

Philosophy of mind and/as the repression of interpersonal understanding

Joel Backström

[Published in: In: J. Backström, H. Nykänen, N. Toivakainen & T. Wallgren (eds.), *Moral Foundations of Philosophy of Mind*. Cham: Palgrave Macmillan, 2019 (pp. 231–266)]

Modern and contemporary philosophy of mind is thoroughly determined by a subject-object problematic. The basic question is taken to be, in different ways, how the ‘I’ with its ‘subjective point of view’ relates to a world of objects describable in impersonal terms, where these objects include the body and brain in which the subject is somehow supposed to reside, or from which it is supposed, perhaps epiphenomenally, to arise, as well as the bodies which somehow house other subjects – although on this view, the very (or the ‘possible’) existence of ‘other minds’ becomes a ‘problem’ for the subject.

This problematic has been with us since Descartes, and although materialism has long since replaced Cartesian substance-dualism as the standard metaphysical view, contemporary materialists accept Descartes’ basic framing of the issue, the radical splitting of our ‘minds’ from our mechanically-materialistically (mis)represented ‘bodies’. They then deny the existence of the other half of the split, the ‘immaterial’ mind – supposedly, we’re all ‘*made of mindless robots and nothing else*’ (Dennett 2003, p. 2) – but this metaphysical denial leaves the logic of the situation intact. The ‘body’ and its ‘behaviour’ are still conceived as merely physical, i.e. as ‘*not-mind*’, and the big philosophical question remains relating this physical brain/body to apparently ‘first-personal’, ‘subjective’ experience, the very possibility of which now seems profoundly mysterious. The widespread idea that we need a ‘Theory of Mind’ (ToM) to understand others also presupposes the Cartesian set-up insofar as the ToM is supposed to enable an inference from visible behaviour to its invisible mental causes (invisible like Descartes’ immaterial mind).

The aim of this chapter is to expose and discuss the basic confusion that Descartes shares with contemporary materialism, namely, the framing of the issue in subject-object terms. Irrespective of the details of the picture one paints within it, and whether one admits this or not, that frame itself excludes *the relationship between human beings*. That relationship is inconceivable in subject-object terms but is, as I hope to show, the very centre

of human intelligibility and inseparable from our relationship to ourselves. ‘I’ exist only in relation to ‘you’. When this ‘you’ is ignored, the ‘I’ turns into a kind of spectre, reappearing as that fantasised ‘entity’ called ‘the subject’ – a notion strictly correlative to that of ‘the object’ that this subject thinks or experiences – whose reality one is inclined alternatively, and fruitlessly, to insist on and to deny, as illustrated by the interminable oscillations, throughout the history of modern philosophy, between various versions of dualism/subjectivism/idealism and monism/objectivism/materialism.¹

The structure of the chapter is as follows. Section One elaborates on what I mean by the ‘I-you-dimension’ of human experience. Section Two brings out (some aspects of) the irremediable incoherence involved in the inferentialist picture of understanding others (e.g., ToM-accounts). If one ignores or denies the dimension of unmediated I-you-understanding, one has to assume inferentialism in some form; hence, showing the untenability of inferentialism shows the impossibility of lucidly – actually, on the level of understanding rather than official pronouncement – denying the I-you-dimension. Section Three sketches how our pervasive difficulties of understanding may be explained, given my stress on unmediated interpersonal understanding. I also argue that the idea that ‘the context’ determines the meaning of expressions shares with inferentialism a misguided objectification of ‘the expression’, and I explicate the essentially *engaged* nature of understanding. Generally speaking, apparent failures to understand others should be seen as forms of *repressing* one’s own engagement and understanding, rather than as mere absences of understanding. Section Four elaborates, mainly through a discussion of our relationship to pain and suffering, how ‘engaged’ means *morally* engaged, i.e., how there can be no satisfactory morally neutral account of interpersonal understanding. Section Five analyses further how lack of understanding and a certain, always ambivalent, split of body from mind and of one mind from another is produced in everyday life through various strategies of repression, and how repression is basically a shared, social ‘project’. The concluding section draws the threads together by indicating how the confused subject-object-perspective in philosophy arises from and mirrors the moral-existential confusion in everyday life created through repression.

1. The openness between ‘I’ and ‘you’

The first ‘object’ infants focus on is no object at all, but the human face of the persons who care for them, addresses and is addressed by them, and with whom they will later explore the world of objects.² This relationship between human beings can be conceived as neither

‘objective’ nor ‘subjective’, nor as some combination of both.³ This is strikingly illustrated by the experience of *meeting someone’s eyes*; no less strikingly philosophers have almost universally overlooked this most basic human experience.⁴ As Wittgenstein pointed out, the eye that sees isn’t part of its own visual field (1951, 5.633–5.6331), but *your* gaze isn’t part of my visual field any more than is my own. That is, when our eyes meet, I’m not looking *at* your eyes, or *at* you; I might not even notice the colour of your eyes or anything in particular about your appearance. I look into your eyes, but not as one might look into a drawer to see what’s inside; my eyes find no object, they find you. To meet someone’s eyes is to be aware of the other person, to be in contact with her. One might say that here the eyes function not as visual instruments, but rather like organs of touch. Think of the arresting experience of unexpectedly meeting a stranger’s eyes; of how the distance between you instantly vanishes as your eyes meet.

When our eyes meet, my eyes find you, and at the same time I’m found by you; while I can look at you, observe you, without your noticing, I can look into your eyes only if you look into mine. This experience isn’t, then, one that might, contingently, be shared. Rather, it *is* the experience of sharing, not this or that, but, one might say, of sharing life with the other, being there with her. In other words, this experience isn’t properly *subjective*, since I cannot subjectively have it without your also having it, being with me in it. In this sense, it cannot be subsumed in, and so breaks the apparently unbreakable hold of, the standard philosophical idea of ‘experience’, according to which it is something a subject has, by itself, and whether another subject has the *same* experience (in the sense that both of them saw the grey cat crossing the street, say) must be determined through their reporting on and somehow comparing their subjective experiences – with interminable philosophical debates and existential doubt ensuing about whether two people can ever ‘really’ have the same experience. When our eyes meet, however, we’re in the same experience, we’re *in* the meeting. As one might say, there’s here nothing for one to do/experience unless the other is doing/experiencing it too. But the meeting of eyes is indeed an *experience*; it isn’t an objectively determined event any more than it is anything subjective. Our eyes cannot meet without both of us knowing; without our partaking in the experience of our eyes meeting there’s no event there for anyone else to register.

The experience of meeting someone’s eyes is, simply, an experience of openness, our openness one to the other. And the point isn’t that meeting someone’s eyes is a strange anomaly, the one experience that’s radically different from all others. On the contrary, it is

only a particularly striking illustration of that pervasive and most basic dimension of experience in which we address and are addressed by each other; a dimension constituted through and as this addressing-and-being-addressed. The distinction I illustrated in terms of what the eyes can do (looking *at* another vs. meeting their eyes) can be formulated in any number of ways in all ‘sensuous’ registers, including the tactile. Hence, the difference between using one’s hand to investigate, perhaps for medical purposes, the form of someone’s skull, and caressing them or being touched by their caress – or that between hearing a noise and hearing the address in someone’s voice, as when they call one’s name.

Our openness to the other’s touch and address is, one might say, the very life of our life; an object is inanimate insofar as it exists beyond any possible address, and death is the moment when the once living other ceases permanently to be reachable by our address. But if this openness to the other’s address is inescapably part, and indeed the very heart, of being alive, it is also the point around which our difficulties crystallize. The range of reactions that being addressed may occasion – embarrassment, shame, bewilderment, etc. – and the felt impossibility, in ‘normal’ circumstances, of looking into another’s eyes, of keeping contact, for more than a few moments, illustrate the existential-emotional charge and the enormous difficulty of the encounter with the other. Our encounters are pervasively deformed by all kinds of wishful and fearful manipulation of self and other, our life lived in the tension between a longing for an open encounter and fear of it, where the fear, too, manifests the very receptivity and openness to the other’s address that one fears. And this means, as I’ll try to bring out, that our understanding of each other and ourselves is, from the start and unavoidably, a morally charged affair where truthfulness is the task and neutrality impossible.⁵

Philosophers try, but cannot in fact quite manage to ignore this dimension of openness: the ghost of the other returns, just as does the ‘I’, when one tries to exorcise it. Thus, Dennett wants to purge ‘the subject’ from philosophy; “A good theory of consciousness *should*”, he says, “make a conscious mind look like an abandoned factory (recall Leibniz’s mill), full of humming machinery and nobody home to supervise it, or enjoy it, or witness it” (2006, p. 70). But a mere page earlier he spoke, *en passant*, of how an almost paralysed Parkinson’s patient’s minimal, and therefore particularly striking, response to the jokes his visitors tell him – “the corners of his mouth turn up in an involuntary smile and [there is] a little crinkling at the edges of the eyes” – “bring home vividly that there is still somebody at home in there, listening attentively to whatever you are saying” (2006, p. 68). Officially, however, *in theory*,

we're told that there's really 'nobody home', so how are we to understand Dennett? It might be said that these are different levels of discourse, different senses of 'the subject' – Dennett himself says "I don't maintain, of course, that human consciousness doesn't exist; I maintain that it is not what people often think it is" (2006, p. 71) – but the point is that the apparent interest and attraction, the shock value, of Dennett's claims hangs on his systematically conflating them. And the important thing is that what stubbornly refuses materialist reductions are not some supposed 'qualia' inaccessible to anyone but myself (a muddled notion rightly criticised by Dennett), but rather the striking presence of another human being, to whom I cannot but respond.

2. Understanding others: the inconceivability of inferentialism

In the world as imagined by Descartes and his contemporary materialist heirs, the living (human or animal) other is supposedly not directly present or 'accessible' to me, or I to them. There is only *one* subject, and the objects it observes. 'I' only perceive 'material bodies', and have to 'infer' the 'inner life', the 'states of mind' of others from the behaviour of their bodies conceived, in C. L. Hull's (1943) phrase, as the 'colourless' movement of mere matter. Different theories in psychology and philosophy of mind disagree over whether such inferences need or can ever justifiably be made, and if so, in what exact form. Behaviourists and eliminative materialists apparently deny this need altogether, while proponents of 'folk psychology' and 'Theory of Mind' think we need 'mindreading' or 'mentalising' in order to make sense of, explain and predict what would otherwise remain random movements. Again, the details of the various proposals for how this is supposedly done vary, but they all implicitly or explicitly assume – whatever this is really supposed to mean – that 'the mental states of others ... are completely hidden from the senses [and so] can only ever be inferred' (Leslie 1987, 139), so that, at bottom, '[the idea that] there are minds or souls attached to some bodies is simply a hypothesis, introduced to account for certain observed facts' (MacIver 1964, p. 308). Human beings are 'just complex objects in our environment whose behaviour we wish to anticipate but whose causal innards we cannot perceive' (Heal 1995, p. 11), understanding them being 'no different, in principle, from ... understanding the behaviour of other, more inert, objects' (Stone and Davies 1996, pp. 126–7).⁶

The problem with this (post-)Cartesian idea of how we 'read' others is that there's no such thing as 'mere, colourless behaviour' from which the mental life of 'other minds' – or,

indeed, any kind of life – could or would need to be inferred. That is: if we try to conceive such an inference, we end up not with an intelligible idea but with nonsense. The mythical element in the myth of the ‘Ghost in the Machine’ isn’t just the ghost, but equally the machine, and as Ryle himself notes (1949, pp. 20–25, 309–11), historically it was the machine-myth’s obvious inability to account for human action that gave rise to the myth of the ghost. What we actually perceive in our dealings with each other isn’t ‘colourless’ behaviour, but behaviour in the very different sense of a living human being’s meaning- and feeling-laden responses. From behaviour in this latter, full sense, life cannot, but also needn’t be *inferred* because it is directly manifested and seen in it, that is, sensed, and in being sensed understood, in various sense-modalities, whether felt in the other’s movement or touch, heard in her words or sighs, or seen in her demeanour or her eyes. But dualism and its materialist progeny are blind to this; they suffer from a fundamental ‘life-blindness’ (Midgley 2014, p. 143).

In abstract terms, the point against the idea of ‘colourless’ behaviour is simple: if one doesn’t perceive the life in the behaviour one also cannot infer it, for what is lifeless remains lifeless – and if one *does* see the life, one need not and indeed cannot *infer* anything either.⁷ Tears, for instance, aren’t just ‘water running from the eyes’, from seeing which I then infer that you are sad. If *that* – ‘water running from the eyes’ – is what I see, I might indeed make an inference or conjecture as to the cause: perhaps what made your eyes water is that you’ve been chopping onions or just came in from the heavy wind outside. What I’ll not infer is that you’re sad, because if you were sad you would be *crying*, which is wholly different from ‘having your eyes water’. To see tears *as tears* is to see them, not as mere water, but as an expression of sadness – or, say, of anger, happiness, humiliation or despair – and this is seeing the emotion itself, that is, the other person’s state of feeling and understanding (or bewilderment) that is expressed in the tears. But once this is seen, there’s nothing left to infer. In other words, tears don’t relate to sadness as an effect relates to the cause it allows one to infer, even if I might (perfectly correctly) say that you cry because you’re sad. Whereas one can perceive an effect without perceiving its cause, if one doesn’t see the emotion *in* the tears, that is, if one doesn’t see tears but ‘eyes watering’, one will, to repeat, *have nothing* from which to infer anything emotional at all. Putting the point in the form of a *reductio*: If one needed inferences, there’d be no room for inferences.

And note that since the supposed ‘leap from observable behaviour to unobservable mental states’ is strictly *inconceivable*, it is no use saying that the leap is ‘so common and

routine that people often seem unaware that they are making a leap' (Epley and Waytz 2009, p. 499). If the supposed move simply cannot be made (or even tried), then *a fortiori* it cannot be made routinely or extremely rapidly through unconscious brain-processes, or again by evolution working very slowly over millions of years. This simple point should be obvious, but a purportedly 'scientific' prejudice makes it seem unacceptable. The objection goes something like this: 'You're saying that we *simply understand* others' expressions; that their meaning *is just there*. But that's mysticism, and we want science; explanations of how complex phenomena/perceptions are built out of simpler elements.' I have nothing against explanations, however; my point is that nothing will ever *be* explained if one tries to construct meaning out of 'colourless behaviour'. That's like trying to find the *news* in a newspaper by chemical analyses of the ink and paper, and can only produce endless mystifications and pseudo-explanations.⁸

The absurdity of the standard (post-)Cartesian picture can be seen from another side when one considers that the 'inner mental state' one would, on that picture, postulate or infer to make sense of the other's otherwise supposedly meaningless 'outer behaviour', can itself only be conceived in terms of the unmediated understanding of the other that the picture denies. Thus, if I try to imagine the sad person (or 'mind'?) whose presence 'inside that head' I'm supposedly lead to infer from seeing those watering eyes, how am I to imagine her ('it')? The only way I could even seem to do this is by imagining the 'person inside' being *sad* in the ordinary sense, i.e., expressing her sadness in various ways, and so the unmediatedly understood meaning I banished, because of theoretical commitments, 'without', returns, surreptitiously, 'within'. To be sure, I might imagine someone having sad *thoughts*. But what makes a thought sad? Merely saying to oneself 'I'm sad' or 'That scene made me sad' doesn't, as such, necessarily express a sad thought. It all depends on *how*, in what *tone* or *spirit*, these words are spoken. If they're spoken absent-mindedly, bitterly, ironically or angrily, they don't express sadness (a sad thought), but rather reveal one's bitterness or anger, or they may leave it unclear *what* one's words express. Similarly, if various pictures – say pictures involving your now absent friend – are in your mind, which emotions and meanings these pictures relate to or express depends on the spirit in which you engage with them. In short, it won't do to pretend that 'colourless' words and other behaviour could get their meaning and emotional significance from some 'inner state' essentially unconnected to the responses of a human being manifested in glance, demeanour, tone of voice etc., for, to repeat, if one pretends not to see the meaning and expressiveness of 'outer' behaviour, one

will in effect have to smuggle it back into the ‘inner’ mental states supposedly inferred from it.

This is the (or one) upshot of Wittgenstein’s discussion (1958, §243 ff.) of the fantasy of a private language, a ‘language’ of inner experience that would have no connection to one’s life with others, at least no connections that would be understandable to others, even in principle. Wittgenstein’s point isn’t that we know what a ‘private language’ – in contemporary jargon: ‘qualia’ – would be, only there cannot (logically) be one, but that *nothing* has really been imagined, conceived under that title.⁹ Correlatively, ‘zombies’, beings indistinguishable from us except for lacking conscious experience, seem conceivable to contemporary philosophers of mind only because what they conceive under the title ‘conscious experience’ (or ‘qualia’) is really *nothing*, so everything indeed seems to stay the same even if one removes ‘that’. But as I have tried to show, what one conceives of as the ‘everything else’ that would be left after one has ‘removed’ this nothing – i.e., all of human behaviour, only ‘without inner experience’, as mere ‘colourless behaviour’ – is also, in fact, a nothing. When one attempts to split ‘inner’ from ‘outer’ in this way, one doesn’t get a split reality, one gets nothing really conceivable at all.

Instead of hopelessly trying to imagine body and soul as essentially distinct, we might say, with Wittgenstein, that ‘if one sees the behaviour of a living being, one sees its soul’ (1958, §357, transl. modified); or, again, ‘... if the play of *expression* develops, then indeed [but not before] I can say that a soul, something *inner*, is developing’ (1982, §947). It is often possible – in some cases, for some time, with some people – to hide one’s thoughts and feelings by controlling or feigning expressions, but the very fact that this is possible only through special effort, that one has to stifle the expressions (e.g., the sob, the smile or the interjection) that ‘want out’, or set oneself to produce the semblance of a genuine expression (e.g., forcing a smile), reinforces the point that we are fundamentally and ineluctably expressive beings, beings who are open to, stand in relationship to each other. This isn’t ‘logical behaviourism’, a would-be ‘position’ that presupposes the inconceivable split into ‘mere behaviour’ and ‘something else’ in its claim that the former is the criterion of ascribing the latter (e.g., a pain) to a person, or that ascribing the latter amounts to nothing more than ascribing the former. Feelings or thoughts cannot be identified with, or reduced to, particular behavioural manifestations; we express them in indefinitely many different ways. If I can express a thought in only one way, stubbornly repeating the same formulation, unable to put it differently to explain and elaborate on what I mean, it becomes unclear whether I really

express a thought at all, as opposed to repeating a formula I don't understand. Similarly, my sadness or happiness will come to expression not just in tears or laughs, but in demeanour, behaviour etc., suffusing everything I do.

Note that, while my discussion shows the incoherence of radical social constructivist ideas of 'emotional expressions' as merely culturally developed interpretations or labels put on behaviour which would be 'colourless' in itself, it doesn't therefore support competing theories of a fixed number of basic emotions with characteristic facial expressions hard-wired by evolution.¹⁰ I take no position on any empirical hypothesis. My point concerns the basic character of interpersonal understanding; an understanding that is presupposed in everything we do, including in our scientific work and empirical hypothesising, hence cannot itself be tested empirically. Note, further, that so-called simulation-accounts of how we understand others (e.g., Goldman 2006) share the same untenable starting-point as other Theory of Mind-proposals, for they accept that we first register a piece of behaviour that in itself means nothing to us, and that becomes endowed with meaning only as we somehow (perhaps unconsciously) 'simulate' being in the same situation as the other, and thus find out what that would be (is) like. But unless we already, directly, understood the other, we wouldn't know which 'situation' she was in, or that she was in any situation at all, or even that there was someone else there in the first place.

3. Misunderstandings, the idea of 'context', and the engaged nature of understanding

It might seem that what I've said about the immediacy of interpersonal understanding is contradicted by the fact that we often read – project – into the expressions of others meanings that weren't actually expressed by the other person, just as we often miss or ignore and repress awareness of what *was* expressed. However, such cases don't contradict what I have said, because they cannot (nor can cases of simple mistakes or errors in perception) be understood as a matter of one's seeing a piece of 'colourless behaviour' on which one then imposes the wrong interpretation. When, say, a distrustful person sees hostility in what is actually a friendly gesture, hears the other person as laughing *at* her rather than with her, etc., there's no neutrally describable perception of the other's gesture that she would share with someone who isn't distrustful and so sees or hears the friendliness that's actually being expressed, with the difference only arising through different interpretations being put on what

is perceived. Rather, her own distrust colours and distorts her experience of the other from the very start. The possibility of different, more or less false or truthful, perceptions of the meaning of the same expression doesn't vindicate the unworkable idea of 'colourless behaviour', then. Rather, it alerts us to how understanding the expressions of another is a process in which one is oneself inescapably *engaged*, and where what one 'can' and 'cannot' see – i.e., what one can allow oneself to see – depends on the character of one's engagement with the other person; on how far fearful and wishful fantasies distort it, etc.

The unmediated, basic understanding between people obviously doesn't prevent pervasive misunderstandings and confusions in our relations. The question is how these constant breakdowns and blockages of communication are to be understood, and the crucial point I'll try to elucidate is that they're basically symptoms of moral-existential difficulties, of our fearfully refusing, that is, repressing our understanding of each other and of ourselves. We don't allow ourselves to become clearly aware of what's going on between us; we block our understanding before it can unfold into full awareness and articulation. This isn't to deny that there are also genuine, innocent, misunderstandings between people where no refusal to understand is involved (e.g., one may see the other's anger but have misunderstood what she is angry about). But they aren't the main thing, and they have the destructive consequences for our relationships that they sometimes have because they're intertwined with refusal and repression – where the point is that refusal of understanding is destructive because it means refusing open relations with the other person who one refuses to understand. For example, in one's proneness to suspicion and resentment one instinctively takes even innocent misunderstandings as proof of the other's ill-will or indifference, and so one thinks, say, that the other wasn't listening when she simply couldn't hear what one said.

It may seem strange that I haven't mentioned the importance of *context* – a favourite philosopher's term – in understanding human expressions. Indeed, in speaking of an unmediated understanding of others I may seem to be denying it, and such obvious facts as, for example, how differently someone's smile may strike one when one realises *what* they're smiling at, which might be a happily playing child, or in another case, a suffering enemy (the example is from Wittgenstein 1958, §539). But I don't deny such facts; the question is only how they're to be understood. It seems to me that the need to insist on giving our expressions a 'context' arises only when one has already, without realising it, reified, one might say fetishised, 'the expression' – as though what was at stake were understanding the meaning of, say, 'the smile' or 'the tears', conceived as some kind of quasi-objects, whereas what one

understands (or fails or refuses to understand, if there is trouble) is really *the person* smiling or crying. And once one has, in one's theorising, turned the expression into a strange 'something in itself' by isolating it from one's understanding of the person who expressed something within a relationship, this abstraction indeed cannot be understood 'in itself', and so one appears to formulate a crucial insight by insisting on the need for 'context' around this reified expressive 'object' to 'determine' its meaning. For example, if one defines 'the smile itself' as what is visible in a photograph taken of just the smiler's mouth, one may well be unable to tell the smile of the person smiling down at the child from that of the person looking down at the suffering enemy. But such artificial conceptions of 'looking /being the same' are irrelevant for interpersonal understanding, which is precisely a concerned understanding of what goes on between people, of how we relate to each other. We smile at *each other*, and it is here, within our relationships, that we understand the differences between different smiles, the friendliness and the destructiveness they can manifest. And here, the cruel smile at a suffering enemy is as different as can be from the heartfelt smile at a playing child.

Now, in reifying 'the expression' in the way described, depriving it of its life and meaning by tearing it out of the living, meaningful interpersonal relationship where it belongs, one has actually rendered the idea of the 'context' lifeless and meaningless, too, and one will be unable to 'generate' meaning and understanding by putting 'expression' and 'context' together – just as one doesn't get a living being, or anything intelligible at all, by putting a Cartesian 'body' and 'mind' together. How, for example, is one to describe the 'context' that is supposedly added to make the smile intelligible? The mere fact that the person is smiling at a happily playing child doesn't yet determine what the smile expresses, for one may smile at children in myriad problematic ways, too: sentimentally or condescendingly, say, or indeed cruelly. I'm not saying that one can see precisely what is expressed in a smile by just intently looking into the smiler's face, that it is 'all *there*'. On the contrary, I'm criticising such reifications of 'the expression'. I'm saying that the meaning of the smile cannot be found by looking somewhere else, either, into a 'context' supposedly distinguished from 'the smile itself'. Rather, one understands the meaning of the smile insofar as one understands the smiler in her relationship to the one she is smiling at, and to other people. And this means, crucially, that one must understand *oneself* aright in one's relation to these people, for insofar as one's attitude to them is deformed by self-deceptive and destructive attitudes such as sentimentality or cruelty, this deforms one's understanding, too. Thus, if one is prone to sentimentality, one sees even what is actually a heartfelt, loving smile in a sentimental light; in a light, that is,

which presents love and human beings in an idealised, ‘prettified’, and at the same time unconsciously cynical way, as in ‘Oh, children are so *cute*; how beautiful and sad that paradise can’t last and they’ll soon be grown up...’ (*i.e.*, that they’ll soon be foul adults like oneself).

What understanding another may in practice involve cannot be determined in advance; it is as open-ended as our relationships. If you enter the room with a smile on your face, I might ask you why you’re smiling. Your smile tells me you’re happy, but not why you are, and your answer (‘I just heard John is visiting’) might be said to provide ‘the context’, or ‘the cause’, of your smile. But it was only because I saw your happiness, *i.e.*, because I already understood you, that I felt a need to ask you for any ‘context’, any explanation at all. And how or whether I understand your answer depends on my understanding of you; for example, given my sense of your relationship to John, I may wonder how you could be happy that *he* is coming, or there may be something in your voice that makes me feel you’re not being sincere, etc. There are endless variations. But whatever they are, rather than saying that ‘the context’ gives meaning to our expressions, one might say that our way of relating to each other, manifest in what we do and express, *is* ‘the context’. But why would one talk of ‘context’ here at all?

I said that one cannot determine in advance what understanding another may involve. But in an important sense this cannot be *determined* at all, insofar as one cannot really say *what* understanding another person involves, or consists in – and not because understanding is somehow vague or uncertain, but because it isn’t built up of discreet factors that could be enumerated and described. What I tried to show in the previous section wasn’t merely that some particular version of inferentialism fails, but that understanding others isn’t basically (but at most in special and marginal cases) a matter of making inferences or drawing conclusions from discreet pieces of evidence of any kind. How do I know that you’re sad? From seeing your tears? Well, one might say that, but then how do I know that those are tears and not just runny eyes, and furthermore tears of sadness, and not of joy, for example (for people do cry from joy, too)? There’s nothing one can point to as ‘the things from which I conclude that you’re sad’. That is, whatever one points to, one can again ask why I see *that* as showing your sadness. Whatever may be proposed as a certain, independently ascertainable ‘sign’ of sadness is bound to have exceptions in both directions – the ‘sign’ may be there, but no sadness, and there may be sadness, but no ‘sign’ – and furthermore, even if there were a sign that we found in fact accompanied all and only sadness, how would we know that these

were all cases of *sadness*? Obviously, the sign didn't tell us that, for, *ex hypothesi*, the presence of the sign is something that we discovered, contingently, to accompany the sadness. Only our *understanding* of sadness, that is, of *people*, others and ourselves, in our sadness and joy, would allow us to say that these were indeed cases of sadness, and this understanding itself isn't based on detecting any sign, but manifests our unmediated interpersonal understanding, which doesn't really *consist* in anything (enumerable signs, elements, etc.).¹¹

It is crucial to see that the root confusion I'm trying to diagnose in this chapter isn't a mistaken conception of *how* we 'gain knowledge' of the other's feelings, thoughts and intentions, where the mistake would be to think we infer them from outer signs rather than immediately understanding people's meaningful expressions. The real confusion is the notion that our relations to others are, most basically, *epistemic*, i.e., relations in which a 'subject' gains knowledge of, or believes or predicts something about, an 'object' (where that 'object' may be conceived of, say, as 'the expression' or as 'the mental states' of another person). While I have spoken freely about understanding, in this context 'understanding' isn't basically an epistemic notion, but rather itself an aspect of our *caring* for each other. When I speak of our openness to each other, I mean this engaged, caring understanding – which, importantly, is typically more or less troubled and repressively deformed. It is this caring that *gives us anything to understand* in the first place. It isn't that we notice an expression on someone's face, and then start caring. Rather, we notice particular expressions because we always-already and inescapably care for, and in this caring understand, each other. Perceiving a heartfelt smile, for example, doesn't mean merely registering, knowing, that the other person smiled; it means, most basically, *smiling back* – not because one has learned to do so in response to facial contortions of a particular configuration, but because the other's smile warms one, moves one to a smile in response. The response is spontaneous, but not blind or automatic; it is a matter of opening oneself, opening one's own heart, in response to the welcoming openness the other manifests. This being-open-and-moved-to-respond isn't something that might follow as an independent reaction on the perception of the other's smile, but is itself the basic form of perceiving heartfelt smiles, that is, of understanding the person who smiles and oneself in relation to them. In other words, the heartfelt 'quality' of a smile – what distinguishes a heartfelt smile from a mere grin, and from the myriad forms of ambivalent, more or less manufactured, strained and/or hostile smiles – isn't really a quality of the smile considered as a quasi-object, but something that is there, when it is, between the

smiler and the one she is smilingly turning to, just as the meeting of eyes is something that happens between two people; a mode of their being in contact, in relation.

This is, of course, also true of all the smiles that aren't heartfelt but polite, sardonic, disdainful, ironic, melancholy, etc. They, too express the smiler's way of relating to the person(s) he is smiling at. The smile is itself a manifestation of his way of relating to them, and so the irony or politeness, say, isn't some mere 'quality' of the smile itself, and to perceive irony or politeness basically means *responding* to the person who is ironic or polite, in one of any number of ways in which one may respond to such attitudes; for example, the other's politeness may provoke an impulse to be rude, or one may feel thankful that the other had the courtesy to treat one politely, even after one was rude to them. What distinguishes the heartfelt smile from the other kinds of smile is that it manifests the smiler's wholehearted opening of himself to the one he smiles at, a longing to abolish every reserve and distance between self and other, whereas in the other cases, there is, consciously or unconsciously, an element of distancing involved, of taking up a *stance vis-à-vis* the other ('I'm over here, and take this attitude towards you, over there'). Such distancing manifests our fearful attempts to defend against our openness to each other, against our own caring, against the way it brings us into contact with the other and with ourselves through the responses the other awakens in us, which we pervasively feel are 'too much to take'. Just think of the often desperate efforts people put into preventing themselves from crying, or, if they couldn't stop a sob from breaking out, into regaining their composure as soon as possible, steeling themselves against their own emotional response to another.¹²

As I will presently try to explain more fully, this whole dynamics between human beings, with its understanding and openness, and its pervasive distancing and urges to close oneself to and *not* understand the other, is essentially a moral matter, i.e., it cannot be understood in morally neutral terms. Or as I put it above: interpersonal understanding is essentially engaged. The central task of this chapter is to make clearer what this means, and how the engaged nature of understanding *renders our philosophical accounts of it morally charged, too*. As philosophers or scientists we don't speak from some morally neutral position 'outside' of the life we wish to account for; rather, we speak from within our life with others, we address them, even if we pretend otherwise, and our morally determined, wishful and fearful difficulties with opening ourselves to and understanding others (our evasiveness, suspiciousness, eagerness to please, etc.) will form and deform our supposedly 'disinterested' accounts of life, too. – As I noted above, there is a diverse literature that is critical, as I am, of

the inferential picture of interpersonal understanding and, more generally, of the neglect of the ‘intersubjective’ or ‘second person’ perspective in standard philosophy of mind.¹³ What I find unsatisfactory in most of this literature is precisely that, notwithstanding valuable individual points and arguments, where one could find close parallels with many points raised above, the critiques tend to overlook, or at best downplay the importance of the moral dimension of interpersonal understanding. To put it more strongly and precisely, they treat it (insofar as they take note of it) as merely a ‘dimension’ of the human reality studied, whereas my central point is that interpersonal understanding *is* essentially moral understanding, and that this means that philosophical accounts of it, too, are drawn into its morally charged field.

4. Compassionate torments: the relation to the other and the moral life of the soul

Our understanding of each other is morally charged from the very start precisely insofar as it is indeed, or exists within, a relationship of caring between people, between ‘I’ and ‘you’. Between ‘subject’ and ‘object’, there is no moral relation; between ‘I’ and ‘you’, the whole relation is, and cannot but be, morally determined. We don’t have an independent sense of what ‘moral’ means that we could then subsume the I-you-relation under; rather, it is our being in the I-you-relation, with the experiences of love, of caring, of hurt and betrayal and bad conscience that this includes, that opens up morality to us, that gives the concept ‘moral’ any meaning at all (cf. Backström 2007 and 2017).

That the I-you-relation is a moral relation means, among other things, that the meaning or ‘effects’ of what is done within it, of how one behaves towards the other, cannot be limited to only this or that aspect of the relationship. Thus, if someone approaches me with an open, heartfelt smile, and I smile back in the same way, I cannot immediately proceed to mistreat them, whereas it is quite possible to smile politely (or regretfully, sardonically, devilishly, etc.) while stabbing someone in the back. This, obviously, doesn’t mean that sardonic smiles or stabs in the back aren’t morally determined, but rather that the difference between the cases is a moral difference, and that this difference is a matter of one’s whole way of relating to the other. The point about the heartfelt smile is that it is an expression of a wholehearted opening-of-oneself-to-the-other, which makes the thought of mistreating them impossible, because opening up in this way means that one’s resentments, envy, vengefulness, callous

egocentricity etc. – all the ways in which one closes oneself up in oneself and hardens oneself against the other; all the motives out of which terrible deeds are done – melt away.

This doesn't mean that opening or closing oneself is a once-and-for-all, on/off affair. But there's a tension, a conflict, between a longing to open up and a fearfully felt wish to close oneself to the other. Thus, insofar as one is determined to close oneself, keeping the determination involves refusing to see and respond to the other's smiles or other expressions as heartfelt, and instead misperceiving them – in a way that is neither deliberate in a conscious way nor independent of one's determination – as ingratiating, as pathetic naivety, etc. One is determined to close one's heart to the other, to harden it against her, and this means hardening oneself against one's own heart, refusing to feel the openness of the other that one perceives in one's own heart. One represses one's own heartfelt response, and because repression doesn't delete understanding and feeling but rather deforms them and makes them superficially unrecognisable, one doesn't simply feel nothing as one closes oneself to the other, but rather feels hard or numb, or feels a murderous rage, etc.

The sense in which interpersonal understanding is an inescapably moral matter can be further elucidated by considering our relation to pain – that crypto-Cartesian philosophers' favourite, because it can seem both indisputably real and essentially private ('I cannot doubt my pain, and it's *mine*; I suffer it, no-one else'). Reflection on our actual responses to our own pain and that of others quickly reveals the absurdity of this idea of essential privacy. Thus, a child's primitive response to pain is crying and looking to us for help and comfort, and the child's pain is no more doubtful – indeed, it may in an obvious sense be felt to be more real, attending to it more *urgent* – than one's own. As underlined above, in seeing you wince with pain I don't register a neutral movement; I see you *wince*. And as we also saw, this perception isn't only unmediated (non-inferential), but *moves* me, e.g., I may wince myself in response to your pain ('Ouch, that hurt!'). That is, I don't first see your pain, with the question of how, or whether, it *concerns me* raised, if at all, only after the perception. Rather, my very perception of you in pain is itself a mode of being concerned for you, affected by what befalls you; your pain pains me (in compassion), just as your sadness saddens me and your gladness makes me glad. Perceiving another's expressions means responding to them; to perceive is to be in relationship with the other.

It might seem easy enough to imagine someone perceiving another's pain without being in any way moved to compassion. But what are we actually imagining here? If there really were no essential connexion between seeing another's pain and responding compassionately,

we would be able to imagine someone seeing that another is in terrible pain, being alive to the terror of their predicament – for that’s what perceiving pain *as* pain, perceiving its painfulness, means – and yet remaining completely indifferent to it. But there’s nothing we can coherently imagine here, for nothing would show that the one who supposedly remains indifferent to the pain they perceive really perceives the pain. Even if they know that giving the other a pill calms them, or that pricking them makes them scream, that still shows no perception of *pain as pain*, for we are, *ex hypothesi*, imagining that these actions (administering pill or prick) are undertaken in a spirit of complete indifference – for purely instrumental reasons, or perhaps with the kind of curiosity one might display in finding out how a machine works. Insofar as someone seemed wholly to lack compassion, then, they would also seem unable to perceive pain (its painfulness), and although their lack of understanding might lead them to actions that caused others great suffering, they would be incapable of callousness or cruelty.

Precisely the possibility of morally destructive responses which, superficially, appears to demand distinguishing the mere understanding of what others feel (‘empathy’) from one’s morally significant response to it – say, compassionate (‘sympathy’) or cruel – actually reveals the untenability of this distinction.¹⁴ That perceiving pain as pain is inextricable from compassion obviously doesn’t mean that we always *act* compassionately towards others. But where a compassionate response in action is lacking this isn’t because there is simply a lack of compassionate understanding of the other; rather, this understanding, this being-moved-by-the-other’s-suffering, has been inhibited, either consciously suppressed or denied and repressed, derailed and perverted in various ways, due to the presence of other motivations and considerations which conflict with the compassion. These conflicting motives won’t only inhibit compassionate action but will also deform one’s feeling-response to the other, so that instead of simply feeling compassion, one feels, say, the malicious satisfaction of cruelty, *Schadenfreude*, disgust, or its more distanced and respectable cousin contempt; or perhaps one feels overwhelmed and helpless in the face of the other’s suffering and so retreats from them, leaving them to their suffering to protect oneself. There are countless variations on the theme of perverted compassion, including apparently ‘positive’ responses like *pity*, that privatisation of compassion where, instead of opening myself to you and understanding you and myself in our relatedness, I reduce my response to you to private ‘sentiment’ and you to a mere ‘suffering being’, the mere ‘object’ of my pity. But the important point is that none of these responses are simply alternatives to compassion, as though one would have them

instead of compassion. Rather, they are perversions of one's compassion; the compassion isn't simply absent in them, but present in repressed form. That is, cruelty and other destructive responses aren't responses to the other's pain directly, but to that pain as revealed by compassion. In cruelty, my very wish to hurt you shows how intimately your suffering affects and concerns me; if it didn't, if I didn't compassionately *care about you*, I wouldn't care to hurt you.

In other words, there's an *asymmetry* between openness to the other – of which compassion is one face or mode; as it were the colouring it receives from the shadow cast on it by the other's suffering – and closing oneself to them in cruelty, indifference etc. The latter are destructive/repressive responses to the former, and their destructiveness consists in the fact that one isn't just destroying some external object – there's in this sense nothing destructive about, say, demolishing a wall, just as such, because here we indeed simply have a person handling an object – but rather one is trying, hopelessly, to destroy the very connection and openness between oneself and the other. In one's destructiveness, one turns viciously on oneself as much as on the other. This can appear in various, and on the face of it very different, modes: one can destroy in a murderous frenzy, in a spirit of cold calculation or sentimental pity, etc. In these and countless other ways, one represses and deforms one's own inevitably felt sense of what one is doing to the other, of what transpires between oneself and them.

Reasoning supposed to justify one's behaviour often plays an important role in creating an appearance that one isn't callously closing oneself to the other – which on one level one cannot help feeling and knowing – but is doing something natural and unproblematic, or good or necessary. But the reasoning, too, takes on an emotional aspect when it is introduced into moral-existential contexts, i.e., into contexts where interpersonal relations are at stake. Thus the spirit of cold calculation is indeed a *spirit*, a mood, a mode of emotionality; hence its *coldness*, its callousness. The inescapability of the openness, the responsive moral concern and understanding between us shows, for example, in facts like this, that when people assume a calculative stance to others, their stance is indeed cold. Calculating isn't cold in itself; there's nothing cold about, say, a carpenter making calculations when making a table. What is cold, what chills, is that someone should try to repress their heartfelt understanding of themselves and the other through adopting towards them an attitude modelled on that to objects in the world.

The coldness shows – is one of countless illustrations of – the way in which it both is, and more fundamentally is not possible to actually model the I-you-relationship on the subject-object-relationship. And it also shows, to put the same point differently, the way in which the very life of our souls, our whole psychology, is a moral affair, is an expression of the way we're related to each other. It isn't just that one may judge our psychological responses as 'good' or 'bad' from an external moral perspective, in the light of some norms or ideals, but that *what* we respond to and what we *feel* is itself morally determined through-and-through; often in the form of these responses in themselves being ways of repressing, i.e., falsifying and perverting, their own moral significance, which is felt only 'unconsciously', in the way compassion is felt in cruelty, say. Morality is, most basically, a matter of our heartfelt responses – ineliminable but repressable – to others. These responses aren't based on norms or principles. On the contrary, the latter borrow whatever moral meaning they may have, for good or ill, from being variously related to the heartfelt interpersonal understanding which may also be called 'conscience'.¹⁵

5. The fear of openness: from polite distancing to brutal repression

We are inescapably open to each other's touch and address; if we weren't, we wouldn't understand each other, or ourselves, at all. My critical discussions of inferentialism, contextualism and the idea of a neutral perception of pain were all meant to show this: how, if one denies or ignores the openness, one can make no sense of our understanding each other. The attempt to imagine mind split off from body and one mind from others, or to imagine them connected only or primarily in an epistemic way, through a relationship of neutral 'knowing', leads to paradox and confusion. However, in a certain sense – always ambivalent and finally only apparent – we indeed *split* mind and body apart because and insofar as we have difficulties in relating to each other openly, that is, because we fear abiding in and welcoming the other in the openness that in another sense we cannot abolish. Succumbing to this fear we try, in ambivalent attempts that never finally succeed but can be endlessly persisted in, to close ourselves to and split ourselves apart from the other, which also means creating a split in *oneself* by hardening one's heart (that is: hardening oneself against one's heart). And as I said, there is a sense in which, in closing oneself to the other, one also splits 'body' from 'soul', 'inner' from 'outer'. Thus, if one feels that someone is closing himself, one might think '*This* is the face he shows the world, inwardly he has another one', but as

Wittgenstein observes, ‘this does not mean that when his expression is *genuine* he has two the same’ (1958, §606). That is, if someone addresses me openly and I respond in the same way, I won’t feel that his ‘outer behaviour’ is an accurate representation of his ‘inner state’, but rather there’s no place for the very distinctions inner/outer, being/appearance; I simply respond to *him*. The distinction, the split, enters with the closing, with distrust. However, if the other’s way of relating to me strikes me as suspicious and I get a sense that he is hiding something from me, then even if his ‘inner’ is now hidden from me, his *attitude* to me is revealed in his behaviour no less than another person’s openness is revealed in their open smile. I know that he is hiding something, even if I don’t know what it is. When one closes oneself to another, this closing will itself be expressed, revealed, in one’s demeanour, and in this sense one’s mind (one’s inner life) isn’t really split from one’s body (one’s behaviour) here, anymore than when one is open.

Distrust and dissimulation aren’t exceptional occurrences, contrasting with a supposedly ‘normal’ trust and openness, but pervasive features of what is regarded as normal, everyday living – although it is part of this very normality not to *present* things (quite) in this way. ‘Normal’ distrust appears in the form of pervasive anxieties over how others will respond to one, how one will be received, manifesting and leading to a fear of any wholly open, unguarded address, even in our most intimate and so-called personal relationships. The fear of openness takes different surface forms, e.g., one might fear that one’s interest in the other will be experienced as intruding, that one’s longing for closeness will be felt as embarrassing or disgusting, or again that the other will reveal something in oneself that one doesn’t want anyone to see, or to have to face oneself. Whatever the case, one tries to adjust one’s expressions in order to manage the impression one makes. Thus, one tries to calm, encourage or flatter as seems demanded by the situation – that is, by one’s own agenda for the encounter, usually a rather inchoate one, more instinctively fantasised and felt than explicitly formulated. In plain terms, one tries to manipulate the other by modifying what has now *become* one’s self-*presentation*, as opposed to one’s simple self-expression to the other.

One standard form of avoiding open, direct engagement is to keep things on a *polite* level, which means choreographing our encounters in such a way as to avoid raising the question of what one’s true feelings and thoughts are; one keeps on the social surface, deliberately leaving the depths of the soul untouched – and there are depths of light, not only of darkness. That is, one substitutes the management of appearances for reality. In uttering ‘How are you?’ you’re not typically asking a genuine question, and in replying ‘Fine’, I’m

typically neither masking how bad I feel nor expressing how fine I feel, although we use a form of words whose primary use is to ask about and express real feeling; ‘primary’ not in the sense of more prevalent, but in the sense that polite conversation ambivalently *plays at* being real. If it is obvious from your tone that you’re not in the least interested in how I feel, your ‘How are you?’ isn’t polite, but rude or mechanical. In our conversations there’s a pervasive play and tension between polite performance and a longing for real, open contact – which, however, is also feared. The function of politeness is precisely to manage this fear in a discreet way, without making its character of fear apparent; conversation becomes a kind of balancing act in which one tries to be ‘natural’ and affable, so that one’s cautious reserve doesn’t become too conspicuous while, at the same time, this reserve, politely called respect for the other, should be clear enough that both parties can relax, feeling secure in their sense that the other won’t challenge them by ‘too’ direct an address.

The fear of openness isn’t only a fear of the other knowing what one really thinks and feels, but a fear of finding out and acknowledging this for oneself. One’s ‘inner’ experience and one’s attitude towards the other are inextricably intertwined. Here’s a banal example: you start expressing your delight at a film you saw, but sensing that the other thought the film bad, you instinctively retract, fall silent or try to explain your delight away as a misunderstanding, etc. This isn’t typically mere concealment of something you yourself are clearly aware of, because in fact your own sense of your reaction, your articulation and development of it – and in this sense the very reaction itself – is something you gain only in the course of expressing it, giving expression to your delight, articulating and exploring it in a process which might also, we should note, end in your modifying or even eventually rejecting your own first reaction. All this you cut short, however, because you sense or fear that the other won’t appreciate it, and so not only they but you yourself are robbed of the opportunity to find out what you really thought about the film. On a less banal level, this same kind of self-censoring self-obfuscation in the interest of preventing truthful self-revelation turns marriages and other intimate relationships into terrifyingly painful comedies of error, where neither party knows who they or the other are – and, of course, even expressing one’s views about a film seems fearful, ‘impossible’, only insofar as it is connected to moral-existential-emotional anxieties, i.e., to things far from banal.

The film-example illustrates how, in trying to hide oneself from others one succeeds only in losing oneself, while *in* one’s very hiding-operations one’s fearfully awkward self stands painfully revealed – to those who have eyes to see. But we often close our eyes. If this

weren't so, there would be no self-hiding games, no comedies of error, for in one's self-obfuscation one needs collaborators; people who *accept* one's self-effacement and falsifying self-presentations, who don't 'notice' that there's anything funny going on. And one generally finds them, because others are as anxious about openness as oneself. Thus, mutual pacts of assistance-in-repression are constantly formed; 'understandings' *not* to understand, agreements to accept appearance for reality, censored utterances and fake gestures for genuine expressions – agreements which are by definition unspoken and whose existence will always be denied. Politeness is pervasively used for the ends of such collective repression; the agreement of various powerful groups to find their disdainful and brutal treatment of people from socially weaker groups quite 'natural' is a darker aspect of the same dynamics. Thus, a rich man can convince himself that his mistreated servant shows genuine gratitude when he for once treats her a little less brutally than he would have a 'right' to, and that what *he* himself feels and expresses is genuine sympathy – even if her anger and contempt for him, and his conceit and lack of wholehearted concern for her are obvious for anyone with eyes to see. But the rich man and his rich friends have agreed to close their eyes to this kind of thing; they help each other uphold a mendacious story about what goes on between them and their servants, and the centrepiece of this collective repression is a falsifying, comforting 'management' of the feeling-responses evoked in them by their encounters with servants; they harden themselves against and malign certain responses, while sentimentally indulging and magnifying others. It isn't that they simply lack any understanding of what goes on between them and their servants, for the openness between us is impossible to simply erase; rather, they systematically pervert and misrepresent their own sense of what goes on in their relationships.¹⁶

In the case of rich and poor, there are obvious material incentives for the rich to 'fail to see' the situation as it is. But opening oneself to another is fearful in itself, even where one stands to lose no money from it – is it not? When you try to close yourself to someone, you may often imagine that what you most fear is being ridiculed and not taken seriously by them. But isn't being taken seriously, having someone turn to you who really wants to know you, know what *you* think and feel, actually an even more daunting experience? After all, the only way in which your words and actions gain *importance* is by being addressed to someone who cares about you, and so about what you say and do, and who will, therefore, respond to you truthfully rather than politely or disdainfully passing it over; in turn forcing a response from you. That is, you'll feel 'forced' to respond only insofar as responding frightens you; insofar

as you welcome the other's address, you will experience your response as being called forth and enabled by it. The central point is that acting, speaking and listening truthfully aren't isolable acts that lead only to a particular, foreseeable effect, but rather truthfulness, wholehearted openness, sets a whole new dynamics in motion: when someone tells the truth in this eminent sense she is, as Adrienne Rich says, 'creating the possibility of more truth around her' (1979, p. 191). This possibility is precisely what we pervasively fear and try to obstruct.¹⁷

6. Lived confusion, philosophical mystification

The self-induced blindness of repression isn't a necessary feature of human relationships. It marks them only insofar as there is – as there pervasively, but *not* necessarily or always, is – destructiveness and fearful distrust, be it intense or ever so slight, between people. Only then is 'inner' ambivalently split from 'outer', and an interminable uncertainty about their relation engendered. This split is the concomitant of one of us – contingently, in succumbing to our fear of openness – splitting off from the other, 'I' from 'you'. The 'inner life' that I pretend is '*only* mine' and that others do not or 'cannot' know – and with it, as the other side of the coin, the life that is 'merely outer', mere behaviour, are ambivalent semi-realities, fantasies that arise out of the rejection of the openness between us. They appear as we close ourselves to each other in suspicion, irritation, shame, self-pity, envy or some other fearful, self-centred and destructive attitude – and closing oneself to others is, indeed, a closing of *oneself* (to oneself).

'The' mind, in the general singular as conceived by philosophers, is in fact an amalgam of aspects of our life as we imagine it from the perspective of self-centred, destructive fear, so that philosophers tacitly promote an alienated and deformed form of human experience and relatedness into the norm for experience as such – thus making its morally and existentially problematic character invisible. The alienated picture of the mind prevalent in philosophy mirrors the existential alienation and mistrust in life.¹⁸ And precisely because the latter isn't some rare anomaly or pathology, but a pervasive feature of what we regard as normal, everyday existence; because the philosophers' refusal to consider openness in their theorising mirrors our everyday refusals to be open, the framing of the problems that dominates – indeed, constitutes – philosophy of mind as we know it, may seem self-evident and self-explanatory. But it isn't.

Self-misrepresentation marks not only the philosophical theorisation of alienated forms of experience, but already the way they are lived. Fearful and destructive attitudes are characterised by the fact that they aren't what we, when we are prey to them, make them out to be. Thus, in envying you, I pretend to despise you, although my very envy shows that I actually intensely wish to be like you; in fearfully distrusting you my distrust comes out in my (projectively) seeing *you* as not trustworthy, etc.¹⁹ However, the mirroring of the lived attitude in philosophical fantasy isn't straightforward; the Cartesian-cum-materialist picture of the mind-body and self-other relation isn't a simple replica of a suspicious mind-set, say. As my discussion of the inconceivability of 'inferring inner mental states' from 'colourless' movements was meant to show, the philosophical picture is senseless *in a way* that a lived attitude isn't. Nonetheless, there are crucial connections, for the senselessness of the philosophical view results from trying to formulate in a logically coherent picture certain aspects of the destructiveness that is lived rather than formulated in everyday attitudes, and that indeed cannot be coherently formulated, for what the destructiveness is set to destroy is precisely the *sense* of life, that is, the expressive openness and vital connection between 'I' and 'you', and thus also, by extension, between our statements and thoughts. Trying to make sense of destructiveness can, in the end, only produce senselessness; the only kind of sense the envious person, say, can make of their own demeanour is self-deceptive, a mere semblance of sense. While we can try to cut ourselves off from and close ourselves to others, we cannot finally succeed in doing so, nor can we produce a lucid account of what it is we're trying to do. My discussion has aimed to shed some light on the topsy-turvy world that results, in philosophy and in life, from these doomed attempts.

References

- Apperly, I. 2010. *Mindreaders: The Cognitive Basis of 'Theory of Mind'*. Hove, East Sussex: Psychology Press.
- Backström, J. 2007. *The fear of openness. An essay on friendship and the roots of morality*. Åbo: Åbo Akademi University Press.
- Backström, J. 2011. 'Wittgenstein and the moral dimension of philosophical problems'. In O. Kuusela and M. McGinn, eds., *The Oxford Handbook of Wittgenstein*. Oxford: Oxford University Press.
- Backström, J. 2013. 'Wittgenstein, follower of Freud'. In Y. Gustafsson, C. Kronqvist and H. Nykänen, eds., *Ethics and the Philosophy of Culture: Wittgensteinian Approaches*. Newcastle upon Tyne: Cambridge Scholars Publishing.

- Backström, J. 2014. 'Touchy subjects: The theme of repression in Freud and Wittgenstein'. *European Journal of Psychoanalysis* 2/2014. Published online at: <http://www.journal-psychoanalysis.eu/>.
- Backström, J. 2015. 'Of dictators and green-grocers: On the repressive grammar of values-discourse'. *Ethical Perspectives* 22, 39–67.
- Backström, J. 2017. 'From nonsense to openness: Wittgenstein on moral sense'. In E. Dain and R. Agam-Segal, eds., *Wittgenstein's Moral Thought*. New York: Routledge.
- Backström, J. 2019. 'Hiding from love: The repressed insight in Freud's account of morality'. In R. Gipps and M. Lacewing, eds., *The Oxford Handbook of Philosophy and Psychoanalysis*. Oxford: Oxford University Press.
- Backström, J. forthcoming. 'Pre-truth life in post-truth times'. *Nordic Wittgenstein Review*.
- Backström, J. and Nykänen, H. 2016. 'Collectivity, evil and the dynamics of moral value'. *Journal of Evaluation in Clinical Practice* 22, 466–76.
- Carruthers, P. and Smith, P. K. 1996, eds. *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Cockburn, D. 2009. 'Emotion, Expression and Conversation'. In Y. Gustafsson, C. Kronqvist and M. McEachrane, eds., *Emotions and Understanding: Wittgensteinian Perspectives*. Houndmills, Basingstoke: Palgrave Macmillan.
- Cook, J. 1969. 'Human Beings'. In P. Winch, ed., *Studies in the Philosophy of Wittgenstein*. London: Routledge & Kegan Paul.
- Davies, M. and Stone, T. 1995, eds. *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell.
- De Jaegher, H. and Di Paolo, E. 2007. 'Participatory Sense-Making: An enactive approach to social cognition'. *Phenomenology and the Cognitive Sciences* 6, 485-507.
- Dennett, D. 2003. *Freedom Evolves*. Harmondsworth: Penguin.
- Dennett, D. 2006. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge, Mass.: MIT Press.
- Ebersole, F. 1967. 'Where the Action Is'. In: Ebersole, *Things We Know: Fourteen Essays on Problems of Knowledge*. Eugene: University of Oregon Books.
- Ekman, P. 1989. 'The argument and evidence about universals in facial expressions of emotion'. In H. Wagner and A. Manstead, eds., *Handbook of social psychophysiology* 58, 342–353.
- Epley, N. and Waytz, A. 2009. 'Mind Perception'. In S. Fiske, D. Gilbert and G. Lindzey, eds., *The Handbook of Social Psychology*. 5th ed. New York: Wiley.
- Foolen, A., Lüdtke, U.M., Racine, T.P. & Zlatev, J. 2012, eds. *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language*. Amsterdam: John Benjamins.

- Gallagher, S. 2008. 'Inference or interaction: Social cognition without precursors'. *Philosophical Explorations* 11 (3), 163-73.
- Goldman, A. 2006. *Simulating Minds*. Oxford: Oxford University Press.
- Heal, J. 1995. 'Replication and Functionalism'. In M. Davies, and T. Stone, eds., *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell.
- Hertzberg, L. 2009. 'What's In a Smile?'. In Y. Gustafsson, C. Kronqvist and M. McEachrane, eds., *Emotions and Understanding: Wittgensteinian Perspectives*. Houndmills, Basingstoke: Palgrave Macmillan.
- Hobson, P. 2002. *The Cradle of Thought*. London: Macmillan.
- Hoque, M. E., McDuff, D. J. and Picard, R. W. 2012. 'Exploring Temporal Patterns in Classifying Frustrated and Delighted Smiles'. *IEEE Transactions on Affective Computing* 3(3), 323–334.
- Hull, C. L. 1943. *Principles of Behavior: An Introduction to Behavior Theory*. Oxford: Appleton-Century-Crofts.
- Hutto, D. 2008. *Folk Psychological Narratives: The Socio-cultural Basis of Understanding Reasons*. Cambridge, Mass.: MIT Press.
- Kristjánsson, K. 2004. 'Empathy, sympathy, justice and the child'. *Journal of Moral Education* 33(3), 291–305.
- Krueger, J. and Overgaard, S. 2013. 'Seeing Subjectivity: Defending a Perceptual Account of Other Minds'. In S. Miguens and G. Preyer, eds., *Philosophische Analyse/Philosophical Analysis: Consciousness and Subjectivity*. München: Walter de Gruyter.
- Leslie, A.M. 1987. 'Children's understanding of the mental world'. In R. Gregory, ed., *The Oxford Companion to the Mind*. Oxford: Oxford University Press.
- Leudar, I. and Costall, A. 2009. *Against Theory of Mind*. Houndmills: Palgrave Macmillan.
- Levin, D. M. 1993, ed. *Modernity and the Hegemony of Vision*. Berkeley: University of California Press.
- MacIver, A.M. 1964. 'Is There Mind-Body Interaction?'. In Vesey, G., ed., *Body and Mind*. London: Allen & Unwin.
- Merleau-Ponty, M. 2002. *Phenomenology of Perception*. C. Smith, transl. London: Routledge & Kegan Paul.
- Midgley, M. 2014. *Are You an Illusion?* Abingdon: Routledge.
- Nykänen, H. 2002. *The 'I', the 'You' and the Soul: An Ethics of Conscience*. Åbo: Åbo Akademi University Press.
- Nykänen, H. 2005. 'Heidegger's Conscience'. *Sats – Nordic Journal of Philosophy* 6, 40–65.
- Nykänen, H. 2009. *Samvetet och det dolda – om kärlek och kollektivitet*. [Conscience and the hidden: On love and collectivity] Ludvika: Dualis.

- Nykänen, H. 2014a. 'Conscience and collective pressure'. *Philosophy, Psychiatry, & Psychology* 21, 51–65.
- Nykänen, H. 2014b. 'Freud's Dangerous Pupil'. *European Journal of Psychoanalysis*. Published online at: <http://www.journal-psychoanalysis.eu/freuds-dangerous-pupil/>.
- Nykänen, H. 2015. 'Repression and Moral Reasoning: An Outline of a New Approach in Ethical Understanding'. *Sats – Nordic Journal of Philosophy* 16, 49–66.
- Nykänen, H. 2018. 'Wittgenstein's Radical Ethics'. *European Journal of Psychoanalysis*. Published online at: <http://www.journal-psychoanalysis.eu/wittgensteins-radical-ethics/>.
- Overgaard, S. 2007. *Wittgenstein and Other Minds: Rethinking Subjectivity and Intersubjectivity with Wittgenstein, Levinas, and Husserl*. New York and London: Routledge.
- Plato (1997). *Complete Works*. J. M. Cooper, ed. Indianapolis: Hackett.
- Ratcliffe, M. 2007. *Rethinking Commonsense Psychology*. London: Palgrave-Macmillan.
- Reddy, V. 2010. *How Infants Know Minds*. Cambridge, Mass.: Harvard University Press.
- Rich, A. 1979. *On Lies, Secrets and Silence*. London: Virago.
- Russell, J. and Fernández-Dols, J.M. 1997, eds. *The Psychology of Facial Expression*. Cambridge: Cambridge University Press.
- Ryle, G. 1949. *The Concept of Mind*. Harmondsworth: Penguin.
- Sartre, J.-P. 1966. *Being and Nothingness*. H. Barnes, transl. New York: Washington Square Press.
- Satne, G. and Roepstorff, A. 2015. 'Introduction: From Interacting Agents to Engaging Persons'. *Journal of Consciousness Studies* 22, 9–23.
- Scheler, M. 1954. *The Nature of Sympathy*. P. Heath, transl. London: Routledge & Kegan Paul.
- Spaulding, S. 2015. 'On direct social perception'. *Consciousness and Cognition* 36, 472– 82.
- Stern, D. 1977. *The First Relationship: Infant and Mother*. Cambridge: Harvard University Press.
- Stone, T. and Davies, M. 1996. 'The Mental Simulation Debate: A Progress Report'. In P. Carruthers and P.K. Smith, eds., *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Thompson, E. 2001, ed. *Between Ourselves: Second Person Issues in the Study of Consciousness*. Thorverton: Imprint Academic.
- Trevarthen, C.B. 1979. 'Communication and cooperation in early infancy: A description of primary intersubjectivity'. In M. Bullowa, ed., *Before Speech*. Cambridge: Cambridge University Press.
- Wittgenstein, L. 1951. *Tractatus Logico-Philosophicus*. C.K. Ogden, transl. London: Routledge & Kegan Paul.

Wittgenstein, L. 1958. *Philosophical Investigations*. G.E.M. Anscombe, transl. Oxford: Blackwell.

Wittgenstein, L. 1982. *Last Writings on the Philosophy of Psychology, Vol I*. G.H. von Wright and H. Nyman, eds.; C.G. Luckhardt and M. Aue, transl. Oxford: Blackwell.

Zlatev, J., Racine, T.P., Sinha, C. and Itkonen, E. 2008, eds. *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam: John Benjamins.

Nykänen (this volume) argues that the subject-object problematic provides the form of philosophy as such – i.e., of philosophy as almost exclusively conceived in our tradition – and that this form is *morally determined* precisely by the exclusion of the I-you-perspective. Leoni (this volume) underlines how the ‘subjective’ is always ‘objectified’, and *vice versa*, rendering the oscillation between them simultaneously inescapable and illusory. In my view, the deadlocks Leoni describes, which tend, as he notes, to turn comprehension into incomprehension, arise because the I-you-perspective is excluded, as Nykänen explains. My chapter aims to bring out aspects of this manufacture of confusion.

² ‘For the first several weeks after birth, the majority of the baby’s awake alert time is spent in and around feeding ... What will he see? It turns out that ... during feeding, mothers spend about 70 percent of the time facing and looking at their infants. Accordingly, what he is most likely to look at and see is his mother’s face, especially her eyes’ (Stern 1977, p. 36). I don’t pretend to decide philosophical questions by appeal to empirical evidence; philosophical puzzlement concerns not what the facts are but how one understands them. Nonetheless, reminders of simple facts sometimes help loosen the grip of apparently ‘natural’, but actually quite crazy pictures.

³ I avoid the standard philosophical term ‘intersubjectivity’, for insofar as ‘subject’ and ‘object’ mutually define each other, one cannot conceive a ‘between’ subjects with no object involved. When ‘I’ and ‘you’ meet, we’re (primarily) neither subjects nor objects.

⁴ Philosophers have typically treated eyesight as a model for all perception and cognition, including the ‘vision’ of philosophical truth itself (cf. the essays in Levin 1993), but I know of no extended philosophical discussion of the phenomenon of meeting someone’s eyes. The experience of looking *at*, and being looked at by, others, has been discussed, notably by Sartre (1966, pp. 340–400), who brilliantly reveals the *aporias* arising from trying to conceive interpersonal relations on the subject-object model. Seeing no other way of conceiving the matter, however, Sartre declares these *aporias* to be inherent in the human condition, rather than arising from confused philosophical fantasies about it. This illustrates the dominance of the quasi-solipsistic subject-object paradigm, or rather delusion, in philosophy, in which encounters between ‘I’ and ‘you’ are reduced to games the ‘I’ plays with ‘its’ perceptions or objects or meanings. For another illustration, consider that Plato, on the (to my knowledge) only two occasions when he mentions the phenomenon of looking into someone’s eyes (*Alcibiades I*, 133a; *Phaedrus*, 255d – both in Plato 1997), fastens on the same curious feature, that when looking into your eyes I may see a small reflection of *myself* in your pupil!

⁵ This is the main idea of Backström (2007), which, exploring the perspective first elaborated by Nykänen (2002), traces some of the endless ramifications of this fact. See also Nykänen (2009).

⁶ For mainstream approaches, see e.g. Apperly (2010), Carruthers and Smith (1996), Davies and Stone (1995), Goldman (2006). There is of course also a diverse literature critical of the inferentialist mainstream, e.g., Hutto (2008), Leudar and Costall (2009), and Ratcliffe (2007). Some of these critics – e.g. Gallagher (2008), Krueger and Overgaard (2013) – think of themselves as presenting alternative, ‘direct social perception’ accounts of interpersonal understanding; see Spaulding (2015) and the other articles in the same journal special issue for a sense of the current state of this debate. As I’ll explain below, the crucial difference between this literature (further references below) and my approach, some obvious similarities notwithstanding, concerns its general neglect of the morally determined nature of interpersonal, and therefore also of philosophical, understanding.

⁷ Cook (1969) makes a similar point.

⁸ Indeed, the very idea of ‘colourless behaviour’ or ‘mere bodily movements’ is perplexing. It isn’t just that we don’t normally see it; it is hard to know what ‘it’ would even be. Cf. Ebersole’s discussion (1967) of the apparent impossibility of finding or constructing examples of ‘mere’ bodily movements that wouldn’t be *involuntary*, like spasms, and even a spasm is a *particular* kind of movement, not ‘mere’ movement. Summarising his discussion, Ebersole underlines the abstraction and specificity – contradicting the *supposedly* basic character – of the philosophical idea of the ‘body’: ‘A domino is a piece of wood seen from a special point of view. A person is not a body seen from a special point of view. Rather, a body is a person seen from a special point of view’ (1967, p. 303).

⁹ For further analysis of the ‘private language’ considerations, see Read (this volume) and Toivakainen (this volume).

¹⁰ On this debate, see, e.g., Ekman (1989), Russell and Fernández-Dols (1997).

¹¹ This is a crucial theme in Wittgenstein’s later writings, as is well brought out in Nykänen (2014b) and (2018), with whose radically ethical understanding on the I-you-perspective I essentially agree. – Let me note that the point about understanding not ‘consisting’ in anything, and my whole discussion, might *seem* to be (but isn’t in fact) ‘disproved’ by the fact that computer-programs are, apparently, to some extent able to discriminate facial expressions, for example to distinguish different kinds of smiles, e.g., to tell smiles masking frustration from smiles of delight (Hoque *et al.*, 2012). This is quite a feat of engineering, but hardly surprising as such, as there are certainly characteristic (although not exceptionless) differences between typical cases of different kinds of smiles; polite or frustrated smiles don’t usually include ‘smiling eyes’, for example, and the latter can be coded in terms of the movement of muscles around the eye. But the salient point is that it is only because we understand and care about each other, are moved by each other in all kinds of ways, for example to smile, that we make, and can make, these distinctions – and that some of us may then be motivated to build machines that can simulate something like our ability here. And even if the machine tended (as reported) to be more reliable than untrained human observers in distinguishing certain kinds of smiles from

others under certain conditions, that doesn't show that it *understood* anything about smiles or people. On what criterion did the experimenters decide the machine was more reliable than human subjects in discriminating between different kinds of smiles? By using their own understanding of the people smiling, obviously. And that the machine was programmed to discriminate cases (not: understand them) based on various measurable indications of muscle-movement etc., doesn't mean that our understanding works in the same way; my earlier discussion should have shown the hopeless paradoxes one gets into if one tries to conceive understanding of others as built up from discreet pieces of information of this kind.

¹² This isn't mere social decorum; people are often determined not to lose their cool, not to allow themselves to feel too much, even in situations where no one would mind, or where there is no-one to witness it.

¹³ This critical literature includes authors inspired by Wittgenstein, e.g., Cockburn (2009), Hertzberg (2009), and Overgaard (2007), and authors in the phenomenological tradition, starting with Scheler (1954) and Merleau-Ponty (2002) and comprising the 'direct social perception'-theorists referred to in endnote 6 above. For further 'intersubjective', 'interactional' or 'second personal' approaches in philosophy, cognitive science and developmental psychology, see De Jaegher and Di Paolo (2007), Foolen *et al* (2012), Hobson (2002), Reddy (2010), Satne & Roepstorff (2015) and the other articles in the same journal special issue; Thompson (2001), Trevarthen (1979), and Zlatev *et al* (2008).

¹⁴ Kristjánsson (2004), for example, distinguishes 'empathy' and 'sympathy' in this way, but there's no well-defined standard use of these terms in the literature.

¹⁵ For more on conscience and its repression through destructively emotionalised and moralised responses, see Backström (2007, pp. 317 ff.), (2015) and (2019); Nykänen (2002), (2005), (2014a), (2015).

¹⁶ In repression, 'collective' and 'private' aren't opposed, but two faces of the same depersonalising movement. One repressively privatises one's responses, distancing oneself from the other by reducing her to the 'object' one 'subjectively' reacts to, but one does this precisely by representing the situation in general terms collectively available in the culture. For example, I turn viciously on someone and excuse my brutality towards her by saying '*I* couldn't help *myself* [1]; when *a person* behaves so rudely [2], *one* just gets so angry [3].' I thus [1] privatise/subjectivise my response, [2] objectivise the other, and [3] refer to collective 'understandings' of how 'one' responds in situations of particular kinds, and in each moment I disclaim responsibility for how I responded to the other; [1] claims that I can't help my responses, [2] that the other was the cause of my response, [3] that everyone responds in the same way. For more on repression, depersonalisation and collectivity, see Backström (2014), Backström and Nykänen (2016), Nykänen (2009) and (2014a).

¹⁷ For further discussion of the pervasive hostility to truthful communication and understanding, see Backström (forthcoming).

¹⁸ For discussion of how philosophical debates quite generally reflect existential confusion, see Backström (2011) and (2013), and Nykänen (this volume).

¹⁹ For more on such essentially misfelt/misrecognised responses, see Backström (2019).