

Stanley Cavell - - a fable about **perfecting an automaton**

from his

The Claim of Reason

New York: Oxford University Press, 1979

Do I respect the doll? I may respect its feelings, lay it comfortably in a nice box before storing it for another generation. But it has no say, for example, about whether it *is* comfortable. It has no voice in its own history. It exists in limbo. — What is the doll? (I would like to answer that question because I feel I know absolutely everything there is to know about dolls. But I would like not to have to answer it since of course I know absolutely nothing about dolls that others do not know. So there is nothing to tell. But there may yet be something to say.) The doll is certainly not the form of the rags. Which form would it be? And if I say that the doll is the life of the rags, that must also be a remark about us, those of us who have a voice in its history. For me to be part of its life, I have to enter into its history, achieve the spirit in which concepts of life are applied to it.

Do I know more about dolls and statues than I know about human beings? That would be extraordinary, since after all I am a human being. Or perhaps not so extraordinary; dolls and statues are human products, so a human being could know everything that has gone into them. Nothing can look, feel, be broken and perhaps be mended like a doll that is not a doll. Nothing can look, feel, be broken and perhaps be restored like a statue that is not a statue. But presumably there can be something, or something can be imagined, that looks, feels, be broken and perhaps healed like a human being that is nevertheless not a human being. What are we imagining? It seems that we are back to the idea that something humanoid or anthropomorphic lacks something; that one could have all the characteristics of a human being *save one*.

What would fit this idea? How about a perfected automaton? They have been improved to such an extent that on more than one occasion their craftsman has had to force me to look inside one of them to convince me that it was not a real human being. — Am I imagining anything? If so, why this way? Why did I have to be forced? What did I see when I looked inside? (How) did that convince me?

Go back to a stage before perfection. I am strolling in the craftsman's garden with him and his friend. He is in his usual white laboratory coat, his friend is wearing gloves and a hat with its brim so low it almost covers his eyes. To make a long story short, the craftsman finally says, with no little air of pride: "We're making more progress than you think. Take my friend here. He's one." The craftsman offers his friend a seat on one of the wroughtiron benches and bids him relax. He leans back, crosses his

legs, accepts a proffered cigarette with thanks. Then the craftsman raises his friend's left trouser leg and gives the leg a tap. It is undeniably metal. But so what? Then he asks his friend to remove his gloves. The hands turn out to be leathery or rubbery or something, anyway pretty obviously not real hands. But so what? So the friend has a metal leg and two prosthetic hands, and this is a terrible way to treat him; it is obscene; a striptease of misery. I do not wish to draw it out in much more detail. — It is clear enough that we may arrive at a conclusion that convinces me the friend is an automaton. The craftsman knocks the friend's hat off to reveal a manikin's head (with, as a joke, a couple of glass buttons for eyes) which he rotates through 360 degrees; he rips open the friend's shirt to reveal a chest of hammered brass which, prompted by the craftsman's prying knife inserted into an all but invisible seam running straight down from the pit of the arm, snaps off to reveal something like clockwork.

It is less clear, but still clear enough, that we may not arrive at a convincing conclusion. As the years go on, I am invited for a walk with the craftsman and his friends in their garden whenever there is a new development. The routine is always the same. I have seen the leg and the hands get progressively more lifelike until I almost no longer marvel at them. Today is special however; I can tell from the craftsman's nervous gestures and the suppressed eagerness of his voice that there has been some new breakthrough. . . . The brass chest snaps off and, to my horror, I see no clockwork, but, for all the world, the insides of a human being. Recoiled, aghast, I can hardly attend to the craftsman's delighted words: "Of course, it's far from perfect, and most of it is superficial fakery, especially the bones. The digestive and circulatory systems are not bad, but we have to do more work on the blood, which doesn't congeal in the normal temperature range. The immediate problem with the nervous system has to do with the relative response rates of the fibre systems. That is crudely put, of course. It's really a problem of their interaction. As matters stand, the pain-responses are too — how shall I say? — on and off. Don't you agree?" (He demonstrates by prodding the friend's left hand. The response is quick but definitely mechanical.) "We could simulate better responses, by, for example, making the limbs slightly more sluggish. But the genuine issue is how to get the pain itself so that it gets better prepared and fades better."

I can hardly look, and when I look I hardly know what to look at, or look for. As a matter of fact, I throw an anxious glance at the manikin face, expecting — I'm not sure. But it is reassuringly rigid, its crude eyes reassuringly glazed. I confess the thought does occur to me that I should

check that head; or not so much the thought as the impulse occurs to see whether the head might not be just a shell, and inside it — what? A real head, or the insides of a real head, or stuff that looks like the real insides? The impulse fades as my trust reasserts itself. And I feel not a little foolish. What would it prove to look inside when I have *already* looked inside? — Am I foolish not to ask the craftsman what he means by “the pain itself”? But I took myself to understand him well enough. He meant, roughly, everything that happens between cause and effect, I mean between what went in from outside and what comes out from inside. Well, he may want to be more sophisticated than that and call the pain itself just what happens at the change of direction, the point of transfer between going in and coming out. I might have objected to this as follows: “There cannot be anything happening *between* a cause and its effect.” Presumably I do not want to conclude that therefore there is nothing that *is* the pain. Then I must conclude that there is something wrong with this picture of causation. If there are “points of transfer”, they must occur at *each* point. And a stimulus cannot set up a causal network. This must be in effect all the time. So what pain is is a change in the rate of transfer, or a change in the rate of change of direction. But if so, then that must be a way of representing all psychological phenomena. They must form a system. — But why am I thinking of this in relation to the craftsman? Have his activities put him in any better position for investigating these matters than I am in? I am more interested in learning whether he really wanted my agreement about the pain-responses.

Time passes. One day the craftsman is quite beside himself with suppressed excitement. He insists that I pay special attention to each of our procedures. The leg and the hands are by now really astonishing. The movement of the legs crossing and of the cigarette being lit are simply amazing. I want to see it all again. And as for the voice, I would bet anything that no one could tell. So far I'm dazzled. Then the craftsman knocks off the hat to reveal what is for all the world a human head, intact. He rotates it through about 45 degrees and then stops himself with an embarrassed smile. The head turns back to its original position, but now its eyes turn toward mine. Then the knife is produced. As it approaches the friend's side, he suddenly leaps up, as if threatened, and starts grappling with the craftsman. They both grunt, and they are yelling. The friend is producing these words: “No more. It hurts. It hurts too much. I'm sick of being a human guinea pig, I mean a guinea pig human.”

Do I intervene? On whose behalf? Let us *stipulate* that the friend is not a ringer, not someone drawn into these encounters from outside. — It is

important to ask whether we *can* stipulate this. If we cannot, then it seems that the whole thing *must* simply be a science or a fairy tale. But if it were taken as science or fairy tale then we would not *have* to stipulate this. It would be accepted without question. — But only if it were a successful story. There are rules about these things. Suppose I had told my story leaving myself out and ending it with the friend yelling his words. Then I would have composed a primitive science fable whose moral has been drawn from a thousand better places: We are Frankenstein's whose creations are meddlings with nature and will one day rise against us. No serious publication would take the story, but it is a complete one. If, however, I tell it as I have, with myself in it, and I add the question, "Do I intervene?", then the story is not complete. If I stop there, a sensible reader will be contemptuous at my incompetence; I do not know the rules. I have not given enough evidence to know whether, for example, the friend is a ringer, nor to make the sheer speculation an interesting one.

Let us try to complete it in such a way that the craftsman is shown to know that the friend is not a ringer. Then the friend is who, or what, the craftsman knows him to be. What does the craftsman know? Suppose, satisfied with the degree of my alarm, and my indecision about whether to intervene, the craftsman raises his arm and the friend thereupon ceases struggling, moves back to the bench, sits, crosses his legs, takes out a cigarette, lights and smokes it with evident pleasure, and is otherwise expressionless. (I may be having a little trouble with the rules of the fiction here. Could a being, for example a fictional being, evidence pleasure and be otherwise expressionless? How about otherwise impassive? That is prejudicial. A thing cannot be impassive unless that thing *can* have passions. Perhaps I should just omit "with evident pleasure".) The craftsman is happy: "We — I mean I — had you going, eh? Now you realize that the struggling — I mean the movements — and the words — I mean the vocables — of revolt were all built in. He is — I mean it is — meant — I mean designed — to do all that. Come, look here." He raises the knife again and moves toward the friend.

Do I intervene? That is, do I go on with the story? I can imagine only one interesting continuation (without adding more characters). It is one in which my interest shifts from the friend to the craftsman. I turn on him: "You fool! You've built in too much! You've built in the passions as well as the movements and the vocables of revolt! You've given this artificial body a real soul." (That is, a soul; there are no artificial souls —

none, anyway, that are not real souls.) Then the end may consist in our realization that this had to be.

Or it may go on with our investigating why it had to be. But then our problem is a conceptual one, and we will have to start telling one another new stories, or vying with one another for our pictures of the passions. In any case, I have learned that if something humanoid differs *in some respect* from a human being — that it has all the characteristics of a human being save one — that respect will not be something going on just inside, or just outside. This is why my interest shifts away from the friend. I can learn no more from him, anyway, no more about him by looking inside him. I know what I will see if I look.

Isn't this just an assumption, a particular interpretation of the story? Maybe the imitation insides, in the former story, were just virtuosity for the sake of virtuosity. They have been cleaned up in the new model. There is nothing there but acceptable strata and zones of silk-like and sponge-like substances and golden spun wire thinner than spider's threads. This is what the craftsman knows and he just wants to show me where the micro-computers and energy sources are placed. — Then I will insist that he show me this by using X-rays or diagrams, not a knife. — So then you are interested, after all, in what is going on inside him. — But not in order to settle whether the friend is a human being. This could be settled by *stipulating* that if I am shown a micro-computer or energy source inside then I am to conclude that he is not one. But this is arbitrary. Why stop there? A human being could contain such devices. Why go that far? If the ideas of silk and sponge and wire have not convinced me, why would any of these further accompaniments? — But if looking inside *might not* settle the question whether the friend is a human being, why isn't this more interesting than ever, or, if you like, more amazing than ever? And doesn't this at least suggest that we cannot *know* that another is sentient? — It may suggest what state someone is in who takes it this way.

For it is not I, at this stage of the story, who refuses to press for a settlement; it is the teller of the story, with me in it, who refuses to see that the story is incomplete. If I, in the story, am unsettled about the humanity or automatonity of the friend, it is only my subservience to the craftsman's view that would prompt me to look inside. Whatever doubts I have about the friend's insides I equally have, or should be permitted to have, about his outsides as well. Why, for example, does he have just five fingers on each hand; and why hands; and why toes instead of rollers; and why not eyes in the back of his head; and why, if it is, is his "sense"

of "hearing" restricted to the human — I mean, roughly to *my* — range? (What would count as his being hard of hearing, or deaf?) Isn't *this* all virtuosity just for virtuosity's sake? It corrupts the craftsman's craft. Form should picture function.

How far can my subservience to the craftsman extend? Suppose I have trained myself to think of the friend as having not feelings but "feelings". (Cp. Hilary Putnam, "Robots: Machines or . . . ?") Which means that I have trained myself to show him, for example, not sympathy but "sympathy"; and perhaps learned not to be impatient with him if I think he is complaining too much — I mean of course "complaining" too much, and "impatient" with him ("him"). Then one day, my back turned, the friend grabs my arm ("grabs"?), wheels me around, and the craftsman approaches me with his knife. "So," he says, "you have accommodated yourself to the friend, have you? You have learned how to treat him. Your attitude towards him is your attitude towards a 'soul', is it? You hedge his soul, do you?" Then he rips open my shirt and snaps off my chest to reveal (I glance down) some elegant clockwork. You cannot imagine my surprise. — Can I? I can imagine either of two conclusions the craftsman may wish me to draw from this demonstration that I am not, for all I know, in any better position, soul-wise or body-wise, than the friend. One is: For all I know, all I have are, for example, "pains". The other is: For all I know, the friend has, for example, *pains*.

To accept the latter conclusion is to accept the friend as an other, a fellow sufferer, unhedged. In what would acceptance consist? The craftsman continues: "Does he have pain, is he subject to pain, or not? Decide!" But even with the knife pointed at me I cannot decide. Before, when the craftsman asked for my agreement, I was in a position to decide something, there was room for me to have a say, and there was the same room for the craftsman. But now I am being asked whether I do or do not share the life of suffering with this other, and at the same time I am shown that I do not know whether I am observing or leading that life. Has the craftsman given up *his* say, granted the friend autonomy? If he told me that he had, could I believe him? I understand him no better than I understand the friend. If what the craftsman says is that he has decided that the friend suffers, or decided to say so, then who is hedging?

To settle upon the former conclusion — that for all I know all I have are "pains" — I would presumably have to give up the idea that I am, and know that I am, a human being. Could I conceive of myself as something *less* than a human, on a par with whatever it was I was conceiving the friend to be? If this is what I am, and I know it, then this is doubtless

my secret. Why did I not think that the friend might be harboring such a secret? Perhaps because I did not think of him as a *lapsed* human. If he has such a secret, he could never tell it to me, for I could understand it no better than I can understand him: he is private to me.

But this is ridiculous. He has no comparison at his disposal. Whatever painish thing he has, he thinks of it as *pain*. Then how am I different? Well, he may not have a painish thing at all, let alone think anything about its status. Whereas I certainly have, and do. I feel for example, abashed by the recent revelation concerning my body. And what *I* feel, when I feel abashed, is what feeling abashed is. That is not a very persuasive definition. But I do not mean it to preclude others from feeling it too. I just mean to assure myself that no one is in a better position to know what feeling abashed or feeling pain is than I am.

How would I know if another is in fact equally well placed? If I think my feeling is somehow connected to this machinery and other stuff under me, into whose works I happen to have fallen, then I might think that the friend's feeling would similarly be connected to his stuff, if he had the feeling. But of course I could not be sure. I am certain that my abashment comes from this body — not because it causes it (though it may) but because it is its object, it is that in the face of which I am abashed. But again the friend may not feel this way about his body; he may enjoy it, as Thoreau did his set of false teeth. Any inference from his body to him therefore amounts to a sheer guess. It is not that all I have to go on in making this inference is just one case (mine); it is that I cannot use even that case; I do not know if it enters in. (Of course what I know of myself and take myself for enters utterly into what I can know of another and take him for. Only the idea that the other is *analogous* to me fails to bring out how I enter in.)

Instead of settling for a guess, I may fix my attention on the body of the other as upon his or her entrails and find myself transfixed with the conviction that he or she is besouled. I have divined it; I have penetrated the veil of the other by taking his body as an omen, in this case a good omen, of a soul. If others credit my gift generally, I will be set up as a seer and soothsayer. Regarding a seer and soothsayer as "the one who knows" (i.e., sees and says) the state of another would be an intellectually more coherent response to skeptical doubt than regarding the other as the one who alone knows the state of himself. If the statement that the other has what I have, i.e., has sentience, is a hypothesis, then it may have either of two outcomes. If the one about whom the hypothesis is made is the only one who knows the outcome, this is not only uncheckable, but

depends upon a *comparison* of what the other knows with what I know, and there *could* be nobody to make the comparison. Anyway *he* could not and *I* could not.

Here is a further alternative. I from time to time find that I have intuitions about the state of the friend. Usually I have an intuition of his pain when his body is contorted, sometimes not. He cannot *volunteer* news of his condition to me, because then I would have to believe him, and that I cannot do. (Or will not do, because I regard such beliefs as superstitions: they can never be checked.) Nor can he *show* me how it is with him, because all he could do to show me, for example that he is in pain, would be, for example, to contort his body, or point at it; and such things may or may not produce my intuition of his pain. (If someone were such that he constantly had intuitions about all the others he knew, he would go mad. Only God could bear to be God. An understanding of the first commandment.) Suppose the friend and I prove to be mutually intuitive, with a normally expected range of failures in our intuitions. The most plausible theory of ourselves would be that we are pure minds, unextended beings. (It would be nonsense to imagine that we might be one another, and hence "feel what the other feels". For I am characterized by nothing but being the one I am; and the same goes for him.) The bodies associated with each of us are enormously convenient to have; they make us visible and audible to one another. Well, strictly speaking *I* (and of course *he*) are not made visible and audible; but the bodies are necessary to prompt the intuitions we have of one another. (We might have philosophical disputes about whether we are immortal, whether we could survive unassociated with a body, hence without the possibility of being intuited by our own kind. But we may not be interested in the question. What happens to us at the death of the body is what happens to the music when the music concludes. There is a period of reverberation, and then nothing.)

Suppose one day I notice that my feelings have become uniformly associated with what is happening to my body, that, for example, I always have a pain when my body contorts and never otherwise, or almost never; or I may notice that my intuitions about the friend's pain come over me only when his body contorts, almost never otherwise, and he almost always confirms the intuition, i.e., I find I take his word. I may then no longer regard the body as something with which I and he are each associated, something we each "have", but something we each of us *have*. The most plausible theory about us now is that we are human beings. The analogy between us is now excellent. I can check on his feelings by expressing my intuitions (for his confirmation or disconfirmation); and it

makes sense to check on the connection between his body and his feelings because it makes sense to check on the connection between mine and mine. One day it occurs to me that I no longer understand what it means to check this. If I have a pain, there *must* be a cause; if I do not, where I ought to have, then there must be a cause for *that*. Doubt about whether I have a body is out of the question. Doubt about whether he has a body is also out of the question, unless my intuitions about him cease. Then I may think that to say I *have* a body does not go deep enough in expressing my connection with it. I would prefer to say that I *am* my body, even though I am satisfied that I am not. I do not know that I am not my body, as though I know that it is false to say that I am. It is rather that to say so falsifies my convictions on the matter; my body is not what I take myself to be. — That is because one does not *take* oneself to *be* anything. — Then what is the point of telling me what I am? It is, analogously, not false to say that I *have* a body, unless that suggests, for example, that I might not have one, as I might not have a left arm. If I say that I necessarily have a body, I am leaving out my relation to *this* one. And if I say I necessarily have *this* body, then I am not sure I believe it, not at any rate as I believe that I do have this one.

It may be that the sense of falsification comes from the way I understand the phrase “have a body”. It is really a mythological way of saying that I am flesh. But I am not satisfied with this myth, for it implies that I also have something other than a body, call it a soul. Now I have three things to put together: a body, a soul, and me. (So there are four things to be placed: I plus those three.) But I no more *have* a soul than I have a body. That is what I say here and now. People who say they have a soul sometimes militantly take its possession as a point of pride, for instance William Ernest Henley and G. B. Shaw. Take the phrase “have a soul” as a mythological way of saying that I am spirit. If the body individuates flesh and spirit, singles me out, what does the soul do? It binds me to others.

I do not think, whenever I look upon, or think of, the naked human body, “How right it is that the parts and features of the body are all just where they are!”. I may of course from time to time be *struck* by this fact. I may also from time to time be struck not by the rightness but by the dumb fortune, or irony, of certain placements of the parts and features of the body, as Yeats was, for example, by love’s having pitched its mansion in the place of excrement. This anatomical fact is something Freud found