

# Human Beings and Automatons

Simo Säätelä, Bergen

1 J.S. Mill has formulated a classical statement of the “argument from analogy” concerning knowledge of other minds: “I must either believe them [other human beings] to be alive, or to be automatons” (Mill 1872, 244). It is possible that Wittgenstein had this in mind when writing the following:

“I believe he is suffering.”—Do I also *believe* that he isn’t an automaton?

It would go against the grain to use the word in both connexions. (Or is it like this: I believe he is suffering, but am certain the he is not an automaton? Nonsense!)

Suppose I say of a friend: “He isn’t an automaton”.—What information is conveyed by this, and to whom would it be information? To a *human being* who meets him in ordinary circumstances? What information *could* it give him? (At the very most that this man always behaves like a human being, and not occasionally like a machine.)

“I believe he is not an automaton”, just like that, so far makes no sense.

My attitude towards him is an attitude towards a soul [*eine Einstellung zur Seele*]. I am not of the *opinion* that he has a soul. (*PI* p. 178)

Here Wittgenstein contrasts opinion (*Meinung*) and attitude (*Einstellung*). How should this contrast be understood? On a view such as Mill’s, to regard someone as a conscious being is to hold certain beliefs about him, beliefs that can perhaps ultimately be grounded in a theory of some sort. To have an “attitude towards a soul” is, on the contrary, to see a person’s gestures and facial expressions as “filled with meaning”. We have an *attitude* towards a soul when confronted with a person, which means that we react to his presence and behaviour in a certain way.

There is a similar remark in the *Nachlass* which develops the contrast between *Einstellung* and *Meinung*, but also involves religious belief:

Instead of “attitude toward the soul” one could also say “attitude toward a human”.

I could always say of a human that he is an automaton (I could learn it this way in school in physiology) and yet it would not influence my attitude toward someone else. After all, I can also say it about myself.

But what is the difference between an attitude and an opinion? I would like to say: the attitude comes before the opinion. (*Isn’t* belief in God an attitude?)

How would this be: only one who can utter it as information *believes* it.

An opinion can be wrong. But what would an error [*Irrtum*] look like here? (*LW II* 38)

By introducing this contrast between opinion and attitude Wittgenstein wants to make at least the following points:

(i) Opinions are often expressed merely by words, while attitudes are not necessarily verbalized or even verbalizable; rather they *show themselves* in the actions and reactions of people, in how I treat the object of my attitude, in “fine shades of behaviour” (*PI* p. 204). It means that we *directly* see actions and expressions of emotions; it is not a question of first

observing “bodily movements” and then interpreting them as meaningful (which will always leave room for the possibility that we are dealing with automatons, not human beings). To put it in a more technical vocabulary we could say that an attitude towards a soul is not a psychological state but rather “a condition of the sense of the ascription of mental predicates” (Gaita 1991, 189), meaning that such an attitude is a condition of behaviour to be seen as the behaviour of a conscious being.

(ii) The difference between opinions and attitudes is perhaps best shown by the fact that changing an attitude involves changing the person with that attitude (his way of reacting, his way of understanding a certain thing, etc.), whereas changing an opinion need involve nothing more than changing just that, the opinion.

(iii) This also means that we can speak about an *internal relation* between the attitude of a person and the object of his attitude (Winch 1987, 148), and between an *Einstellung* and the reactions indicative of it. In other words, the attitude is constitutive of the object. Opinions only have an external relation to their objects; i.e., our understanding of what the object *is* is independent of our opinions about it (the object is what it is regardless of our opinions about it). In contrast to this, we can say that an attitude is something that is manifested in our life, which means that our reactions indicate and presuppose a certain attitude towards the object of the reaction.

Thus, to regard someone as a human being is not to believe that he is *not* an automaton; instead, we have to do with “an attitude towards the soul”. It is important to keep in mind when discussing Wittgenstein’s notion of an attitude that it is not a feature of an experience discoverable by an empirical investigation. That is, it is not a psychological stance or special mode of perception, but rather something that characterizes the grammar of the concept of a person.

2 On the basis of what we have said so far, Wittgenstein seems to deny the very possibility of regarding other human beings as automatons. He did, however, sometimes speak or write in a way that contradicts this. What are we to make of such passages? There are at least three different cases to be considered:

(i) In the *Nachlass* passage quoted earlier he introduces the puzzling possibility of actually saying of a human being that he is an automaton. He writes (seemingly in contrast with the example in the *Investigations*) that “I could always say of a human that he is an automaton (I could learn it this way in school in physiology) and yet it would not influence my attitude toward someone else”. That is, if the teaching in physiology changes (e.g., as the result of new scientific discoveries) this might influence my *opinions* about human beings (including myself), but would not change my attitude toward a human being. To be more precise, what we seem to have to do with here are facts and beliefs about the workings of the human body. It might become my opinion that the human beings can be thought of as automatons in this sense without this fact influencing my *attitude* towards people.

(ii) Wittgenstein also introduces a possibility of seeing a human being (and not just the body) as an automaton. This is a case of aspect perception: to see a human being as an automaton is possible, but Wittgenstein describes this as a kind of "limiting case". Such a possibility is derivative upon the fact that we do *not* normally regard human beings as automatons, that our *basic* attitude towards them is an attitude towards a soul, which is *not* a case of aspect perception. This allows us to (imaginatively) see human beings as automatons. However, it is impossible to hold on to this kind of aspect perception in our day-to-day dealings with others:

But can't I imagine that the people around me are automatons, lack consciousness, even though they behave in the same way as usual?—If I imagine it now—alone in my room—I see people with fixed looks (as in a trance) going about their business—the idea is perhaps a bit uncanny. But just try to keep hold of this idea in the midst of your ordinary intercourse with others, in the street, say! Say to yourself, for example: "The children over there are mere automatons; all their liveliness is mere automatism." And you will either find these words becoming quite meaningless, or you will produce in yourself some kind of uncanny feeling, or something of the sort.

Seeing a living human being as an automaton is analogous to seeing one figure as a limiting case of another; the cross-pieces of a window as a swastika, for example. (PI §420)

The point here is that we have to do with a limiting case, a possibility of aspect perception that is derivative upon something more basic, in this case, our attitude toward human beings.

(iii) Norman Malcolm does, however, report that Wittgenstein (in lectures) attempted to give sense to the idea of *consistently* regarding other people as automatons. He introduced the example of "a tribe of people who had the idea that their slaves had no feelings, no souls—that they were automatons—despite the fact that the slaves had human bodies, behaved like their masters, and even spoke the same language". The masters would observe the slaves as if they were machines, and "if a slave received a mortal injury and twisted and screamed in agony, no master would avert his gaze in horror [...] any more than he would if the ceiling fell on a printing press". Malcolm concludes that "here is a difference in 'attitude' that is not matter of believing or expecting different facts" (Malcolm 1954, 548–549). If this really is a difference in attitude, then it would be impossible for us to say that the slave-owners have got something *wrong*, since it is out of order to talk about errors when attitudes are concerned, *pace* Wittgenstein's example involving belief in God. How could a belief in God be shown to be wrong, if it really is an *Einstellung*? Of course it is possible that such an attitude changes; one might fall out of faith, say. But this does not make one's former belief an *Irrtum*; there are no independent facts to be mistaken about (thus proofs of the existence or non-existence of God are quite irrelevant here). So what about the case of the slave-owners? We would like to say that they are terribly wrong about their slaves, but here, too, it seems that it is out of order to speak of an error. What is lacking in the case of the slave-owners is not more *information*. It is not as if they would have got the *facts* wrong. But there *is* something lacking. When discussing sensations, Wittgenstein writes: "Our attitude to what is alive and to what is dead, is not the same. All our reactions are different" (PI §284). Indeed, in the case of the slave-owners, we can say that their reactions are so utterly different from ours that it is questionable whether we would

like to describe *them* as human beings whom we could criticize from a moral point of view. It is only within a web of human relationships that an attitude towards the soul can be revealed (Winch 1987, 143), and what is lacking in the example of the slave-owners is precisely such a web, or form of life. Thus we can say that it is questionable whether Wittgenstein seriously entertained the possibility of there being such a difference in attitudes towards human beings. As Cockburn notes, if we think that this case is clear then we just fail to appreciate how extraordinary it is. Indeed, such a tribe would be so completely alien to us that it is unclear that we could even describe the situation in terms of judgements these slave-owners do make about their slaves (Cockburn 1990, 48-50). I would like to claim that Wittgenstein here, too, describes a kind of "limiting case". The very difficulty or even impossibility to fill in the details of this case and still keep it intelligible shows us something important about the limits of the concept of a person.

3 But does not all this boil down to a kind of idealist position? Isn't Wittgenstein saying that the attitude decides what things are? It is important to realize that we are not talking about aspects here, or "emergent properties". It is not the case that if I take a certain attitude towards a body, the mental properties somehow emerge or dawn as an aspect. Indeed, it is question-begging to speak about "taking an attitude" in these contexts; an attitude is rather something I *find* myself in, and something that I am unable to abandon at will. It has something to do with how we see and treat things; the light in which we understand them; the perspective in which things make themselves perceived by us. To have an attitude towards a soul means that I react to a person's words and behaviour in a way that indicates that he has a soul; I am not of this *opinion*, i.e., I do not interpret his movements and behaviour and then infer that he is indeed a self-conscious being. But this does *not* make the position an "idealist" one. The attitude does not turn a body into a human being; nor is it up to us to simply choose whether to have such an attitude or not. "Our attitude to what is alive and to what is dead, is not the same. All our reactions are different" means that I am not, e.g., of the opinion that a fly is alive and a stone is not: this is something that shows itself in the way I react to the object in question, how I treat it and what I am prepared to say and do about it. But it is not my *attitude* in itself that makes the fly a living creature! Similarly, it is not my "attitude towards a soul" that makes something a person or a human being. But to say that *eine Einstellung zur Seele* is basic to our concept of a person is a way of showing how this concept is dependent upon the web of human relationships within which such attitudes are meaningful and possible.

## References

- Cockburn, D. 1990 *Other Human Beings*, London: MacMillan.  
 Gaita, R. 1991 *Good and Evil*, London: MacMillan.  
 Malcolm, N. 1954 "Wittgenstein's *Philosophical Investigations*", *Philosophical Review*, 43, 530–559.  
 Mill, J.S. 1872 *Examination of Sir William Hamilton's Philosophy*, London: Longmans, Green, Reader, and Dyer.  
 Winch, P. 1987 *Trying to Make Sense*, Oxford: Blackwell.  
 Wittgenstein, L. 1958 *Philosophical Investigations*, Oxford: Blackwell. (PI)  
 Wittgenstein, L. 1992 *Last Writings on the Philosophy of Psychology*, Vol. II. Oxford: Blackwell. (LW)