

# Myself and I

John Perry

October 1997

[In Marcelo Stamm, editor, *Philosophie in Synthetischer Absicht* (A festschrift for Dieter Heinrich), Stuttgart: Klett-Cotta, 1998, pp. 83-103.]

## 1 Introduction

In this essay I distinguish three kinds of *self-knowledge*. I call these three kinds *agent-relative knowledge*, *self-attached knowledge* and *knowledge of the person one happens to be*. These aspects of self-knowledge differ in how the knower or agent is represented. Most of what I say will be applicable to beliefs as well as knowledge, and to other kinds of attitudes and thoughts, such as desire, as well.<sup>1</sup>

Agent-relative knowledge is knowledge from the perspective of a particular agent. To have this sort of knowledge, the agent need not have an idea of self, or a notion of himself or herself. This sort of knowledge can be expressed by a simple sentence containing a demonstrative for a place or object, and without any term referring to the speaker. For example, “There is an apple” or “that is a toaster”.

(Ideas of specific objects I call *notions*. Ideas of properties and relations I just call *ideas*. A judgement involves an idea being associated with a notion. A notion together with all of the ideas associated with it is a *file*.)

In self-attached knowledge, the agent has an idea of self, which is associated with a notion, which I call a *self-notion*. This is the kind of knowledge that expressed with the word “I”—what Shoemaker calls “first-person knowledge.” For example, “I am a philosopher,” or “I see a toaster” or “I have a headache”. In the last section of this paper, I try to explain why that is an apt name—why it is that the word “I” is so intimately connected with the expression of this sort of “self-thought”.

In knowledge of the person one happens to be, the agent is represented to herself in just the same way that other people are represented to her. The agent

---

<sup>1</sup>The first seven sections of this paper present are intended to present basically the same account that was developed in [6], from which it borrows some examples and prose. I believe this essay is clearer about the structure of agent-relative knowledge and the relations between the three kinds of knowledge, and connects with the ideas developed in [7]. ==

has just the same kind of idea of herself as she has of other people. This kind of knowledge can be expressed with a name or third person demonstrative. For example, “John Perry is a philosopher” or, pointing to myself in a mirror, “That man is a shabby pedagogue.”

## 2 Agent-relative knowledge

### 2.1 Two facts about the human condition

Everything we learn about other objects we learn by employing methods that are appropriate because those objects stand in certain relations to us. And however remote from us the object we are ultimately learning about may be, our inquiry will involve detecting the properties of things in our immediate vicinity.

I may learn that Bill Clinton visited the Bay Area by reading the paper, or seeing pictures on television, or hearing the radio. I learn about Clinton by reading what is written on the paper, seeing what is portrayed on the television screen, or hearing the sounds coming from the radio. Being the radio I am listening to, or the television I am watching, or the newspaper I am reading, are all what I call *agent-relative roles*: roles that other individuals play in the lives of agents. These are agent-relative roles, because an object plays or doesn't play such a role relative to a given agent, at a given time. For example, my computer is playing the role of *object in front of me* right now, relative to me, but not relative to you.

When I read about him, or watched him on TV, or listened to a report of what he was doing on the radio, Clinton was also playing an agent- relative role in my life, one that was derived from the role these other objects were playing, and his relation to them. For example, Clinton was the object read about, because the newspaper was the object read, and Clinton was the object the newspaper story was about.

This is the first of two very general facts I want to emphasize: any object we learn about, plays some agent-relative role, basic or derived, in our life. We learn about the object by using an epistemic method connected to the role, a way of finding out about the object or person playing that role. The way to find out about the object in front of you is to look at it, or perhaps to walk up to it and touch it. The way to find out about the object that the document in front of you is about is to read the document. Finding out about the objects around us is a way of finding out about other objects, given general facts about the way things work and specific facts about how things are related. If Clinton is the source of the image on my television, then I can find out things about him by finding out things about that image.

The second fact is that however complex our lives are, everything we do comes down to performing operations on the objects around us—objects in front of us, behind us, above us; objects we are holding; objects we can see.

By doing these things, we do things to objects in less basic relations to us. By speaking into the phone I hold, I speak to the person I called, the person to whom the signals that pass through the phone I hold are ultimately directed. I know how to move my body so as to effect objects around me, and I know how effecting those objects will effect other objects related to them in certain ways.

There are then two kinds of methods connected with agent-relative roles, epistemic methods and pragmatic methods. These two kinds of methods are the key to all human intelligence and purposive activity. We know how to find out what kinds of objects occupy these roles, and we know how to perform various operations on them. Technology extends the methods, so that we can find out about things in more and more complex relations to us, and do things that will change them in predictable ways.

Our practical knowledge then, the knowledge that enables us to do things, forms a structure at whose base is information about the objects that play relatively basic agent- relative roles in our lives.

## 2.2 Knowledge concerning the self

Consider a simple successful transaction involving such a basic agent- relative role, and epistemic and pragmatic methods associated with it. I am hungry. I see an apple before me. I pick it up and eat it. The complex movement of arm, hand, fingers, neck and jaw was successful in getting the apple into my mouth, because of the distance and direction the apple was from *me*. What I learned from perception, then, must have been the distance and direction of the apple from *me*. Or consider a transaction with a fax machine. To press certain buttons on it, I have to move my fingers a certain distance and direction from *me*. It isn't enough to know where the buttons were relative to one another, or where the fax machine was in the building or the room. I had to know where these things were relative to *me*.

It seems then, that these basic methods already require me to have some notion of myself. For it seems I need to know who it is, for example, from whom the apple is a certain distance and direction. If it is that distance and direction from you, or President Clinton, then moving the way I did would not be a way of eating it.

However, I think this is misleading. A natural way for me to report what I saw would be to simply say:

That's an apple

or

There is an apple there

There is nothing in this remark that refers to me. And after all, why should there be? I didn't see myself, I saw an apple. But, one might reply, we saw above

that I got the information that the apple was a certain distance and direction from me. Otherwise, how did I know that I could reach it?

When we perceive, we learn how things are around us. But that remark is a bit ambiguous. Suppose I say that when you look at an accurate clock, you learn what time it is in the time zone you are in. That's true. But is it true if said of a child, who doesn't know what time zones are? It's still true in a way. The child learns what time it is in the time zone she is in, as opposed to learning what time it is in some other time zone. But there is nothing in her thinking that reflects that she is this time zone rather than that one, so it is misleading.

Is there something defective in the child's approach to time? That depends. Given the child's life, does she need to keep track of time-zones? Does it matter to her that there are other time-zones? Perhaps it does; perhaps she talks to her grandmother in Denver, and finds the whole thing very confusing. But perhaps it doesn't. Perhaps she never talks to anyone outside her time zone, and never travels. In that case, there is no point in her thinking, as she sees the clock, "It is 3 p.m., Pacific Time".

The general point is this. Sometimes all of the facts we deal with involving a certain  $n$ -ary relation involve the same object occupying one of the argument roles. In that case, we don't need to worry about that argument role; we don't need to keep track of its occupant, because it never changes. We can, so to speak, pack it into the relation. For centuries people in Europe assumed that *being a summer month* was a property of months. July was a summer month, December was not. Once they started to visit the Southern Hemisphere, they had to take account of the relativity to places. July was a summer month in the Northern Hemisphere, but not in the Southern Hemisphere. A child who is unconcerned about and even unaware of the weather anywhere but where he is, can treat the issue of whether it is raining or not as a property of a time, rather than a relation between times and places. He says, "It is raining now" rather than "It is raining here now". (In this case the argument role is not always occupied by the same place, but always occupied by a place with a fixed relation to the agent, the place he is at.) The child we thought of above says "It is *now* 7 o'clock p.m.," treating *being 7 o'clock p.m.* as a property of the present time, rather than a relation between that time and a place or time zone. Before Einstein, we could treat simultaneity as a 2-ary relation between events, rather than as a 3-ary relation between a pair of events and an inertial frame, because in our daily life we never need to worry about alternative inertial frames.

In all of these cases, I say that the judgement *concerns* the fixed, unarticulated object, even though it is not explicitly about it. (See [5]). The judgement concerns the object because its truth-value depends on the object, even when it is not explicitly represented in thought. The child's is right when he thinks "It is 7 o'clock" because it is 7 o'clock Pacific Coast Time; he is right when he judges "It is raining now," if it is raining where he is.

Let us then return to the remark that I said was ambiguous, that when we

perceive we learn how things are around us. When we perceive how the world is around us and act upon it, we need to judge what distance and direction things stand relative to ourselves. But we do not need to keep track of who it is that we are judging things to be in front of or to the left of, at least as long as we are basing our actions on simple perceptual knowledge. In this case, our knowledge concerns ourselves but need not involve an explicit representation of ourselves.

Of course, humans use a wide variety of knowledge, not only the input from immediate perception. They combine this input with all sorts of facts and general principles that they know from previous experience and communication of various sorts. All of this requires a notion of themselves, and once we have one, there is an easy transition from “There is an apple” or “Apple in front” to “There is an apple in front of me”. But if our cognizing were confined to discovering facts about the objects around us and acting upon them, we would only need selfless thoughts. There are systems that perceive, and use the information about their circumstances they get through perception, that do not know that it is *their* circumstances they are learning about. During our formative stages and in certain moods later on, we may be such systems.

This then is the first aspect of self-knowledge, agent-relative knowledge of things that play various roles in our lives. This kind of knowledge is self-knowledge, in that it embodies knowledge of the relations things stand in to the agent; the thoughts are true because of facts about the agent. But it does not require that the agent have an idea of self or a notion of itself.

### 3 The detach and recognize information game

The concept of agent-relative knowledge fits into what I call the “detach and recognize information game” ([7]). We live in a world where we encounter the same objects on different occasions. On each of these occasions we are in a position to learn some facts about the object. If we can accumulate this knowledge, then in later encounters we will be able to deal with the object in light of this whole file of information, rather than simply what we can pick up on that occasion.

Suppose I am talking to you at a conference. You are occupying a number of agent-relative roles in my life: the person in front of me, the person I am talking to, the person talking to me, the person I see, and so forth. I am taking in a lot of information about you and deciding what to say to you and in general how to treat you. I accumulate information—the way I am thinking about it, this means I associate ideas to the notion I have formed of you. During this whole period, this notion is *attached* to the perceptions I have of you, to roles you are playing in my life and hence to the epistemic and pragmatic methods connected with those roles. It is what I call a *buffer*.

The conference is over; you go one way, I go another. I still have a notion of you. But now it is *detached*. The notion is no longer a buffer, but what

I call an *enduring notion*. It is not connected to any agent-relative roles and epistemic and pragmatic methods. I may want to ask you a question, or tell you something, or hit you or shake your hand. The ideas in my file may give me good reasons to do all of those things. And I do know how to ask questions, tell people things, hit them and shake their hands. But only when they are playing certain roles in my life. To do these, things, I need to get my notion of you attached to the appropriate roles.

Suppose then we meet again. At first I don't recognize you. I see you, form a notion, a buffer, and begin to collect information. After a while I recognize you as the person I saw before. This time the buffer doesn't endure; the information is transferred to my old notion of you, which now attached to the stream of perceptual information.

Now I am in a position to ask you a question or shake your hand. That's what recognition is: getting one's file on a person or thing attached to the roles that the person is playing in one's life, so one can bring one's information about the object to bear on one's decisions about what to say and do. Misrecognition is attaching one's file to roles that the some other object is playing; failure to recognize is not attaching one's file of an object to a role the object is playing.

In our lives, of course, we are not dependent on perception of objects to form notions of them and learn a lot about them. We can learn of people, and learn a lot about them, by talking to third parties about them, and by reading things that they have written and that have been written about them. Suppose you think W.V.O. Quine is a great philosopher; you have read many of his works and many articles about him. You understand he is going to be at a philosophy department reception; you go. You have all sorts of things you want to say to him. You have a notion of Quine, formed when you first read about him or heard your philosophy professor talk about him. You have all sorts of ideas associated with your Quine notion—that he wrote *Word and Object*, that he is sort of a modern Pythagorean, arguing that all we need to believe in is set theory; that he has a pleasant, friendly face, looking a little like David Hume might have had he gone on a successful diet and taken up vigorous walking. You have a rich mental file on Quine.

This same notion is also involved in certain desires you have. You'd like to shake Quine's hand, and tell him how much you enjoy his works, and ask him if he is serious about his Pythagoreanism.

You read that Quine will be at a reception given by the Philosophy Department. You go to the reception. You are standing in front of Quine. Let's say for a while that you don't recognize him—he looks a little older than the picture of him on your book. There are a couple of other people that you think might be Quine. But eventually, partly on the basis of your memory of how he looks, and partly on the basis of how one of the candidates seems to talk and act like a modest but great philosopher might, things sort of click and you take one of them to be Quine.

Let's focus on this period when you have seen Quine, and are noticing things

about him, but haven't quite recognized him. During this period, you have two notions of the same person. One of them is the one we talked about earlier. You acquired it when you first heard of Quine. Since that time information has been accumulating around it, and desires have been forming on the basis of this information. One thing associated with that notion is the desire to shake his hand. But of course there is no simple *method* for shaking Quine's hand. You have to know where he is relative to you to do this. This sort of information, where Quine is relative to you, isn't part of your notion. Or rather nothing very specific is—you think that he is at the same reception as you are, but that's not specific enough to support handshaking.

You have a second notion of Quine, a buffer, that you formed when you entered the room, and noticed him as one of the Quine candidates. Now this notion is associated with his position relative to you. As he moves or you move you "track him," keeping his position relative to you associated with the notion, so if you conclude that it is Quine you can strike.

Now you are standing next to Quine. The next-to relation is connected with epistemic and pragmatic methods. That is, there are certain ways of finding out more information about the person next to you, and there are certain ways of effecting the person next to you. You can find out more about him by looking. You can make him move by shoving. And you can shake his hand by turning towards him, extending yours, and smiling.

At this point you want to shake Quine's hand, and Quine is the person standing next to you, and you know how to shake the hand of the person next to you. So why don't you do it? Not because you don't see him there; you are looking at him very intently, almost staring. Because you don't quite recognize him; you're not sure. The desire to shake Quine's hand is associated with one notion of him, your enduring notion of him, the method for shaking hands is connected with another, your perceptual buffer of him as the person next to you. Only when you bring these together do you turn, smile, and extend your hand.

Suppose that  $F$  is an idea. That is,  $F$  is a cognitive particular in the mind of some person that combines with notions in that person's mind to produce judgements about the objects the notions are of. We might think of notions as names and ideas as predicates in a language of thought. Or we might think of notions and ideas as different sorts of nodes in a network that can be associated by edges of some sort. All I require of an idea is that it provide a way that notions can be modified.

What makes it the case that  $F$  was an idea of some agent-relative role  $R$ ? It is by being associated to the epistemic and pragmatic methods that are associated with  $R$ . I'll call these "normally  $R$ -informative ways of knowing" and "normally  $R$ -effecting ways of acting".

A notion that is associated with  $F$ , an  $F$ -notion, will serve as the *repository* of information that is normally acquired by methods for finding out about the object that plays the  $F$  role, and the *motivator* of actions that are directed at

(whose success depends upon) the nature of things that play the  $F$  role.

A notion, whether buffer or enduring notion, that is associated with the idea of a role  $R$ , I call an “ $R$ -notion”, for as long as the association continues. Of course, this is not an enduring trait of notions, since typically the same object will not play a given agent-relative role in our lives for very long. The apple is in front of me now, but won’t be a moment from now. My apple notion is an *In-front* notion now, but won’t be a moment from now. There are exceptions, of course. The role of being the planet I live on the surface of, for example, is one that has been occupied by Earth since I was born, and presumably will continue in that role until I die.

## 4 Mach and the shabby pedagogue

In 1885 Ernst Mach wrote:

Not long ago, after a trying railway journey by night, when I was very tired, I got into an omnibus, just as another man appeared at the other end. ‘What a shabby pedagogue that is, that has just entered,’ thought I.

In this case, Mach had agent-relative knowledge of a certain person, which he describes as “that man”. He knows how to find out whether or not a man he sees at the other end of a bus is a shabby pedagogue. He is thinking of the man he sees as a man at in front of him some distance away, for to find out more about that man he uses the methods appropriate for someone playing that role. Now in this particular case Mach wasn’t motivated to do much of anything. We can suppose that he saw a bunch of lint on the man’s vest. There are methods for getting the lint removed from the vest of a man a few yards in front of one, if one cares a lot about it. One can shout, “Brush that lint off your vest, you shabby pedagogue,” or one can walk forward and stretch out one’s hand and brush it off for him. Mach no doubt would have used those methods, had he cared that much about the lint he saw, and not gone on to make an important discovery:

It was myself: opposite me hung a large mirror. The physiognomy of my class, accordingly, was better known to me than my own. ([2], p. 4n.)

During the period when didn’t know who he was looking at, Mach could be said to have known something about himself in the following sense. He knew something about a certain person, the man he was looking at, the man he referred to as “that man.” And this man, that he knew something about, was in fact Ernst Mach. Let’s assume Mach’s belief was true. What made it true was that Mach himself was a shabby pedagogue. It was a case of Mach knowing something about Mach, and in that sense, self-knowledge.

But it wasn't what we would ordinarily call self-knowledge. Mach only had that after he recognized himself, and was ready to say, "I am a shabby pedagogue." What Mach had, before he recognized himself, was knowledge of the person he happened to be. He knew about someone whom he thought of as "that man", and that person happened to be him.

This was a failure of *perceptual* recognition; Mach failed to recognize himself as the person he was perceiving. Another kind of case of recognition and failure of recognition is what I'll call *documentary*. One fails to recognize someone that one knows as the person one is reading about (or seeing a video about or a portrait of).

Suppose that Mach, when he gets off the bus, hits his head and as a result has amnesia. "He doesn't know who he is," we would say. When he wakes up in the hospital he sees a story in the paper, "Famous Scientist is Missing". It goes on to say the Ernst Mach never returned home, and his family and colleagues at the university are upset. We can imagine Mach reading this and not recognizing himself as Mach, not knowing that he is the missing person, not remembering his name and his profession. This is a failure of documentary self-recognition. This is another case of Mach having knowledge of the person he happens to be, but not what we would ordinarily call self-knowledge.

## 5 Self-Attached Knowledge

Ordinarily all ones knowledge about oneself is integrated around a special sort of idea or notion of oneself that we express with "I". While my perception that the beer is in front of me may not require a representation of myself, the information I acquire is immediately integrated into self-attached knowledge, that I might express with "I see a beer" or "there is a beer in front of me". And when I read a piece of email, that says that John Perry's paper is overdue, I integrate this information into self-attached knowledge, "My paper is overdue," and I realize that it is me that has to get to work. I would think, "There is a beer in front of *me*, but *I* have a paper to do. So I want to turn to the question of this kind of knowledge. What did Mach lack, when he "didn't recognize himself"?

The view I advocate is simply that identity is a basic relation, and that our idea of self ("being me") is the idea of the agent-relative role, *is identical*. This is the role we each play in our own lives. That is, identity, like being in front or behind or above, is a basic relation relative to which we have epistemic and pragmatic methods. There are certain methods for picking up information about the person identical with us, and certain methods for having an effect on that person. The notion that is the repository of information gained via those methods, and the motivator of actions associated with that relation, is our self-notion. The person this notion is of, is the person we take ourselves to be.

An uncharitable summary of the view I am putting forward would be that

my idea of *me* is just is the idea of *the person identical with me*, which seems clearly circular, since the defining idea seems to contain the idea of *me*, the very idea being defined. It is agent-relative knowledge that keeps the account from circularity. It is not the idea, *person identical with me* that I need, but only the role-idea, *person identical*. My idea of *me* is not a part of this idea.

But why should we think that “being identical” is one of these epistemic-pragmatic relations?

Suppose I am at a party. I bend over to pick something up, and I hear a ripping sound that is characteristic of trousers splitting. I suspect that the trousers that have just ripped are my own. But they might not be. You can hear other people’s trousers splitting. Then I feel a hot flush in my face. So I am aware that someone is blushing. But who? It’s a silly question. Of course it is me. It is my own blushing of which I am aware. I can be aware of the blushing of others. But I can’t be aware of it in the same way that I can be aware of my own blushing. I feel myself blush, and anytime I am aware of blushing by feeling it, I know that it is I who am blushing.

This is an example of what I shall call a *reflexive method* of knowing. This is a method for finding out whether someone has some property or does not, that we can each use to find out about ourselves, but can’t use to find out about others. What one finds out may be accessible to others, using different methods. But the particular method in question can only be used by the person in question to find out about himself or herself. Feeling hunger is normally a way of detecting that one’s own stomach is short of food. Feeling thirst is a way of knowing that one’s throat is parched or that one’s body is short of water. There is a certain feeling, that children are trained to recognize, that signals that one’s bladder is full. In each case, someone else can determine the same thing, using a different technique. This alternative technique may even be superior. Perhaps you can tell that I am blushing, by looking, when I am not sure. Perhaps you can be sure that my stomach is full, having noticed what I have put into it, when I am still in that charming interval between being full and feeling full. Parents are often better judges of the states of their children’s bladder than the children themselves are. So the point isn’t that our reflexive methods of knowing about ourselves are always infallible or superior to any other methods. It is that only we can use them.

In these examples, the fact that these methods of discovering a person’s states are for the exclusive use of the person in question, is a matter of quite reliable but not quite necessary facts. We could imagine, for example, cases involving spinal columns that are connected across bodies which had the result that one person knew about the state of another person’s stomach in the way that we normally know only of the states of our own stomach.

We also have reflexive methods of knowing our own mental states, and in these cases it seems quite plausible to suppose that this is a matter of necessity. There are ways that I have of knowing whether I have a headache, or a throbbing tooth, or believe that Berkeley is west of Santa Cruz, that I employ to find out

about my own mental states, and others employ to find out about theirs. It is very difficult to imagine even a science fiction case in which one uses these methods to find out about someone else's mental states.

On Locke's theory of personal identity, at least according to one way of interpreting and developing it, this special way of knowing one's own mental states guarantees identity by bestowing identity. An instance of being aware of an experience, and the experience of which one is aware is known, necessarily belong to the same person, because it is in terms of this relation, that "same person" is defined.

Locke says that this method of knowing experiences and actions may be extended back in time. It seems that what he has in mind is what we might call first person memory. Compare:

- (1) I remember Mach wrote *The Analysis of Sensation*.
- (2) I remember Mach writing *The Analysis of Sensation*.
- (3) I remember writing *The Analysis of Sensation*.

(1) might be said truly by any of us who remember that Mach wrote the book; (2) might be said by any of his family and colleagues who remember him laboring on it, but (3) could be truly said only by Mach. It seems to be getting at a way of remembering, "remembering from the inside," in Shoemaker's phrase, that we each have of remembering our own past experiences, and with which we can remember no others [8].

But in the case of memory from the inside, it is not so clear that the link is a necessary one. In his influential paper "Persons and Their Pasts," Shoemaker develops some examples in which it seems we can remember, <sup>2</sup>in this way experiences we did not have.

There are also reflexive ways of acting. These are ways of bringing it about that someone has a property, that each person can use to bring it about that he or she has a property, but cannot use to bring it about that others have it. Towards the end of the movie "Spellbound" Leo. G. Carroll points his gun at Ingrid Bergman as she walks out the door of his office, having just disclosed that she knows that he framed Gregory Peck. We see this from Leo. G. Carroll's perspective. Then we see the hand holding the gun turn slowly, until the barrel of the gun is all that is visible on the screen. Then it fires. We know what Carroll has done, and to whom. He has killed someone, and the someone is him. The way Carroll held and fired the gun was a reflexive way of killing.

---

<sup>2</sup>Or at least "quasi-remember"; one quasi-remembers an experience if one fulfills all of the conditions other than being the original experiencer. See also [9] and [10]

## 6 What is special about self-notions

So far I have emphasized the analogy between self-notions, notions that are associated with the role-idea of *being identical*, and notions that are associated with ideas of other agent-relative roles. But there is one important disanalogy. As we saw above, the other cases (or most of them), the attachment between notion and role cannot be permanent, because different objects occupy the same roles relative to us at different times. We noted one exception, the role of being the planet lived on. And of course there are indefinitely many others based on the same general idea: being star that warms, being the part of the universe relied on most, etc. Clearly, the person with whom I am identical falls into this category of non-shifting roles, but in this case the reliability is not a contingent matter. It is logically possible that I will move to Mars, or even to a planet in the solar system of another sun. But it is not possible that I will ever be identical with anyone other than me. If I have picked out the right notion to be my self-notion once, it will continue to be the right one.

This is not to say people cannot be wrong, nor that they might not change their self-notion. My parents might have, as sort of a practical joke, raised me believing I was really Al Smith, former Governor of New York and Democratic Presidential candidate. They tell me that in an extreme use of the witness protection program, I was shrunken and made child-like. My memory was obliterated and then I was passed off as John Perry, the real John Perry having died in 1944 as an infant. They recommend that I go along with the story, and pretend to believe that I am really John Perry. At some point I do a bit of research, uncover the prank, and realize that Al Smith really did die in 1944, and I am John Perry, just as everyone but me and, as I thought, my parents, had always believed. At that point my Al Smith notion would cease to be my self-notion, and my John Perry notion take over that job. And I might become convinced, at some point in the future when I sink into madness or senility, that I am really Napoleon and not John Perry. One can be mistaken about who one is.

But for most of us most of the time, possessed of sober parents and a sound mind, there will never be any reason to detach our self-notion from our enduring notion of the person we are, and we and everyone who knows us takes us to be.

## 7 Back to Mach

Now let us see if this analysis provides us with a plausible candidate for what Mach lacked in the two cases we considered.

When he looks to the far end of the bus, Mach gets information about himself in a way that is not normally self-informative, but normally “person-in-front-and-looked-at” informative. So this information doesn’t pass into his self-notion; it is not combined with information gotten in reflexive ways. And

it doesn't motivate normally reflexive actions.

Suppose Mach looks down at his own vest and sees a big piece of lint. (Mach himself provides us with a picture of the way one's front characteristically looks to oneself.) He would have associated the idea of having a large piece of lint on one's vest with his self-notion. That's what I mean by saying that the self-notion is the repository of normally self-informative perception. Now if Mach had desired not to have large pieces of lint on himself, he would have reached out and removed in a way that works when the piece of lint is on one's own vest. If he has this desire, and the idea of having lint on one's vest is associated with his self-notion, we would expect him to take such a normally self-directed action. That's what I mean by saying that the self-notion is the motivator of normally self-directed actions.

But when Mach sees a piece of lint on the vest of the person in the mirror he does not act in this way. The information is not gotten in the normally self-informative way.<sup>3</sup> So it is not combined with the other information in the self-notion, and doesn't lead to the action that works to remove lint from oneself. At the beginning of the episode, Mach formed a notion for the person he saw, whom he took to be getting on the other end of the bus. This was a notion of himself, but not a self-notion. We assume Mach knew who he was, and so that he had a notion of Ernst Mach as having all of the well-known properties of Ernst Mach that was also a self-notion. Mach's beliefs change, during the episode, in that he transfers the information associated with the new notion formed when he got on the bus to his old self-notion. If, after he has made the transfer, he notices that the person in the mirror has a piece of lint on his vest, he will pick the lint of his own vest in the normally self-dependent and self-effecting way of picking lint off one's vest.

Now let us turn to the case of Mach the amnesiac, reading about himself in the paper. He plays the role in his own life of "person being read about." He forms a notion of this person; he knows he is called "Mach," is a scientist, and is missing. Even in the middle of a bout of amnesia Mach would have had at least a self-buffer, a notion tied to normally self-informative action and perception. He realizes that the table he sees in front, is in front of him, the same person who is reading the paper and can't remember who he is. He does not associate the ideas he picks up from reading the paper, of being named "Mach", being a scientist, and being missing, with his self-notion; that is, that role remains unattached to the self-notion.

---

<sup>3</sup>In many situations, looking into a mirror is a normally self-informative way of getting information. In Mach's case the mirror was far away and not set up for self-viewing, and he didn't seem to realize at first that he was looking into a mirror.

## 8 Self-notions and “I”

The word “I” refers to the speaker or writer. Thus the meaning of the first person associates it with a role in the situation of discourse. One reason for having such a word is that it puts a modest cognitive load on the hearer in a variety of common speech situations. One is a face to face speech situation, in which some of the information a speaker is providing will be likely to motivate actions towards the speaker. So, for example, you and I are sitting at a table and I say, “I’d like some salt.” I want you to hand the salt to me. You, being an agreeable sort, will hand the salt to anyone who wants it. In order for my request to work, I simply have to depend on you knowing English, and being able to tell that the person speaking to you is the person playing a certain perceptual role. You don’t have to know anything more about me.

If, on the other hand, I were to say, “John Perry would like some salt,” there is no telling what you might do. Perhaps you would say, “Good for him”. My request puts a larger cognitive load on you; you need to know more in order to be expected to accede to it. You need to have recognized the person before you as a person named “John Perry”. Even if I’m pretty certain that you do know who I am—perhaps we are old friends—it will still sound odd and sort of pompous to say “John Perry wants salt.” The only fact relevant to your passing the salt would usually be that I am a human being who wants the salt.

On the other hand, use of the first person is inappropriate in other situations. If I call you on the phone, and you clearly don’t recognize my voice, I need to tell you my name, and it’s impolite to do otherwise.

The first person also puts a relatively light cognitive load on the speaker. When Mach had amnesia, he still referred to himself with “I”—or “ich” at any rate. To know that he was doing so he needed to know i) the meaning of “I” and ii) that he was the utterer of the words he was speaking. That is, he needed to realize that that the words he was going to speak, the one’s he was planning to say, would be spoken by him. There is in fact a certain way of knowing whom words are spoken by, when the words are the result of one’s own planning and articulating; they will be spoken by the planner and articulator, and hence, if the word is “I”, will refer to that person. Thus, a person who has forgotten who they are, and so no longer has an enduring notion associated with the self-role, may nevertheless successfully and confidently refer to himself or herself.

Most of us don’t get amnesia, so having a referring device that enables us to refer to ourselves without knowing very much about ourselves isn’t all that important. But in philosophy, having such a device is often of great value. The point isn’t that we don’t know much about ourselves, but that we don’t want to assume much about ourselves. Why should Descartes say “I think” rather than “Descartes thinks?” [1]. When he says “I think” he does not assume that there is a person with a past like Descartes has and a name “Descartes”. He can use the method of doubt to bracket all of that knowledge he has of himself.

Now what Descartes can’t quite do is to refer to himself without assuming

anything except what he is entitled to at the beginning of the Second Meditation. For to refer to himself with “I” (or ‘je’) he needs to assume that there is a language with a word in it that refers to the person who uses it. But, to be fair, to write his *Meditations*, as opposed to merely meditating them, Descartes does need to assume that, or at least play along with his inclination to believe that, there is a language.

The first person, then, does not give philosophers as secure a way of talking about themselves as their self-notions give them of thinking about themselves. I can think about myself so long as I exist and have a self-notion, even if there is no language at all, although in the case I cannot refer to myself. But first person pronouns, like “I”, “ich” and “je”, give philosophers a pretty secure way of referring to themselves, one that should satisfy all but the most dedicated solipsists. The most dedicated solipsists don’t write very much anyway.

## References

- [1] Anscombe, Elizabeth. The First Person.
- [2] Mach, Ernst. *The analysis of sensations*, translated by C.M. Williams and Sydney Waterlow, (Chicago & London: Open Court, 1914), p. 4n.
- [3] Nagel, Thomas. The Objective Self. In Carl Ginet and Sydney Shoemaker, eds., *Knowledge and Mind*, 1983.
- [4] Perry, John. *The Problem of the Essential Indexical*. New York: Oxford University Press, 1993.
- [5] Perry, John. Thought Without Representation. *Supplementary Proceedings of the Aristotelian Society*, vol. 60 (1986): 263–83. Reprinted in [4].
- [6] Perry, John. Self-Notions. *Logos*, 1990: 17-31.
- [7] Perry, John. Rip Van Winkle and Other Characters. *European Review of Philosophy*, Volume 2, 1997: 13-40.
- [8] Shoemaker, Sydney. Persons and Their Pasts. *American Philosophical Quarterly*, October, 1970.
- [9] Shoemaker, Sydney. *Identity, Cause and Mind*. Cambridge: Cambridge University Press, 1984.
- [10] Shoemaker, Sydney. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press, 1996.