried in forgetfulness. When I use my memory, I ask it to produce whatever I wish to remember. . . . Finally that which I wish to see stands out clearly and emerges into sight from its hiding place. . . . They return to their place of storage, ready to emerge again when I want them.[15]

This, it will now perhaps be evident, is a compendium of visually driven descriptions of what is taken to be the uniform act of remembering, in what Augustine calls 'the great storehouse of memory' and 'the vast cloisters of my memory'. He also says, tellingly, that in it 'I meet myself as well. I remember myself and what I have done . . .'.[16] And the retrieved visual image, in his attempt to capture the essence of memory with theoretical concision, occupies a central place: 'the images of all the things of which I speak spring forward from the same great treasure-house of the memory'. He underscores what he takes to be the epistemological power of such images and their priority to language: 'And, in fact, I could not even mention them at all if the images were lacking.'[17] Wittgenstein's investigations into these issues, and then, if perhaps in a less direct way, Davidson's also, read as a sustained critical commentary on Augustine's stated position. Each segment of Augustine's position is, as we have now seen, subjected to intense critical scrutiny: (1) the reduction of memory to memory-image storage and retrieval; (2) the centrality of the visual image to the experience of memory; (3) the very idea of the 'experiential content' of memory; (4) the 'inner process' that is the alleged referent of the words 'to remember'; (5) the subservice of the linguistic to the visual image and the resultant descriptive conception of autobiographical language; (6) the passive nature of our relation to that internally shown past; and, more generally, (7) the unification of all variegated memory usage into an essence-as-inner-process.

Wittgenstein's labors serve to free our thinking about reflexive thinking, about what kinds of divergent and active engagements constitute self-investigation; this is similar to the freedom from picture-driven

[15] Augustine, *Confessions*, trans. R. S. Pine-Coffin (Harmondsworth: Penguin, 1961), 214.   [16] Ibid. 215.
[17] Ibid. 216.

inquiry of which Davidson wrote. Of course, Augustine, like Wittgenstein and then Davidson (in some projects) after him, was struggling to shed light on our human condition vis-à-vis self-directed scrutiny. He wrote of his labors: 'For I am not now investigating the tracts of the heavens, or measuring the distance of the stars, or trying to discover how the earth hangs in space. I am investigating myself, my memory, my mind.'[18] That also describes Locke's and Hume's equally humane projects, as well as, as we have seen, the more recent cases. But the difference—a profound one made possible in and after Wittgenstein's work—is to proceed in a manner constantly mindful of the misleading power of conceptually embedded pictures or templates, and a vigilance with regard to the misdirecting influences of grammatical nuance and rhetorical phraseology in nourishing these illusions concerning what it is, in Wittgenstein's sense, to 'descend into ourselves'.[19] With that difference, we are enabled to see the countless cases of autobiographical writing for what they in truth are: particularized literary 'reminders', as Wittgenstein uses the term,[20] that are of the first importance in gaining a perspicuous overview of (and thus also of the first importance to achieving conceptual clarity about) our psychologically reflexive endeavors.

[18] Ibid. 223.

[19] It is important to bear in mind that, for Wittgenstein, the unearthing of such illusions and the tracing back through the complex grammatical genealogies that led us into them is by no means a matter of purely 'professional' interest: this philosophical work was itself what he called 'work on oneself'. The deeply personal nature of this drive to escape linguistically motivated illusion is well described by Norman Malcolm. Having just referred to Rush Rhees's mention of Wittgenstein's desire to come to 'recognize his own nature, ceasing to disguise it from himself', Malcolm writes: 'This need to understand his own nature was connected, for Wittgenstein, not only with his wanting to be a completely honest person, but also with the quality of his philosophical work. If he was not truthful about himself then his writing would not be truthful. In a notebook of 1938 he wrote: "If anyone is unwilling to descend into himself, because this is too painful, he will remain superficial in his writing." In the following year he wrote: "The truth can be spoken only by one who rests in it; not by one who still rests in falsehood, and who reaches out from falsehood to truth just once"' (Rush Rhees (ed.), *Recollections of Wittgenstein* (Oxford: Oxford University Press, 1984), p. xix). The kind of work undertaken by Wittgenstein as considered here—like a philosophically mindful reading of autobiography—is thus just such a 'descent into ourselves', and the intricate removal of misleading pictures, or 'disguises', thus earns the possibility of honest self-recognition.

[20] Reminders are, in this particular sense, detailed examples that show the fine-grained employment of the word, phrase, or concept in question. Such reminders simultaneously display the use of the word, the circumstances within which we show an unproblematic mastery of it, and the considerable distance between such real cases and overgeneralized and oversimplified pronouncements concerning their alleged unitary nature or essence.

3. ON PHILOSOPHY AS THERAPY:
WITTGENSTEIN, CAVELL,
AND AUTOBIOGRAPHICAL WRITING

At this stage we can see very clearly that, in his later philosophical writings, Wittgenstein was exquisitely sensitive to the misleading implications housed within the formulations of philosophical questions. The question with which he opened the *Blue Book*, 'What is the meaning of a word?', the question 'What is thinking?', and the question 'What constitutes understanding?' each put into place, as we have seen, presumptions concerning how an answer is going to proceed: in the first, that the meaning will be an entity of some kind—likely of an ontological kind different from the acoustical or calligraphic sensory property of the word (to which we will then think the meaning *attached*); in the second, that thinking will be in essence a determinate mental process or event—likely one of an ontological kind different from that of *speaking* (to which we will then picture thinking as prior); in the third, an inner process culminating in a state that is metaphysically hidden from all things outer and that is the inner condition that lies unreachable behind what we will then call the behavior that is contingently correlated with it. And as we have seen, each of these questions has been subjected to thorough scrutiny throughout the writings in *The Blue and Brown Books*, *Philosophical Investigations*, *Remarks on the Philosophy of Psychology, Last Writings on the Philosophy of Psychology, Zettel*, and even more extensively in the typescripts and manuscripts from which these are taken.[21] But it is still perhaps less widely realized than it might be that the very phrase 'Wittgenstein's method' can easily prove as misleading as each of the above questions, and there is an aspect of irony waiting to dawn on those who reflect that the modern philosopher most concerned to maintain a philosophically relevant mindfulness about front-loaded implications is perhaps too often not afforded precisely the mindfulness about his own later philosophical writings that he was most concerned to teach.

The phrase 'Wittgenstein's method' might well lead us to expect, in the first place, a unitary way of contending with philosophical difficulties that is employed across the board; in the second place, it might lead

---

[21] For a guide to the *Nachlass*, see von Wright's catalogue in his *Wittgenstein* (Oxford: Blackwell, 1982).

us to expect an acceptance of a set of perennial problems as framed by philosophers through the ages that are treated with that unitary method; and in the third place, it might lead us to expect a set of solutions to those time-honored questions providing answers formulated in a way parallel to, if different from, those worked out by Frege, Russell, Quine, or many others before and since.

There are numerous ways of attempting to describe Wittgenstein's method in brief scope, many of which contain perhaps a kernel of truth—but they also carry the danger of miscasting the spirit of the entire philosophical self-interrogatory undertaking. For example, one might say (1) that, if this philosophical-methodological stance is generically anti-metaphysical because the intricate particular cases in the real contexts of lived human life are so complex and variegated that they cannot successfully or accurately be subsumed under general categories or type-concepts, is *that* then not a quite overarching metaphysical thesis? Or one might say (2) that this is in effect a method carrying more than a trace of positivism, since they believed that a verificationist criterion could provide a sharp and failsafe criterion for separating out meaningless metaphysical pseudo-propositions from the meaningful propositions of science. Is this not a somewhat less unitary way of attempting just the same thing? Or one could say (3) that this methodological stance actually harbors a nominalist metaphysic at its core, placing so much emphasis on the study of language that under it all it asserts that, simply put, saying what it is makes it what it is. Or one could say (4) that there is an underlying animus against the postulation of mental entities, processes, and realities, and that it thus carries a materialist or physicalist assumption at its core. Although the rest of this chapter (like much that came before it throughout this book) is an attempt to characterize the spirit of the undertaking that these methodological reductions fail to capture, each of these might be given an answer here in short form.

To (1), one might answer that what one sees by turning to cases that is both philosophically valuable and methodologically salutary, whatever else it is, is *not* a general thesis *concerning* particularity—this has, I hope, been seen in very many examples throughout this book. (Although this is not the place to pursue the matter, this, if in a nutshell, is why what is presently discussed under the heading of moral particularism, for all that it in fact is, is not an extension of the Wittgensteinian tradition.) Nor is it a *general* polemical attack on metaphysics in any generic sense or on metaphysics *in toto*. It is rather a set of investigations into very particular metaphysical *pictures*, in

Wittgenstein's sense, as they would preclude taking complexity seriously, cut short the patience required to comprehend significant particularities, and allow drastic oversimplification to masquerade as an 'account'. (It is true that, throughout this study, I use phrases such as 'metaphysical baggage' to refer briefly to such pictures and the misleading analogies, motivated impatience, etc. that generate or nourish them—but again, these individual critical–diagnostic discussions do not themselves cohere into a single larger anti-metaphysical 'account'. A *blanket* critique would indeed here again prove antithetical to the spirit of the enterprise.) To (2), one might answer that, while indeed the kernel of truth is right (there is a continuing concern with distinctions between the meaningful and the meaningless), the contextually emergent details that serve to shed light on questions of the meaningful and meaningless will consistently be *sui generis*, and thus not transferable even as one of a collected set of meaning-criteria of a neo-positivistic kind. More importantly, the *very question* of the meaningful versus the meaningless is itself not a generic and context-transferable question template that automatically applies to all contexts or all propositions (indeed as though we utter 'propositions' in all contexts of human verbal interaction). To (3), here again it is true that there is an abiding concern with language, but to see nominalism in that is at the same time to fail to see that in so many cases (again as I hope we have seen throughout this volume) Wittgenstein is rigorously concerned with unearthing the sources of confusion as generated by language—language *itself* is hardly taken unquestioningly as an arbiter of what is the case or what does and does not make sense. (And more importantly, this in fact puts it far too generally and implies that we are transcontextually and always concerned with a generic question concerning what is and is not the case, and that we might find a single epistemological litmus test to apply to all cases.) To (4), one might say that the perceived generic animus is just what led to the misperception of Wittgenstein as a reductive behaviourist of the kind discussed in Chapter 3—as we have seen, his concern is to remove *confusion about* the mental, not the mental. But more importantly, if we stand back and ask if, on the whole Wittgenstein seems more concerned (*a*) to redescribe the mental in physical terms because he takes extended things to be real, unlike mental phenomena, or (*b*) to show both the *complexity* of the relations between the language-games of physical objects, bodies, and embodiment, and the language-games of the mental as well as the very great difference between the mental and the physical (where many positions in philosophy are objectionable

because they make that difference seem too *slight*), the answer would have to be the latter, that is, (*b*). And that is clearly incompatible with (and the point is, fundamentally different in methodological spirit from) any such governing physicalistic assumption.

Expectations of neat and concise methodological uniformity of a kind that generates similarly neat and concise position-formulations have often outlived the recognition that Wittgenstein's own writings are presented in a piecemeal way that seems not to offer any such overarching method or any such problem-accepting concisely formulated answers; some have simply assumed that, while Wittgenstein himself did not or could not provide such results, the raw material is there, and the task of interpretation, in his case, just comes to the task of argument-and-position formulation. To employ a common spatial metaphor that is loaded into epistemology, nothing could be further from the truth. What Wittgenstein's later writings offer us is—to put the matter in a different way misleadingly briefly—a way of seeing philosophical problems that constitutes, as has been claimed throughout this book, a *radical* departure from the approaches of Frege, Russell, Quine, and so many others. (But this is, of course, not to suggest that we will not gain much from keeping those authors, as foils, in view as we assess Wittgenstein's methodological radicalism—quite the contrary.) And in enacting this departure throughout his writings from the *Blue Book* and the Cambridge lectures of those years through to his final remarks in *On Certainty*, he himself, wary of too simply or too conventionally characterizing his own ways of working, appeals repeatedly, as we have now seen, to the notion of therapy. This is a fitting concept for a moment in philosophy's history of radical departure and an ensuing process of radical change.

It is true that, to return to the early work not seen since the beginning of this study in Chapter 1, in the *Tractatus* Wittgenstein wrote, 'Most of the propositions and questions of philosophers arise from our failure to understand the logic of our language' (4.003), and it would not be wrong to call attention to the continuities running through his thought from the earliest to the latest periods (indeed the deep and abiding interest in the self, the first person, and the nature and possibility of self-understanding is one such unbroken thread). But that should not obscure important differences as well; in the early philosophy he wants to articulate correctly the logic of our language such that illegitimate questions will not—indeed cannot—find a voice. In the later philosophy, the questions call not for theory, but for a therapy that

employs a multiplicity of voices within the work, where Wittgenstein enters into dialogue with the impulses of the traditional philosophical methods and formulations. (It does not follow from this difference that the early philosophy does not exhibit therapeutic aspects or house therapeutic elements—it does.) Fitting with the analogy to therapy is not only the dialogical dimension of this kind of work, but also the fact that the interlocutor's voice (or voices) is very often brought into the dialogue through an explicit expression of a more deeply motivated impulse (as we saw in Chapter 2) to voice a matter a certain way (e.g. 'I want to say . . .'; 'One here wants to say . . .'; 'One would like to say . . .'; etc.) than to present it as a reply or counterproposal in the context of a philosophical argument. And the sense that such impulses *are* deep, i.e. not merely surface-level, grammatically induced misapprehensions that can be quickly unseated—as quickly as they are identified—and replaced by correct formulations not under the influence of simplifying conceptual pictures or misleading analogies—also fits the notion of therapy: it takes a distinctive kind of *work*. Thus, Wittgenstein writes, in *Philosophical Investigations* §111, in the context of a set of remarks (§§89–133) on the nature of philosophy as profound as those found in Kant or Aristotle: 'The problems arising through a misinterpretation of our forms of language have the character of *depth*. They are deep disquietudes; their roots are as deep in us as the forms of our language and their significance is as great as the importance of our language.' It is vital to an understanding of Wittgenstein's point here, however, to go on to characterize more fully the nature of that sensed depth. But knowing how to go on at this juncture has not proven to be a simple matter.

It would, given a cohort of philosophical pictures concerning meaning, the mind, and the self, be easy to assume that this depth is beyond or beneath the reach of language, that it is perhaps pre-verbal, and that the depth of the disquietudes is such that any attempt to give them an articulate voice would, whether we are aware of it or not, constitute an attempt either to disguise or to transform their true nature. The trouble, we may too easily be inclined to think, is deep in the well from which language springs, and not in that language itself. This construal of sensed depth is of course common to our reflections on self-knowledge, where the image of interior containment includes a sector above, with content that is amenable to propositional articulation, and a sector below, which houses a repository of dark forces and impulses. And it is but a short step to a belief that the experience of sensed depth is caused by the stirring

of content behind or beneath the inner partition that is made to move, to vibrate noticeably, through the voicing of the propositional content, the linguistically capturable content, on the surface. Such deeper resonances, as we might call them in accordance with this picture of the self and its divided content, would thus be a measure—indeed a sounding—of the human significance of the utterance that conveys this sense of inarticulable depth.

The brute fact of human life—that there *are* such deep disquietudes—does not, however, necessitate this entire way of thinking, this way of seeing the issue. Wittgenstein follows his remark above with a brief but extraordinarily potent remark: '—let us ask ourselves: why do we feel a grammatical joke to be *deep*?' The potency is in the power to reorient: the sense of depth *need* not be characterized, or modeled, in the foregoing dualistic way *at all*. The sense of depth, to put it one way, may well be entirely *within* language, and the sense of depth may be caused by the wholly verbal, wholly articulable associations an utterance may awaken, or by the implications an utterance may hold that are, again, wholly within the realm of the sayable. That such associations and implications contingently may not, as they are sensed, be voiced is no argument that their content is necessarily unvoiceable. Indeed, autobiographical self-investigation, the investigation of the content of autobiographical consciousness, may be far better characterized precisely as—one here wants to say, if in a misleadingly reductive form—*grammatical* self-investigation, just as Wittgenstein characterized his new way of working on philosophical problems, his set of 'methods', as grammatical investigations. The depth of a grammatical joke is, in a telling way, *in* language, and Wittgenstein concludes §111 with the parenthetical punchline of this entire direction of thought: '(And that is where the depth of philosophy is.)' The sense in question has, as we say, employing the organic analogy that Wittgenstein uses, roots that run as deep as, not our pre-verbal subterranean inner content, but the *forms of our language*. And because language is profoundly important in human life to our self-constitution, our self-identity, we ought not to feel a sense of loss at recharacterizing the sense of depth—in philosophy generally or in the questions of self-knowledge more particularly—in the linguistic terms of grammatical investigation. For this *itself* is what is deep; it is not that we are, in changing our way of seeing this entire issue and leaving the dualistic model with its bifurcations behind, abandoning a genuine sense of humane depth. On the contrary, we are—or would be, were we to follow out all of Wittgenstein's extensive

therapeutic investigations on this topic—rather giving that sense of depth a stronger position. If we do feel a sense of loss at this suggestion, then that itself would call for its conceptual therapy in turn—where such a therapeutic process would take place not beneath or behind language, but rather, and *only*, within it.

In *Philosophical Investigations* §104 Wittgenstein writes: 'we predicate of the thing what lies in the method of representing it. Impressed by the possibility of a comparison, we think we are perceiving a state of affairs of the highest generality.' This exactingly describes the situation we find ourselves in when we represent ourselves to ourselves—where the mind forms an image of itself—as bifurcated entities with two distinct kinds of categories of content, with the propositional on the surface and the nonpropositional in the depths. We represent the self that way to ourselves—tellingly, *in language*—and then predicate of the self the properties and structure as given, not by it, but *to* it, by the picture. Impressed by the comparison between outward things (wells, depth-versus-surface geology, hydrodynamics) and inward senses (of depth in the present case) we too quickly—that is, in the very first formative stages of the inquiry, where again many of the expectations as to how that inquiry will proceed and what form an answer to it may take are laid down, often unwittingly, in those problem-formulations—believe ourselves to be articulating nothing short of the human condition, indeed a state of affairs of the highest generality.

If we sense loss at the unearthing and—after some philosophical work—removal of a 'vertical' sense of the self's autobiographical content, revealing to perspicuous view what lies before us as what we might call a more 'horizontal' one in language, that felt loss is only a measure of the strength of the grip of the picture of selfhood we would better do without, and not a measure of what is truly predicated of the thing—in this case, the self—itself. And our desire, our impulse, to speak in that 'vertical' way, as Wittgenstein's imaginary interlocutor so frequently does, is what is called, in *Philosophical Investigations* §254, 'raw material'. He writes: 'What we "are tempted to say" . . . is, of course, not philosophy; but it is raw material. Thus, for example, what a mathematician is inclined to say about the objectivity and reality of mathematical facts is not a philosophy of mathematics, but something for philosophical *treatment*.' Wittgenstein follows this with the methodologically dense single sentence of §255: 'The philosopher's treatment of a question is like the treatment of an illness.' And the analogy between philosophy and therapy is now again apt because the self-investigative

work required to unearth the only-indirectly manifested influences on our thought, such as misleading analogies, grammatical similarities, the falsifying and oversimplifying conceptual pictures that result from these—taken together, the deep—i.e. deep-in-language—sources of the impulses to speak in a metaphysical voice in accordance with the dictates of captivating pictures—*takes time*. Like (for all the lines of continuity) Wittgenstein's profound change in his way of seeing language from his early to late philosophy, in which a struggle (what he famously calls a battle against the bewitchment of our intelligence by means of language in *Philosophical Investigations* §109) for conceptual clarity is undertaken against his earlier self in his middle period, it is, again, not a matter of just recognizing an error and changing one's mind accordingly. Thus, in *Zettel* §382 he writes: 'In philosophizing we may not terminate a disease of thought. It must run its natural course, and *slow* cure is all important.'

It is significant that Wittgenstein includes the first-person pronoun among other philosophically troublesome words ('knowledge', 'being', 'object', 'properties', 'name'), recording our impulse to 'grasp the *essence* of the thing' in *Philosophical Investigations* §116. The kind of philosophical labor he is describing and recommending in these sections on the nature of philosophy, in so far as it requires a sustained vigilance to the picture-driven impulses, analogies, grammatical convergences, and so forth that we have discussed throughout this book, is indeed one kind of self-investigation, and this gives sense to his remark that working in philosophy becomes a way of working on oneself. Thus, to mimic the methods of science in philosophical work would be to fail to apprehend, or acknowledge, the distinctively *personal* nature of this kind of philosophical work; if the impulsions to speak in a metaphysical voice or a particular conceptual issue do not manifest themselves for a given individual, for that person those particular problems do not arise (while others certainly may). This points to the great gulf between the idea of a research program and this kind of therapeutic investigation. If we do grasp at the essence of selfhood—what we might there think of as the referent of the first-person pronoun—Wittgenstein suggests that we ask ourselves (beginning a process of therapeutic conceptual disentanglement) 'is the word ever actually used in this way in the language-game which is its original home?', and he follows this with what became the slogan of (the unfortunately named) ordinary-language philosophy: 'What *we* do is to bring words back from their metaphysical to their everyday use.'

We might then—one certainly feels an impulse to say this—assert that this recommendation only concerns words in their first-order 'homes', but that we are interested primarily in a second-order level, where the hidden reality behind the first-order usage of the first-person pronoun—indeed its essence—is revealed by the utterance of the metaphysical voice. We might say this of 'knowledge', 'propositions', and the other terms on Wittgenstein's list, along with many others, and we might, voicing opposition to Wittgenstein's 'horizontalizing' observations, say that philosophy itself has a second order, and *that* is what allows impersonal research programs and transcends this personalized–therapeutic model of conceptual work. Anticipating—well, no doubt, feeling—this impulse, Wittgenstein writes in *Philosophical Investigations* §121: 'One might think: if philosophy speaks of the use of the word "philosophy" there must be a second-order philosophy. But it is not so: it is, rather, like the case of orthography, which deals with the word "orthography" among others without then being second-order.' But one would find this convincing, and find the remark therapeutically viable, only after, or while in the process of, a changed way of seeing, in which what he calls the command of a clear view of the use of our words is genuinely valued and thus pursued with patience.

If, as Wittgenstein says in the following remark in §122, a main source of our misunderstandings is our lack of a command of a *clear view* of the use of our words, and if, as he says, this lack of perspicuity is precisely what generates and extends the linguistically misbegotten lives of our deep disquietudes, we naturally want to know how to characterize the perspicuous representation that is rendered elusive by misleading analogies and related bewitchments of language, or how, in short, to describe the clear view toward which our conceptually therapeutic labors aspire. Wittgenstein writes that what we might call the content of the perspicuous representation is a matter of 'seeing connexions' (of the kind we discussed in relation to Iris Murdoch and in connection with autobiographical memory); this itself, we should observe at once, is a way of describing the clarified view in profoundly *contextual* terms. The very concept of a connection requires that a given thing—a particular case or an example of a remark, a gesture, an utterance, an observation, and so forth—*not* be taken alone, that it not be considered in isolation as (so we might think of it) a narrowly circumscribed specimen for analysis. For any process of self-investigation, i.e. in autobiographical writing, the significance of this is plainly evident: any determinate episode or action (speaking far too generally here to convey any sense of

the kind of close attention to particulars required for the achievement of the clear view of which Wittgenstein is speaking) in one's life that is the present focus of autobiographical inquiry needs to be seen for what it is, where 'seeing it for what it is' runs precisely *counter* to the first picture of investigative scrutiny that may strike us on hearing this phrase. Seeing a slice of life for what it is requires—if we value fidelity to psychological life, to human *experience*—that we see it relationally, with links, interconnecting strands, indeed 'connexions', to related experience that in a multiplicity of ways informs and inflects it (as we also discussed in relation to Davidson). Thus, Wittgenstein added: 'Hence the importance of finding and inventing *intermediate cases*.' The assembly of cases and serious, sustained reflection upon them is the content of the therapeutic process that yields perspicuity; in the case of autobiographical writing, one can fairly readily or intuitively apprehend how *finding* what Wittgenstein calls intermediate cases might proceed. They would proceed just by looking at the episode or action in question, looking at it relationally, positioning it into a sequence, or indeed into a pattern that is emergent over a larger frame of time, and coming to see the connections that integrate it into the trajectory of a life, that—as we say tellingly if again too succinctly—make sense of it. Narrative threads are—in patterns—interwoven.

By contrast, *inventing cases*, as Wittgenstein also recommends, might seem initially unrelated to any serious autobiographical undertaking, since—if we want the truth—invention would seem the last thing we want. But this would be an error: the full contextualization of a case can benefit enormously from imaginary cases similar to, but still (perhaps only minutely) different from, the case in question. Determining, for example, whether a particular action, or pattern or sequence of actions, was genuinely altruistic, or selfish, or an amalgamation thereof, can proceed in just this way, that is, by imaginatively changing some of the morally relevant nuances of the case and asking ourselves how we would describe the case then. Such a process of invention in a sense *is* fictional, but should not be summarily dismissed for that: the light it throws on the adjacent, real case is factual, and the imaginary cases can greatly help to describe exactingly, to understand exactingly, the case in question to which they are its close neighbors. This is true of autobiographical investigations, just as it is true of philosophical investigations in a broader sense, and the therapeutic process of gaining a clear view is common to both. This parallel should not be underemphasized. Wittgenstein states its importance himself: he closes §122 with: 'The concept of

a perspicuous representation is of fundamental significance for us. It earmarks the form of account we give, the way we look at things.' And he adds, not in the errant voice of the inner interlocutor who recurrently expresses the philosophical recidivism of the nonperspicuous view, but in a speculative voice, one speaking more generally and outside the bounds of his focused discussion but in a way that casts light on his remarks, contextually situating them into a larger pattern, and making salient, through relational comparison, their fundamental imperative to bring about a change in our way of seeing, the parenthetical question '(Is this a ''*Weltanschauung*''?)'.

In the opening section of part II of *Philosophical Investigations*, Wittgenstein provides an example—or a hint of an example, if we are to go by the standards of elucidatory particularity he encourages in his sustained discussion of the nature of philosophy earlier in the book—of a human experience that is only comprehensible as an emergent pattern from a life (and not as a bounded, atomistic experience). '''Grief'' describes a pattern which recurs, with different variations, in the weave of our life. If a man's bodily expressions of sorrow and joy alternated, say with the ticking of a clock, here we should not have the characteristic formation of the pattern of sorrow or of the pattern of joy' (p. 174). We next see a contrast between the sentence 'For a second he felt violent pain' and the sentence of a very different kind but dangerously grammatically similar 'For a second he felt deep grief', with Wittgenstein asking why the latter does not sound at all right. The mistake here would be to accept unwittingly the analogy between the two kinds of experience because of the grammatical parallelism in the sentences, which would mislead us in turn to picture grief, like a sudden very sharp pain that lasts a second, as a sensation that is either 'on' or 'off'. The avoidance of, or the working one's way out of, such conflation is essential to commanding the clear view of which Wittgenstein spoke with special regard to any autobiographical investigation. In the grip of this picture encouraged by the conflation of the grammar of a sudden sharp pain with the far more conceptually involved human experience of grief, we would not only mistakenly characterize grief as an episodic sensation but then picture autobiographical true description as of a kind that bears a one-to-one correlation to such sensations—a correspondence theory of truth turned inward, as we saw such a view in action in previous chapters. And there is a question concerning grief that *looks* similar to a parallel question concerning the feeling of a sudden pain in the form 'But don't you feel grief *now*?' The intelligibility is derived *not* from

its analogy to 'But don't you feel that sharp pain *now*?' that the doctor asks while pressing the abdomen in an examination (that analogy would engender the *illusion* of intelligibility), but rather from the case in which someone asks us if we aren't feeling grief now that we rightly apprehend the grief-striking circumstances that earlier we misapprehended and uncomprehendingly replied that those circumstances caused us no grief. Of this question, Wittgenstein writes, 'The answer may be affirmative, but that does not make the concept of grief any more like the concept of a sensation,' adding that the question is of a personal kind, and not the misinduced 'logical' one wanting to ask about the presence or absence of the inner momentary sensation. Gaining clarity in such matters just is the process of philosophical investigation that is at one and the same time autobiographical clarification; it is how we achieve a perspicuous overview of the language of self-description, of autobiographical narration.

If (1) Wittgenstein's conception of philosophy is usefully elucidated (in large part—but to say this is not to say that his later work does not also display in various sectors of his investigations a more conventional philosophically argumentative *aspect*, if undertaken in the interest of therapeutic freedom and, again, radical methodological departure) in terms of conceptual therapy, if (2) the product of that therapeutic work is illuminatingly characterized in terms of the attainment of conceptual perspicuity, and if (3) as a result of the interaction of (1) and (2) we find ourselves in a position to grasp better the philosophical nature of an autobiographical undertaking, we might then reasonably ask for a still-deeper understanding of perspicuity. In his 'The *Investigations*' Everyday Aesthetics of Itself',[22] Stanley Cavell writes in pursuit of just such an understanding. And it is, I believe, for reasons that will emerge as we proceed, well worth returning to his work (last seen in Chapter 2) precisely at this point where the question of philosophical method—and particularly where that is linked to the question of autobiographical writing—has explicitly resurfaced.

That the task at hand, the gaining of perspicuity, is *in language* Cavell records in this way: having said that *Philosophical Investigations* provides its own self-description as work toward a perspicuous presentation, he describes this as 'an articulation of a task of writing'. And that work was described by Wittgenstein as leading words back from their metaphysical

---

[22] This piece is perhaps most conveniently found, and illuminatingly introduced, in *The Cavell Reader*, ed. Stephen Mulhall (Oxford: Blackwell, 1996).

to their everyday use. But if so, Cavell asks if the kind of perspicuity Wittgenstein is pursuing is the kind found in a formal proof. And if the answer to that is affirmative, is Wittgenstein's centrally significant methodological remark concerning perspicuity 'meant to signal an ideal of lucidity and conviction that he cannot literally expect in a work made of returns to ordinary words?'[23] This question, itself fundamental to a grasp of the methodology—not to suggest that this reduces to one unitary thing—of the later Wittgenstein, asking if the therapeutic result of perspicuity is equivalent to, or is modeled upon, the perspicuity of a proof, *is* fundamental precisely because, as Cavell makes clear, if the answer is affirmative, then Wittgenstein's most mature work should be understood as projecting 'arguments of formal rigor, even though its surface form of presentation does not, to say the least, spell them out'.

A model of perspicuity drawn from formal proof would indeed lead us then to go on to attempt to draw out and formalize, very possibly as positions against positions, as *counter*arguments in the mode of philosophizing conventional at least since medieval disputation and probably since Plato's dialogues, Wittgenstein's positions (and it is but a short step from here to the word *theory*) on word-meaning, naming, ostensive definition, family resemblances, linguistic privacy, rule-following, aspect-perception, speaker's intention, and the like. Cavell puts forward a *very* different approach, suggesting that 'Wittgenstein is claiming for the ordinary its own possibility of perspicuousness, as different from that of the mathematical as the experience of an interesting theorem is from the experience of an interesting sentence.' And we might add that the imposition of the model of formal-proof perspicuity onto the conception of contextualized, connection-seeing perspicuity that Wittgenstein is developing would indeed be to fall prey to a misleading analogy, to a misleading parallel on the surface of grammar, ironically while in pursuit of a kind of reflexive perspicuity, i.e. a perspicuous overview of perspicuity itself. Wittgenstein's remarks effect a change in our way of seeing, *very* unlike the theory-construction of more conventional modes of philosophical work; in doing so, they do not, as Cavell suggests, reduce to an exegetical extraction yielding formalized counterarguments. Cavell describes the experience of the sudden gain in conceptual perspicuity that can occur as the result of a view-transforming sentence as 'the movement from being lost to finding oneself'.[24] He adduces a number of such reorientations that can at once (1) give pleasure (for

---

[23] *The Cavell Reader*, 377; this and following quotations.   [24] Ibid. 380.

their literary power but not only that), (2) give a shock of conceptual freedom upon, I would add, having one's view of a problem, or an entire problem-field, transformed by suddenly seeing connections to a different range of examples, by feeling a freedom from the previously undiagnosed tyranny of misleading analogies and the expectations they enforce, by seeing through a particular bewitchment of language, *and* (3) give a sense of anxiety, since, as Cavell puts it, 'they treacherously invite false steps' of the reader, e.g. taking Wittgenstein's remarks about a word striking a note on the keyboard of imagination not as a vivid phrase that gives voice to a misleading philosophical picture but as a vivid encapsulation of Wittgenstein's own (not yet formalized) position. Cavell's chosen specimens include Wittgenstein's remarks about the aim of his philosophical work being to show the fly the way out of the fly-bottle, the turning of one's spade at reaching bedrock, the human body serving as the best picture of the human soul, the shutting of one's eyes in the face of doubt, and the marvelous image from *Philosophical Investigations* §107: 'We have got onto slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk.'

Tellingly, each of these is informative in connection with the elucidation of the kind of perspicuity in question here: (1) the fly-bottle analogy intimates a radically different way of seeing the problem before one; (2) the reaching of bedrock suggests an acknowledgement that what one has before one is *already* the bottom—nothing is hidden beneath it requiring deeper logical analysis; (3) the dualistic picture of inner-plus-outer selfhood is a picture that is not, against our initial philosophical intuitions, inevitably *given* as the human condition; (4) like the pragmatists, we should remember the option of shutting our eyes to *false* doubt (the doubt that, as Peirce said, if we do not first feel in our hearts we should not entertain in philosophy); and (5) we do not—against the presuppositions of Wittgenstein's own early thinking as well as those of many others before and since—*want* a logically perfect 'language' that formalizes in abstraction from the contextualized, grainy particularities that allow us to walk, i.e. to gain the special kind of perspicuity that comes from a knowledge of *cases*.[25] Cavell observes

[25] I explore the philosophical relevance of just such a turn to cases in literature in *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge* (Ithaca, NY: Cornell University Press, 1994), and in other arts in *Art as Language: Wittgenstein, Meaning, and Aesthetic Theory* (Ithaca, NY: Cornell University Press, 1995).

that a fully considered appreciation of such sentences, such philosophic gestures, marks a point of convergence between aesthetic and philosophic efforts, writing, 'Differences in the work philosophy does and the work that art does need not be slighted if it turns out that they cross paths, even to some extent share paths,' and he adds, '—for example, where they contest the ground on which the life of another is to be examined, call it the ground of therapy.'[26] This crossing of the paths of philosophical work and aesthetic appreciation or understanding takes place also with self-understanding, with autobiographical work in which the life of oneself is to be examined. And the achievement of perspicuity in any such reflexive endeavor will exemplify all of the characteristics of philosophical therapy that Wittgenstein, in his later philosophy, articulated—including the very circumstances that call for therapy, i.e. the myriad dangers of linguistically induced conflations that became confusions. The investigation of the self thus *is*, in Wittgenstein's special sense, a grammatical investigation; it is thus misdirected to claim that the close and exacting study of the language of self-description is 'merely semantic', and that the real study of the self is both logically and ontologically prior to any such verbal considerations. The analysis of what we might here call the 'textual self' is an investigation neither secondary to, nor prefatory to, the study of selfhood, and such an investigation shows another clear aspect of the meaning of Wittgenstein's remark that working on philosophy becomes a way of working on oneself.

In *Philosophical Investigations* §123 Wittgenstein writes, 'A philosophical problem has the form: "I don't know my way about (*Ich kenne mich nicht aus*)",' which Cavell translates with a different inflection as 'A philosophical problem has the form "I cannot find myself".'[27] Cavell calls attention to the kinship such a philosophical condition displays to Dante's loss at the beginning of what we might call his grand project of self-narration. But I would add that we should not forget that the distinctive loss Dante suffers, a destabilizing loss of way in a dark wood at the middle of life's journey, is somewhat ameliorated by the presence of Virgil. Our Virgil would thus of course be not personified, not Wittgenstein, but rather the collection of tools, the various

---

[26] *Cavell Reader*, 380.   [27] Ibid. 379.

methods, the ways of shaking free from the grip of conceptual pictures, in short, the extremely nuanced grammatical vigilance that would shield us from the disorienting influences of grammatically induced, linguistically induced, bewitchment.[28] These together can take the name 'philosophical therapy', and—to go too far—our Beatrice would thus be the state of complete grammatical clarity at which, as Wittgenstein puts it in *Philosophical Investigations* §133, our problems completely disappear.

The voicing of our disorientation in the phrase 'I cannot find myself' is also heard, Cavell reminds us, in Emerson's beginning of 'Experience', where he calls out: 'Where do we find ourselves?'[29] Strengthening the connections we are now positioned to see between philosophy and autobiography, Cavell writes:

It is, accordingly, as the philosophical answer to this disorientation that Wittgenstein proposes the idea of perspicuousness—outside the realm of proof, and by means of a return to what he calls the ordinary, or 'home' (I place the quotes to remind ourselves that we may never have been there). The section that names perspicuous presentation mentions 'intermediate cases,' hence suggests that the idea of understanding as 'seeing connections' is one of supplying language-games—as in the string of cases of 'reading' (§§156–178), or in comparing the grammar of the word 'knows' to that of 'can' or 'in a position to' and also to 'understands' (§150) or, more generally, in showing grammatical derivation, as of the grammar of 'meaning' in part from 'explaining the meaning,' or in showing grammatical difference, as between 'pointing to an object' and 'pointing to the color of an object.'

Such differences—which Wittgensteinian therapeutics would teach us how to teach ourselves—of philosophically relevant grammar we forget at our Dantean peril, and when such considerations converge on the topic of self-investigation, philosophy does, as we have seen, indeed become a way of working on oneself. Cavell sees the common point of departure in Wittgenstein, Dante, and Emerson as calling for a hard-won, genuinely earned perspicuity; he concludes the passage with: 'Perspicuous representation is accordingly the end of a philosophical problem that has *this* form of beginning.'

[28] For a microcosmic specimen of precisely this kind of therapy (by which, as Wittgenstein puts it in *Philosophical Investigations* §133, our problems 'completely disappear'), see *Philosophical Investigations* §154, which uncovers the ineliminable role of 'particular circumstances'. [29] *Cavell Reader*, 379.

Cavell's excursus on Wittgenstein's perspicuity positions us so that we can indeed see connections between autobiography and philosophy (and incidentally we should thus see Cavell's own recent turn to autobiographical writing[30] as a continuous *part of*, and not a second-order recounting of, his larger philosophical project). To see autobiography as philosophy is to see it as an ineliminable source of language-games of narrative self-description, and to see philosophy as autobiography is to see it, in turn, as the distinctive kind of self-analysis—the intricate, layered disentangling of the mind's grammatically fueled impulses to misspeak, to mischaracterize itself—that Wittgenstein's remarks on therapeutic philosophy articulate.[31] And to attempt to reduce these ideas to a succinct statement, or even to a *theory* of philosophical method, is of course to falsify one distinctive mode of conceptual engagement by forcing it into a mold—a picture—of a unitary methodological manifesto that, once forcefully stated, we would then follow formulaically. As Wittgenstein puts it in the closing sentence of his sustained inquiry into the nature of philosophy in *Philosophical Investigations* §133: 'There is not *a* method, though there are indeed methods, like different therapies.'

This study has been an attempt to bring together a number of those methods, and to follow out, and extend, the courses of a number of those therapies as they may contribute to the achievement of a perspicuous grasp of autobiographical, self-revelatory, and self-constitutive language.

---

[30] See *A Pitch of Philosophy: Autobiographical Exercises* (Cambridge, Mass.: Harvard University Press, 1994). The character of Cavell's philosophical project is well articulated by Arnold Davidson when he writes: 'Cavell writes not primarily to produce new theses or conclusions, nor to produce new arguments to old conclusions, but, as Kierkegaard and the later Wittgenstein did, to excavate and transform the reader's sensibility, to undo his self-mystifications and redirect his interest. This is a distinctive mode of philosophizing with its own special rigor, in which the accuracy of description bears an enormous weight. In aiming to transform a sensibility, one must capture it precisely, and if one's descriptions are too coarse, too rough or too smooth, they will hold no direct interest, seeming to have missed the mark completely. Cavell's writing places extraordinary pressure on itself to describe, undistractedly, and specifically, the forces of the mind' (quoted in Timothy Gould's fine study *Hearing Things: Voice and Method in the Writing of Stanley Cavell* (Chicago: University of Chicago Press, 1998), 24). Writing under a self-imposed pressure to describe undistractedly the forces of the mind, as a distinctive philosophical vocation, is, I am suggesting, at the same time and for the same reasons autobiographical work.

[31] See the foundational writings of John Wisdom in *Philosophy and Psycho-analysis* (Oxford: Blackwell, 1953), esp. 'Philosophy, Metaphysics, and Psycho-analysis', and *Paradox and Discovery* (Oxford: Blackwell, 1965), esp. 'A Feature of Wittgenstein's Technique'.

And we would do well, I think, to remember—to keep alive both in our philosophical reflections on autobiography as well as in our own autobiographical reflections—that it is perhaps only with such perspicuity that we may describe ourselves with a reflexive accuracy sufficient to meet the demands of what Socrates conceived as the life worth living.

This page intentionally left blank

Index