

## ON LEARNING THROUGH CONFUSION

Lars Hertzberg

1.

To do philosophy, a man must be able not only to see questions where those not given to philosophy see none, but also to look on those questions in a particular way. Not wanting to dismiss the questions, nor get rid of them through any sort of answer, or to show that they are a sort of needless worry to be put out of mind... Trying rather to understand these questions – and from this angle or in this sense to understand human thinking and human investigation and human life; to understand how they arise in, and in one sense belong to, our thinking about questions that we ask and answer. This goes with *contemplation* of the ways in which people think and inquire – e.g. in trying to solve problems of physics, or in connexion with moral problems. And this is difficult. Perhaps especially so in a culture which has become as technological as ours – as much preoccupied with getting things done, with how to do things, with results.

(Rush Rhees, "The Fundamental Problems of Philosophy", *Philosophical Investigations* 17 (1994), pp. 578 f.)

Philosophical inquiry is inquiry into the nature of philosophy. When trying to get clear about whatever question we are up against we are at the same time trying to get clear about what a philosophical treatment of the question would amount to. Indeed once we are clear about that, we are done with the question, more or less. If this is so, that means that in philosophy there is no "method" in the sense of a procedure taking you from where you are to where you want to go. Once you know *where you are*, the rest is easy.

Even so, apart from the way in which the nature of philosophy becomes an issue in the doing of philosophy, philosophers will sometimes discuss the nature of philosophy as a question in its own right. We may feel the need to do so, for instance, in teaching introductory philosophy. What is the status of such "meta-philosophical" discussion? I should like to suggest that we are strongly tempted to misunderstand it. It might be thought that if our collective endeavours in philosophy are to make sense, it must be agreed from the outset what the problems and the methods are. We must, as it were, be clear about matters of jurisdiction and procedure. The temptation to think about

philosophy in this way, I believe, comes from our regarding philosophy as analogous with other academic disciplines.<sup>1</sup>

Philosophers often take the easy way out, by declaring that philosophy is “conceptual inquiry”, the contrast being made with disciplines such as history or the natural sciences, which, it is said, are concerned with empirical inquiry. Saying that philosophy consists in conceptual inquiry is supposed to tell us something about its subject matter as well as its methods. Factual questions are *a posteriori* and can only be settled through records, observation and experiment, conceptual questions are *a priori* and we are to get clear about them through conceptual analysis.

On the whole, this answer only invites a new set of problems. (For one thing, the story only holds up as long as we do not inquire too closely into the notion of the *a priori* that is invoked here.) The answer fails to make it clear in what way philosophy is different from formal disciplines like mathematics or symbolic logic. In any case, this view has had some impact on the way many analytic philosophers have thought about their discipline. Thus, philosophy as traditionally carried out has seemed to them simply to be a kind of preparation for the real work, which would consist in setting out the formal structure of our concepts in the shape of a logical calculus. Others have thought that the task of philosophy is the faithful recording of ordinary linguistic usage. To others again, it has seemed that the distinction between philosophy and empirical science cannot really be upheld, since there is no principled way of making the distinction between conceptual and empirical truths. Philosophy is simply the most general branch of natural science. The best-known expression of this view is W. V. O. Quine’s influential essay “Two Dogmas of Empiricism”<sup>2</sup>. This view, too, has had a vast impact on the way philosophers see their own activity. Either way, it has come to seem hard to maintain the identity of philosophy as a discipline in its own right.

It could be argued that invoking the empirical-conceptual distinction has done more to obscure the character of philosophy than to clarify it. The focus has been on identifying a particular subject matter when it should have been on the *attitude taken* to the subject matter. For one thing, we obviously are not taught to *do* philosophy by being given a field of inquiry and a set of tools. There is no way I could acquire, before I start doing philosophy, the particular way of looking at questions that Rhees talks about. Nor is this way of looking at questions something that I will one day possess

---

<sup>1</sup> Still, the reservations we might have about describing philosophy along those lines might apply to many of the sciences too.

<sup>2</sup> In Quine, *From a Logical Point of View*, 2<sup>nd</sup> ed. (New York and Evanston: Harper & Row, 1963).

once and for all. Rather, becoming versed in this field means coming to accept the need to recapture this “particular way of looking at questions” over and over.

We learn to do philosophy in the doing of it. I believe Rhees would have said that learning to do philosophy is not so different from learning to participate in discussions. Part of what happens is that I become aware of the limits of meaningful discussion. Some of my remarks will be met with counter-questions or counter-arguments: I may be asked whether I actually mean this and whether I’d also accept that apparent consequence of what I am saying, etc. Other remarks may be met with edge-marking responses like, “Well, if you say so”, or “That’s something we’d have to look up”<sup>3</sup>, or “All right, that’s your opinion”, indicating that if I wish to press this particular point, discussion will turn into something else: a fact-finding affair, a shouting match, name-calling or what have you.

Similarly, in learning to do philosophy our faltering attempts at argument (if our teachers are any good) will not be met with the reaction “That isn’t philosophy”, or “You can’t say that here” (which would be rather like countering argument in a political discussion with, “That’s a bourgeois view” or “That’s socialism”); rather, they will be met with counter-arguments or edge-markers. In either case, there are no specific rules or skills that you learn to master. What you learn, in a sense, is staying the course. In the process, you arrive at something you are ready to acknowledge as what you really meant to say.

Since philosophy is learnt in the doing of it, then, I would suggest that remarks about the nature of philosophy are not comparable to the remarks by which a disciple of, say, chemistry or art history is brought into his subject. General remarks about the nature of philosophy are not much use as guidelines aimed at someone about to enter the field. Rather, they are reflective remarks, belonging in a discussion between people who are engaged in the activity, baffled by it. Above all, they are made *in the course* of the activity.

A similar point is expressed by Sören Stenlund as follows:

... it is something characteristic of philosophy that the questions “What is philosophy?” and “How shall a philosophical problem be dealt with?”, are themselves philosophical questions on the same level as other philosophical questions.

---

<sup>3</sup> Any discussion *presupposes* that the facts are not contested, or better: that there is something that, for the purposes of the discussion, is recognized as “the facts”, and which is left uncontested.

He goes on to make another important observation:

And I want to add that a philosophical work is alive only when these questions, directly or indirectly, are present in it. Philosophy is alive only when it is open to the problem of its own nature and task. Philosophy is alive when it is problematic in its own eyes.<sup>4</sup>

2. If philosophy is not a knowledge affair with its own method and subject matter, some have argued, maybe it should be likened to a form of intellectual therapy, the activity of ridding oneself of philosophical confusion. To others, this has seemed to belittle the worth of the activity. It does not explain why we should not simply try to stay clear of these worries, putting them out of our minds as quickly as possible. But on the other hand, if these are important worries, *what makes* them important? Is it the fact that we may learn something from them, and if so, what? In order to get a grip on these questions, let us take a closer look at a couple of cases of philosophical confusion.

I have two examples in mind. The first case is the predicament of someone who does not see why computers could not be made to speak a language; or rather, why there should be any significant difference between an artificial speaker and a human speaker. After all, what is speaking other than displaying distinctive behavioural reactions to various physical stimuli? Provided the end result is indistinguishable from what is produced by a human speaker, nothing else should count. It should not really matter by what processes the result is brought about.

Of course, this person will admit that the computer should be able to do more than simply produce speech sounds that we could recognize as sentences of our language. It would have to have the capacity to emit those sounds in a meaningful way. How he thought about that would depend on his philosophical outlook. He might think that what is required in addition is a sensitivity to the truth-conditions or to the assertion conditions of the sentences emitted. Then again, he might think that paying attention only to truth-conditions or assertion conditions would betray an impoverished conception of language. What is required in addition, he might argue, is a capacity to operate in accordance with the rules of interaction associated with various linguistic and extra-linguistic practices. In formulating this requirement, he might feel he was being true to the insights conveyed

---

<sup>4</sup> Sören Stenlund, "Philosophy and Contemporary Science" / [www.bu.edu/wcp/Papers/Cont/ContSten.htm/](http://www.bu.edu/wcp/Papers/Cont/ContSten.htm/).

by Wittgenstein in *The Brown Book* and *Philosophical Investigations*, and focused around the notion of a language game.

Whatever his philosophical outlook, this person might agree that the technical problems are huge, even insurmountable for the foreseeable future, perhaps forever. That does not matter to him. What he cannot see is why the distinction between a human and an artificial speaker should be anything *other* than technical.

It might be thought a common sense response to this idea is to say that the computer would be lacking something essential: an inner life. No matter how skillfully it simulated the linguistic behaviour of ordinary humans, it would still not be expressing beliefs, wishes, confidence or hesitation, etc, in the sense in which we do.<sup>5</sup> To this a retort might be found, again, purportedly, in Wittgenstein's coffer: any appeal to the inner is out of bounds here, unless outward criteria can be assigned to it. And this gets us right back where we started, since, it would seem, anything outward can be simulated.

This interchange, however, would be missing the point. The important absence, in the case of the artificial speaker, is not some inner correlate to the spoken words. Rather, it is the idea of whatever makes a speaker's words *his*. Learning to speak is making language one's own. For someone's utterances to make sense is for it to make sense for him to make them. Words must have a place in our life. That is not something that could be supplied through a connection to some current inner state. It presupposes *having* a life, and that is precisely what the computer is lacking. Having a life is not like mastering a task.

The problem, generally speaking, lies in thinking that "making sense" can be identified with fulfilment of a condition that can be specified once we have settled on a correct theory of meaning.

Rhees was clear about this. His criticism of Wittgenstein's suggestion that the builders' game could be the entire language of two speakers is connected with this. He wrote:

When I say that [someone learning to speak] learns what it makes sense to say, I do not mean that he learns the correct way of using various expressions ... But I want to emphasize, in the first place, that he learns to *tell* you something... And this is not

---

<sup>5</sup> As John Searle has famously put a related point, even if a worker hidden in a room were successfully to mimic the behaviour of a Chinese speaker, he still would not understand what he was saying if he did not know Chinese.

something to which anyone can be trained, as an animal might be trained to react to an order... If he tries to tell you something, he is speaking to you; and this is connected with the ideas of addressing you and of greeting you...<sup>6</sup>

For words to come to have a place in a child's life is connected with the significance to her of the people she is addressing. Only if someone matters to you will you have something to say to her. This does not simply mean that you may have use for her. If I cannot reach the apples I may ask you to give them to me, but thus far *you* may not matter to me at all: a ladder might have done as well, or better. This is the situation illustrated by Wittgenstein's builders. When you speak to someone and are not just trying to get a certain thing done, you do so because the listener matters to you, maybe as a friend, lover, parent, or just as another human being who is ready to lend you her ear. *Because* the other matters to me, I may have something to say to her.

On the other hand, *why should* the computer address anyone? When the computer "speaks", there is no context of conversation that could lend sense to its utterances. The notion of having something to say has no application where computers are concerned. The baby whose body is convulsed with pleasure and surprise when her mother suddenly appears in her field of vision, even the cat mewling eagerly on returning from a bout with a flock of crows, are infinitely closer to having something to say than a computer, no matter or how many grandmasters it is able to beat at chess or how many tidbits of common sense it has acquired.<sup>7</sup>

---

<sup>6</sup> Rhees, "Wittgenstein's Builders – Recapitulation", in *Wittgenstein and the Possibility of Discourse*, p. 186.

<sup>7</sup> Consider the following passage from an article presenting the attempt to produce a computer that can simulate human thought:

Austin, Texas-based Cycorp Inc. claims to be "the leading supplier of formalized common sense". CEO and founder Doug Lenat has labored 17 years to codify facts such as "Once people die, they stop buying things"...

[Lenat explains] "We've put in 600 person-years of effort, and we've assembled a knowledge base containing 3 million rules of thumb that the average person knows about the world, plus about 300,000 terms or concepts. [For instance] terms like 'first date' and rules of thumb like 'People are more polite on their first date than they are on their *n*th date'. A lot of these things were true 50,000 years ago, like 'If you are carrying a container that's open on one side, you should carry it with the open end up'. The idea is to represent these in formal logic as opposed to English sentences.

You want the machine to be able to crank through the logical deductions—the consequences of these assertions—the same way you or I would....

A typical person knows about 100 million things about the world. I see us crossing that point in five years. It's difficult to predict the course thereafter.

(Gary H. Anthes, "Computerizing Common Sense", *Computerworld*, 8 April, 2002.)

Setting aside the question whether such a computer could be used as an aid in carrying out certain tasks, the idea that it might come to resemble a human speaker, and that the people building this data base are even remotely on the right track when it comes to reproducing human intelligence, is preposterous. One cannot help thinking that those involved in the project might have use for a bit of common sense themselves.

Learning to speak involves coming to do what others are doing, but this does not mean replicating their vocal behaviour. The relevant sense of “doing what others are doing” only has application in the context of a life in which things matter. Rhees repeatedly speaks about the “sort of sense one remark may have in connexion with another” (ibid. and passim), and I believe the connection he has in mind is to be understood as depending on the place of the remarks in the life shared by the speaker and his listeners.

Rhees’s remarks (although written with a different purpose in mind) go some way toward showing the nature of the confusion involved in thinking that the difficulties facing someone trying to produce an artificial speaker are only technical in nature.

What tempts us to accept such a line of thought, first of all, is too narrow a conception of what speaking is, one that focuses merely on what happens at the moment of utterance and the way it is related to its surroundings. What keeps us from questioning this conception, on the other hand, are probably certain features of our culture, e.g. the cognitive bias which puts thought and knowledge at the centre of human existence, and on the other hand, the instrumentalist bias which construes all forms of meaningful action as aimed at some external purpose. In addition to this, there may be a fear of looking dumb through questioning the limitless power of scientific and technological progress.

3. The second example I have in mind is the predicament of someone who is puzzled about the possibility of genuinely unselfish actions. Even when we act from concern for another, she may argue, we would not do it unless we thought the action would give us greater satisfaction than any other thing we might do.

Having taught introductory philosophy classes for nearly three decades, I have found it to be very easy indeed to elicit support for some such line of argument from beginning students, and very hard to make them give up that support. One way of trying to do that is to make the students aware that they are confusing two different ways of speaking about satisfaction. On the one hand, in deliberating on what to do, I may try to predict which alternative will *bring me the greatest satisfaction*. Having chosen one alternative, I may discover that my expectations were mistaken. In that sense, it turns out that I made the wrong choice. Here, satisfaction is independent of my decision. On the other hand, when I act with some goal in mind, my intention defines what will

*satisfy my endeavour*. This may or may not be something that I believe will give me a feeling of satisfaction. In fact, we often act in ways that we realize may give us little pleasure and a great deal of misery. Once we are aware of this distinction, this should help us realize that there are no logical limits to the kinds of goal a person may set for himself. Genuine unselfishness cannot be ruled out *a priori*.

However, even having this pointed out to us may not be sufficient by itself to relieve us of the problem. The case is not like one in which we are made aware of some *double entendre*, after which we simply laugh and move on. Perhaps it could even be said that the linguistic unclarity becomes a pretext we cling to in order to hold on to a line of thought that we do not have the energy or imagination to resist. A whole range of considerations discourage us from examining the problem too closely. The idea of genuine goodness poses a challenge that we may not wish to acknowledge. It threatens to expose our own motives in all their paltriness. Also, we fear that in showing our faith in the goodness of another, we expose our gullibility. As Kierkegaard noted:

It mortifies our vanity and our pride to have thought too well of a swindler, to have been foolish enough to believe him ... One is vexed at oneself, or at least one finds that it is ... “so stupid” to have been fooled... [H]ere in the world it is not “stupid” to believe evil of a good man; it is a superciliousness by which one adroitly gets rid of the good; but it is “stupid” to think well of an evil man, that is why one protects oneself – since one fears so much being in error.<sup>8</sup>

This does not mean that those who accept this line of reasoning will necessarily regard their fellow human beings in accordance with it. They are not particularly calculating or cynical, rather they represent the normal range of human strengths and shortcomings. Maybe they simply lack the courage to declare themselves for the good.

The denial of unselfishness is connected with other considerations as well. It has the ring of science: the idea that each one of us is ultimately looking after her own interests seems to mesh with our ideas of human beings as products of the struggle for survival.<sup>9</sup> Besides, a great deal of science is based *on* that idea; e.g. the economists’ notion of *homo oeconomicus*. This is bound up with the deeply entrenched idea that meaningful actions are always aimed at some goal.

---

<sup>8</sup> Søren Kierkegaard, *Works of Love* (Princeton, N.J.: Princeton University Press, 1946) trans. David F. and Lillian M. Swenson, p. 184. I have in some points deviated slightly from the Swenson translation.

<sup>9</sup> It should be obvious that the sociobiologists’ notion of altruism in the service of the gene offers a caricature of human goodness.



Again, questioning the possibility of genuine unselfishness is facilitated by the real uncertainty or indeterminacy characterizing particular cases – the ways in which hypocrisy or self-deception will distort our understanding of human motives. As Wittgenstein formulates it:

It is hard to understand yourself properly since something that you *might* be doing out of generosity and goodness is the same as you may be doing out of cowardice or indifference. To be sure, one may act in such & such a way from true love, but also from deceitfulness & from a cold heart too. Similarly not all moderation is goodness. And only if I could be submerged in religion might these doubts be silenced. For only religion could destroy vanity & penetrate every nook & cranny.<sup>10</sup>

Rhees addresses a similar difficulty in one of the texts published in *Moral Questions*, discussing the case of someone who is worried whether his refusal to lend another person money was truly concern for the other or just meanness. Rhees emphasizes that there are no facts that would settle the issue; it is a matter of coming to see the facts in the right light.<sup>11</sup> The fact that the issue ultimately requires taking a moral stance makes it easier to brush off the possibility of unselfishness: there is no danger, one may feel, of flying in the face of the facts. Facing a dilemma, not knowing which way to turn, we may settle for the “safe” alternative of simply shrugging our shoulders at the idea of genuine unselfishness.

4. Thus, it appears, the bewilderment of someone who is puzzled about the possibility of genuine unselfishness may be bound up with a complex variety of motives. All cases of philosophical confusion, however, do not have the same countenance. For instance, St Augustine’s famous puzzlement concerning the nature of time, or the confusion of someone who thinks the construction of an artificial speaker presents only technical problems, will not have the same kind of immediate bearing on how a person views her own life. A puzzle may put on different faces at different times; thus, it is a characteristic of the present time that in thinking about philosophical questions we feel the need to square them with what we take to be a scientific picture of the world. This, of course, was not always so. Again, what considerations carry weight will vary between individuals. The lesson to be drawn from this, if anything, is that we should abstain from generalizing about the ways in which philosophical confusion may arise.

---

<sup>10</sup> Ludwig Wittgenstein, *Culture and Value*, revised ed. (Oxford: Basil Blackwell, 1998), p. 54 (14.8.1946).

<sup>11</sup> Rhees, “Self-Deception and Mistakes”, in *Moral Questions* (ed. by D. Z. Phillips), Basingstoke: Macmillan, 1999, p. 242.

If philosophy is the disentangling of confusion, could it be called a form of therapy? Or should we say it involves learning about the possibility of discourse? We should be clear first and foremost that these are different metaphors which may have a use in reflecting on the nature of the philosopher's activity, and that everything depends on *how* they are used. Obviously, both metaphors have been badly misused and misunderstood, so perhaps we could say that each of them can be useful as a corrective to the other. (As for the idea that the confusions dealt with in philosophy are rooted in a lack of sensibility about human life, it is clearly too narrow to fit all cases.)

The reason one might want to resist the notion of philosophy as a form of therapy is that it seems to exclude the possibility that contending with philosophical confusions can lead to a genuine deepening of one's understanding of things. There is such a deepening both in the overcoming of particular confusions, and, more important, through the repeated experience of having to acknowledge that one's thinking is confused, and of learning about the complex moves required to find one's way out of those confusions. For one thing, in getting clear about the problem of an artificial speaker or of unselfish actions, for instance, we are helped to reflect on the richness and complexity of the life we live with language, or of the nuances of human motivation. In this way, becoming unconfused means getting in touch with our own thinking on particular issues (here one might see an analogue to the way in which psychoanalysis aims to bring us in touch with our own feelings). On the other hand, in learning to recognize that, whether I know it or not, my thinking on various issues is liable to be beset by confusion (indeed that, if the issue is complex enough, it is virtually certain to be), *I learn something about the difficulty of thinking seriously about things.*

To recognize that one's thinking is confused is to recognize that one is in a state of helplessness, that one is dependent on things beyond one's power of control. Being confused, in this way, is different from being misinformed or having made an error in calculation. If I suspect I am misinformed or have miscalculated, I usually have some idea how to go about trying to remedy it. This means that, if I try hard enough, I shall be able to get out of my error. But coming to believe I am confused means that I do not know how to go on. As was famously said, "A philosophical problem has the form, 'I don't know my way about'." In this way, the wish for clarity is like a hope for grace.

The fact that achieving clarity is dependent on things I cannot control, I would argue, is the most important lesson learnt by doing philosophy. In doing it I will come to recognize, time and again,

that when I thought the answer to some question was obvious, or that those arguing the other side of some issue were obviously wrong, this was an illusion due to the fact that my thinking was confused. In this way, philosophy teaches us a form of intellectual humility, one that brings to mind the idea of Socrates proving his wisdom by recognizing how little he knew.<sup>12</sup>

When Wittgenstein speaks of the results of philosophy as “bumps that the understanding has got by running its head up against the limits of language” (*Philosophical Investigations* § 119), then what he has in mind could be the humility taught by the experience of confusion. Rhees says he “never could get into the simile of ‘bumps and bruises’”<sup>13</sup>, but Rhee and Wittgenstein may not be so far apart here.<sup>14</sup> For, coming to realize that I may be confused is learning about the possibility of thought and discourse in the only way one *can* learn about it, i.e. through the bitter experience of trying to ignore the conditions of discourse.

Maybe the ideas of therapy and learning should not be thought to exclude one another. There is some truth and some falsity in each. What speaks in favour of the idea of therapy is the fact that each one has to go through the recovery for himself: confusions cannot be disentangled once and for all. What speaks against it, on the other hand, is the fact that someone who has gone through the steps of being confused and recovering from confusion may feel that she is better off than before the sickness took hold.

---

<sup>12</sup> Could someone understand what philosophy was about if she was never confused? While I don't think imaginary questions of this sort can really be seriously asked, my inclination is to say that philosophy would be beyond this person's grasp; maybe in the same way that morality would be beyond the grasp of angels.

<sup>13</sup> Rhee, “The Fundamental Problems of Philosophy”, p. 585.

<sup>14</sup> On the other hand, Wittgenstein seems to be exaggerating when he asks, sarcastically, in *Culture and Value* (p. 98; 1950): “Philosophy hasn't made any progress? – If someone scratches where it itches, do we have to see progress?”