

ON KNOWING WHO I AM

Keywords: *Castañeda, amnesia, efficiency, Kaplan, context, character, basic actions, self-informative, self-effecting, other-informative, primitive self-knowledge, self-notion, self-attached, essential indexicals.*

(August 5, 2022)

1. On Knowing Who One is. Most of us would say that we know who we are. But there are exceptions. In my version of Hector-Neri Castañeda's (1968) wonderful example a soldier – I'll call him 'Elwood Fritchey' – is involved in a battle, and performs some heroic deeds.¹ Late in the battle he is injured in a way that causes amnesia of an appropriately philosophical sort; memories of his past are inaccessible to him, but remain deep in his brain.² He also loses his dog tags and wanders far from the battle. He is clearly a GI, so he ends up in a military hospital. But no one can figure out who he is. The heroic soldier had blue eyes, as does Elwood, but due to an error, the records said that Elwood Fritchey had brown eyes. So the possibility that the patient was the missing hero Elwood Fritchey was mistakenly eliminated.

Consider Elwood's first few hours in the hospital. He can't remember his name or where he is from or anything else of that sort. He is given a list all of the soldiers in the recent battle, with a few key facts about each, but he can't say which one of them he is. He says, "I'm sorry, I don't know who I am".

But in an important way Elwood *does* know who he is. The nurse comes in with a tray of food. Elwood says, "Good, I am hungry." He has feelings of hunger. He knows who is hungry. He knows into whose mouth he needs to put the food, in order to relieve the hunger he feels. He knows whose hands

1 Castañeda's hero is Quintus. The sequence of events Castañeda presents is a bit puzzling. Consider, for instance, the case of a man, to be called „Quintus,“ who is brought unconscious to a military tent, but on gaining consciousness suffers from amnesia, and during the next months becomes a war hero and gets lost in combat and completely forgets the military chapter of his life. (p. 446)

I suspect he had the sequence of events I attribute to Elwood in mind, where the heroism comes before the amnesia, but I may be missing something.

2 See Perry (1993).

and arms will move, when he decides to do so. He knows that the person who is speaking when he says, “I am hungry” is the same person whose hunger will be relieved. So, in a way, he does know who he is, in spite of not knowing his name and history before wandering away from the battle.

2. Efficiency. Ellen, the soldier in the next room, will know things about *herself* in the same way that Elwood knows about himself. If she has feelings of hunger, she will know that she is hungry, and she will know whose mouth she needs to put food in to relieve that hunger. Perceptual knowledge is *efficient*, in the sense Jon Barwise and I gave the word in *Situations and Attitudes*.³ Different people, in different circumstances, can know or believe different things, by being in the same perceptual state. Our actions are also efficient. Different people, in different circumstances, can bring about different results, by performing the same basic actions, that is, basically moving their bodies and limbs in the same way.

Elwood and Ellen are both sitting at a bar near the hospital. A mug of beer is placed in front of each of them. They are in the same perceptual state, more or less, the state one is in when one sees a mug of beer in front of one. Elwood believes there is a mug of beer in front of him; Ellen believes there is a mug of beer in front of her: efficient perception. Elwood moves his arm and hand, picking up the glass and bringing it to his lips, opens his mouth, tilts the mug a little more, and swallows. He brings it about that he has a drink of beer. Ellen makes basically the same movements, bringing it about that she has a drink of beer: efficient action.

It's clear that Mother Nature appreciated efficiency. A species of animal has a repertoire of methods for picking up information, and a repertoire of basic actions they can execute. The billions of members of a given species will pick up different information, about themselves and the context they are in, and perform different actions with different results, on themselves and the things in a position to be effected by them.

3. Context and Character. We can borrow terminology from David Kaplan's theory of indexicals and demonstratives to elaborate on this.

An agent, location, time, and circumstances, such that the agent is in the location at the time, and the circumstances obtain in that location at that time, is a *proper context*. The *character* of an indexical expression is a function from contexts to suitable *contents*. The character of an indexical expression takes us from a context to the referent of the expression in the context. The character of 'now' takes us to the time of the context, of 'I' to the agent, of 'here' to the location. The character of a sentence with an indexical or indexicals is a function from contexts to singular propositions that includes the referents of the indexicals. So “I am hungry now” in the context consisting of Elwood, his

3 Barwise & Perry (1983).

room r , and the t moment the nurse enters, yields the singular proposition that Elwood is hungry at t .

Sentences with indexicals provide ways for different agents and agents in different contexts to say different things in the same way. In a debate about action I might say to Michael Bratman, probably falsely, "I am right and you are wrong". He might reply, probably truly, with the same sentence. We use the same sentence to disagree with one another.

David Kaplan's system of context and character provides a good account of the efficiency of language with indexicals. His account can be adapted for other kinds of efficiency. Philosophers of action call the movements we can normally make at will in any circumstances *basic actions*. The same basic actions, made by different agents and/or in different contexts, have different results. You mow your lawn with basically the same sequence of basic actions I do. But I mow my lawn, you mow your lawn. Different agents, different contexts, the same sequence of actions (more or less), different lawns mowed. The basic actions can be thought of having an important character-like property, that gets at a key element of their causal role. An agent who *knows how* to do things basically knows which basic actions will bring about those things in different contexts.

Elwood knows how to walk and eat and say that he is hungry using 'I'. But he doesn't know how to say that he is hungry using his name. When I wake up in the morning, look out the window, and see that it is sunny, I know how to say so using the indexical 'Today': "Today is sunny". But until I look at a calendar I won't know how to say this by using the date of the day of my waking and looking.

4. Self-Informative perceiving and self-effecting acting. Elwood has a number of ways of knowing things about himself and the things around him. There are his five external senses, by which he can know that various sorts of things are happening around *him*. For example, he knows by means of vision that the nurse he sees has tray of food. We think of the external senses as ways of finding out about other things. But they also provide information about the perceiver; Elwood knows that there is a tray of food in front of *him* – using italics as short for "in a way he would express with the first-person".

But we also have internal senses. Elwood knew that *he* was hungry through interoception, which along with proprioception, provides us with *normally self-informative* ways of knowing about what is going on inside of us and how our limbs are arranged. I say "normally" because a philosopher can imagine exceptional cases. A mad scientist sitting behind us at a movie manages to connect nerves coming from your stomach to my brain and vice versa. The more I eat, the fuller you feel. Definitely not normal. There are normally self-effecting ways of acting, as illustrated by Elwood and

Ellen's beer-drinking. But who knows that could happen in an AI lab, or a philosopher's imagination.

Introspection is a self-informative way of knowing what's happening in our own minds. Do we need to say "normally"? The mad scientists in my imagination haven't figured out how to wire us up so that I know what you are thinking by introspection and vice-versa, but give them time.

I'll say that seeing, hearing, and our other external sense provide *normally other-informative* ways of knowing about things around us, which serve as the basis for inferences about more distant things. But they are also *normally self-informative*, since one learns that one stands in various relations to the perceived objects or inferred objects. So Elwood learns in a normally other-informative way, when he sees the nurse, that the nurse *he* sees has a tray of food. He learns, in a normally self-informative way, that there is someone in front of *him* with a tray of food. He infers that there is someone who cooked the food and gave it to the nurse.

Normally other-informative ways of knowing can also be self-informative in a different way in the right circumstances. I can look at a person, read about a person, google a person, and so on. Most of the time one uses these methods to find out about other people; that's their normal use. But the same methods can be used to find out about oneself, without a philosopher or mad scientist playing a role. If I don't remember where my class meets, I'll go the appropriate university website and enter my name, and find out where *my* class meets in the same way I would find out where Michael Bratman's class meets. If I can't remember my phone number I can look it up in the phonebook – if I, unlike Elwood, know my name.

To use normally other-informative methods in this way, one needs to know how the sources of such information identify him – by his name, or library card number, or social security number, or whatever. The university library may post a list of users with the worst record of returning books on time, listed by their library card numbers. I may be amazed at the number of times the holder of #1234567 has been late in returning books, without realizing that it is me, at least until the Stanford Library Police show up at my house.

This is Elwood's situation as he looks at the information under "Elwood Fritchey" on the list of missing soliders: he is learning facts about himself without realizing that he is doing so. He learns that Elwood Fritchey was born in Broken Bow, Nebraska. He is Elwood Fritchey, but he doesn't realize that he is getting information about himself in this normally other-informative way, and has no idea where *he* was born.

A normally self-directed action is a sequence of basic actions normally used to have an effect on the agent. There is a way of scratching one's back that we each use to scratch our own backs. Such actions are often precipitated by self-informative perception, as when one has an itchy back.

I use the phrase “primitive self-knowledge” for what one knows via normally self-informative methods, and what one knows how to do using normally self-effecting actions. It doesn't include what one learns about oneself through normally other-informative methods. If my wife Frenchie tells me that I have a bit of lettuce caught in my teeth, that's not part of my primitive self-knowledge. If I learn this by feeling the lettuce with my tongue, it is.

I use the concept of primitive self-knowledge in a very broad way, to get at what I see as a key aspect of evolution, common to animals from humans to snails and beyond. Any species of living things is equipped with a range of states that, relative to constraints and circumstances, carry information about crucial things occurring in its body and environment. Being in these states cause behavior that, given that information, contributes to survival, that is, behavior that increases the probability of the living thing surviving, and/or its genome surviving, and/or the species surviving. The key point is that Nature is *efficient*. A creature's being in a state at a time indicates something about *that* creature's environment at *that* time and this state will lead to behavior useful for *that* creature (and/or its genes, and/or its species). A different member of the same species may be in the same state. It's being in that state indicates something about that creature's environment. It may lead to the same behavior. Chicken-1 sees a kernel of corn in front of her, walks forward and pecks, with the result that chicken-1 gets some nourishment. Chicken-2 see a kernel of corn in front of her, walks forward and pecks, with the result that chicken-2 gets nourishment. Mother Nature or God or whoever or whatever got our world going, realized the importance of efficiency.

Animals with primitive self-knowledge typically need to integrate information from different senses, or primitive precursors of senses, and remember information at least for a short time, and act in light of this integrated and remembered information. I call this collection of integrated information a *primitive self-notion*, when it contains only information acquired in normally self-informative ways. The primitive self-notion is not an idea that is common to all the bits of information, but more like a mental file folder into which all the bits are put. Elwood's amnesia doesn't prevent him from remembering the things that have happened to him since he wandered off from the battle, so Elwood has a primitive self-notion that is filling up with things he has observed and inferred since arriving at the hospital: *I was fed yesterday; nurses bring me food; they will probably continue to bring me food, etc.*

Most of us humans also have a lot of what I call “self-attached knowledge”, which is what I take to be part of what we usually have in mind when we use the phrase “self-knowledge.” If we know our name, or our social security number, or the number on our library card we can pick up information about ourselves in normally other-informative ways, the same ways we might

use to pick up information about others, and integrate with our primitive self-knowledge. Before the battle Elwood might have looked at a list on the Barracks Bulletin Board each morning to see whether he had KP duty or some other assignment that day. If he saw that Elwood Fritchey has KP duty and he felt hungry, he would know that he was both hungry and had KP duty; both will be part of his self-knowledge.

Even though non-human animals tend not to read or google or look for their names on lists on bulletin boards, they can learn about themselves in ways that are normally other-informative. A mammal or a bird can see its reflection in a pond or a mirror. In general, these bits of information will not be incorporated into the self-notion so as to interact with primitive self-knowledge and effect normally self-effecting actions.

But some animals – chimpanzees, various birds, and others – can integrate this information from a mirror with their primitive self-knowledge so their actions are motivated by by the combination. This seems to be what is involved in passing Gallup’s mirror test.⁴ This integrated collection is not a primitive self-notion, but something more sophisticated, which I just call a “self-notion”.

If we know English, and have primitive self-knowledge, and are not worried about being trapped in the lab of a mean psychologist, we can use ‘I’ with confidence to express what we learn in that way. If I feel pangs of hunger I can confidently say “I’m hungry”. If I know by introspection that there is some thinking going on --- worrying about whether I exist, perhaps --- I know that *I* am thinking, and so, it seems, that *I* exist. Elwood can’t say much of anything with confidence about where he came from or how he has spent his life. But if he limits himself to his primitive self-knowledge, he can be pretty confident that what he says is true. When he tells the nurse, “I am hungry” the nurse won’t say “How do you know? You don’t even know who you are!” But he eventually does learn about his past, and as his self-knowledge grows he can use ‘I’ to express it.

5. “Essential indexicals.” The connection between self-knowledge and ‘I’ and other indexicals is easily overstated. ‘I’ is an expression of natural language, the main purposes of which is communication and retention of information. ‘I’ provides a way of conveying information about ourselves to people who in a position to see who is speaking. As a result, it is the standard way of referring to ourselves when people can see that we are speaking, even if they do know our name.⁵ I argued in “The Problem of the Essential Indexical” that identifying ourselves (or our location, or the time of our utterance) indexically is often *essential* to explaining something to the person to whom we are speaking. I visit a doctor; she asks, “Why are you here today?” “John

4 See Baker, 39n

5 See Korta & Perry (2011), section 3.5 for a discussion of such cases.

Perry is having a lot of headaches” isn’t a helpful answer, but “I am having a lot of headaches” is a pretty good explanation. Similarly, if the doctor had asked, “Why are you at 795 El Camino Real on the 6th of April” it would have just confused me, but “Why are you here today” worked just fine.

I did not mean, and did not say, that whenever a thought or action has a *character* or something like it, there must be an indexical involved, so that our brains are full of some kind of mental indexicals. I wasn’t claiming that indexicals were essential to thought and action. My point was that the contribution that indexicals make to explanations, as the one I gave in the last paragraph, is hard to explain on the traditional theory of propositional attitudes. In the example “John Perry is having a lot of headaches” and “I am having a lot of headaches” express the same singular proposition. What differs is not what is said, but how it is said. We need to make the same distinction between what and how with belief states and types of actions, and so the context/character distinction is helpful. But the difference in ways of believing and ways of acting isn’t typically a matter of indexicals.

If a new-born sees a nipple in front of it, it has primitive self-knowledge that will cause primitive self-directed action that fits the situation: sucking on that nipple. Other infants will do the same: efficiency. The perceptual state is efficient, the action is efficient, and the causal connection between them is to be explained in terms of general facts about infants, nipples, and nutrition. But its not efficient because indexicals are involved.

Finally, suppose you are having dinner with President De Gaulle and some others. After his first bit of steak, De Gaulle says, “Please pass the salt to President de Gaulle”. Everyone knows who de Gaulle is, but his assumption that this is so seems a bit self-important. If he has just said “Please pass me the salt” it wouldn’t have had this implication. Same proposition expressed, but different natural inferences made by the listeners. (CP, pp. 35–36; 83–85.)

6. On Knowing Who One is (continued). On my view, we all know who we are in basically the same way that Elwood does his first day in the hospital. Knowing who I am is a matter of knowing who I am relative to a population of identified objects, typically agents. In his room, Elwood can identify two humans, himself and the nurse. Perhaps later, at dinner, there will be dozens he can identify perceptually. If he is in doubt who he is, he can follow Wittgenstein’s suggestion, raise his arm, and look around to see whose arm goes up. But if you give him a list of names, say of the soldier in the battle, he won’t be able to say which one he is.

The difference between Elwood and I is that there are many more sets of identified individuals relative to which I know which one I am. The Palo Alto phonebook, for example, lists thousands of individuals; I know I am the one identified as “John Perry”; if there are more than one of them, I am the one who lives on Hilbar Lane. But given a list of emeritus Stanford faculty, identified by their university ID numbers, I no idea which one I am.

8. On Who We Are. I think all of us know who we are in basically the same way that Elwood does; we are the person about whom we have primitive self-knowledge. We are the person who is having the experiences that we are having, who is having the internal events they disclose the occurrence of, and who is perceiving external things around her via external senses: primitive self-knowledge. As we learn more about these things, including their relation to other things inferred rather than perceived, we learn more about ourselves, and develop a richer and richer self-notions, all ultimately grounded on self-informative ways of perceiving.

But this leaves some interesting questions open, questions that may not have occurred to Mother Nature but bother philosophers. Do we have essences, combinations of ordinary properties that are necessary and sufficient for being who we are? If not, do we have haecceities, non-ordinary properties, with no empirical implications, necessary and sufficient for being who we are? If neither of these things, what makes it the case that the same person does different things in different possible worlds – the problem of “trans-world identity”. I’m thinking about these things. But, more and more frequently, I find myself asking, “If Mother Nature didn’t have to worry about these things, do I have to?”

References

- Almog, J., Perry, J., & Wettstein, H. (1989). *Themes from Kaplan*. New York: Oxford University Press.
- Barwise, J., & Perry, J. (1983). *Situations and attitude*. Cambridge, Mass.: The MIT Press. (Reprinted with a new introduction by CSLI Publications, 1999.)
- Castañeda, H. N. (1968). “On the logic of attributions of self-knowledge to others.” *The Journal of Philosophy* 65, 439–456.
- Kaplan, D. (1979). “On the logic of demonstratives.” *The Journal of Philosophical Logic* 8, 81–98.
- Kaplan, D. (1989). “Demonstratives.” In Almog, J., Perry, J., & Wettstein, H., (eds.), *Themes from Kaplan*. New York: Oxford University Press, 481–563.
- Korta, K., & Perry, J. (2011). *Critical pragmatics*. Cambridge: Cambridge University Press.
- Perry, J. (1993). *The problem of the essential indexical and other essays*. New York: Oxford University Press. (Enlarged edition, Stanford: CSLI Publications, 2000.)
- Perry, J. (2002). *Identity, personal identity and the self*. Indianapolis: Hackett Publishing.
- Perry, J. (2019). *Re-visiting the essential indexical*. Stanford: CSLI Publications.