Thinking about the Self

John Perry
University of California, Riverside
Stanford University

INTRODUCTION

Suppose that when Bill Clinton moved to the White House he was unaware that the Secret Service used the acronym "POTUS" to refer to the president. In his morning briefing summary he sees the sentence "POTUS meets with the Queen of England at noon." This upsets him. Who is POTUS? Why is this person meeting with the Queen, instead of Clinton? He says to Hillary, "POTUS is meeting with the Queen; who in the world is POTUS, and why is he meeting with the Queen instead of me?" It seems that the person Clinton is asking about is Clinton himself; the right answer for Hillary to give is, ``You are POTUS." Clinton *said* something about himself, without realizing it. And, indeed, it seems that Clinton *believed* something about himself, without realizing it. I will say that Clinton had a belief that was *merely* about the person he happened to be. This is a self-belief, in that the truth-conditions of the belief impose conditions on the believer. But it is not what we would normally call having a belief about oneself.

It seems then, that we can make a distinction between saying something about yourself in the sense in which that merely said something about the person you happen to be, and saying something about yourself in the sense of making what we might call a "self-assertion", which implies not only that the person you referred to was you, but also to that you knew this and intended to convey it. Use of the word 'I' indicates that one is self-asserting. And an analogous distinction may be made at the level of belief. We have the phrase "self-knowledge"; I will supplement it with the phrase "self-belief". Clinton had a belief about the person he happened to be, but it wasn't' a self-belief.

Hector Neri-Castañeda [Castañeda, 1966] noticed that reports like "Clinton believes he is meeting the Queen for lunch" and ``Clinton said he was meeting the Queen for lunch," are in some way ambiguous. These reports could be construed so that they are true in our opening case, even before Clinton realized that he was POTUS. Usually, however, we would take such reports to imply more than this, that is, we would take them as reports of self-belief and self-assertion. Castañeda uses an asterisk to mark the difference, and so will I. Thus, "Clinton believes that he* is meeting the Queen for lunch" would not be true before he learned that he was indeed POTUS, but would be true after he learned this. And in the imagined situation, ``Clinton said that he* was meeting the Queen for lunch" would not be true, but if he had said ``I am meeting the Queen for lunch," it would have been true. Castañeda thought that `he*' was a different word than `he', or at the very least marked an ambiguity. One might be skeptical about this, and suppose the added information is a matter of Gricean implicature. This issue won't matter in this essay.

Self-beliefs are typically expressed with self-assertions, and hence with the first person. Here is simple account of how the first person works: an utterance of the word 'I' stands for the person who utters it. The word 'I' is thus one way that Bill Clinton has of referring to himself. He can also call himself 'Bill Clinton', or even, in the appropriate circumstances, 'that man'. If he is expressing a self-belief, then using the word 'I' is the natural way for him to refer to himself. If, on the other hand, he is expressing a belief that happens to be about himself, but is not a self-belief, we would not expect him to use the first-person.

But it doesn't seem very satisfactory to explain self-belief in terms of the first-person. It seems like there is something about self-beliefs that leads makes them correctly and naturally and appropriately expressed with the first person; the difference between self-beliefs, and beliefs that are merely about the person one happens to be, as I shall put it, is somehow reflected in the difference between first and third-person ways of referring to oneself, but not constituted by it.

My first goal in this paper is to construct an account of self-beliefs that distinguishes them from beliefs that are merely about the person one happens to be. Such an account should explain the close connection between self-belief and the first person. My second goal is to consider issues raised by David Rosenthal's in his paper in this volume [Rosenthal, forthcoming] about what he calls the ``essential indexicality'' of HOTS. HOTS are Higher Order Thoughts about our own mental states and activities. Rosenthal's account involves giving up Sydney Shoemaker [Shoemaker, 1968] principle that a person is immune to certain sorts of errors about her own mental states, namely, whether it is she* that is having them. I will argue that we should not give up Shoemaker's principle, and that we need not do so to account for the feature of HOTS that Rosenthal calls their ``essential indexicality''. These issues are crucial for those who are inclined to accept Rosenthal's account of consciousness, in which HOTS play a crucial role.

THE CAUSAL ROLE AND CONTENT OF BELIEFS

Before looking into what is special about self-beliefs, we need to fix some ideas about beliefs in general. I take it that beliefs are particular states of the brain; I assume that they are physical states, but a property dualist, who thinks that belief states are not physical in some important sense, could agree with everything I say here, so long as she is not an epiphenomenalist.

I use the word `state' in two ways, for particular states in the mind of a particular person, and for types of such states, that get at their internal, non-circumstantial properties. In either sense, states can be total or partial.

I conceive of brain states, in the first sense, as non-basic particulars of a certain sort for which some philosophers in the seventeenth and eighteenth centuries used the flexible term *mode*.[1] In this use, a mode is not basic particular, but one that involves a more basic particular (God, or Nature, perhaps), or more basic particulars (atoms, perhaps), or complexes of more basic particulars, or successions of complexes of more basic particulars. I won't assume that the more

[1] See [Locke, 1694] Book II, Chapter XII, section 4; [Spinoza, 1677] Part I, Definition V.

basic particulars have to be completely basic particulars, or even that there are any of those. Earlier stages of a mode of given kind will pass along traits and patterns of traits to later stages in more or less predictable and often quite useful ways.[2] The ripple that flows across the lake is a mode, involving successive complexes of water molecules, rising and falling in predictable ways. A dent in a fender is a mode; left alone, it won't change much with time. Ripples and dents are not universals but particulars. They begin to exist at a certain time, say when someone tosses a rock into a lake, or backs into another car in a parking lot. They last for a while, and then may cease to exist, say when the ripple reaches shore, or the fender is repaired, or the car goes to the junkyard and is crushed.

Another example of a mode is bump in a carpet, of the sort that inevitably occurs when an amateur tries to lay a large carpet, and gets a pocket of air trapped under it. You can push the bump all around the carpet. But it is the same bump. You can refer to it: ``I've been trying to get rid of that bump for an hour." You are not likely to think of it as a metaphysically basic object; you could get by without referring to it, and just talking about the carpet, and how various parts of it rise and fall over time, as the trapped air is pushed about. But it's natural to talk about it and refer to it.

Modes that preserve wave-forms, ---sound waves, radio waves, and the like---can preserve information about their sources. The movement and preservation of the waveforms, and thus information, occurs naturally. The technology of preserving computer-generated files to disks uses preservation of structure, at some appropriate level, to preserve information, even as the file is moved about the disk by the operating system, in ways that lose other kinds of structure --- a single file may end up stored in various discontinuous regions, for example. So such files are modes.

Computer files change over time. When things go right this is due to additional input by those who use them; when things go wrong it is due to

[2] Everything I say is consistent with Dan Giberman's ``Glop Theory," for example [Giberman, 2010].

malfunctions of various sorts, and the files are corrupted. Although the files are designed and created to preserve information, the information they carry is not an essential property of them; it may be changed or lost, while the file continues to exist. I think beliefs and other cognitions, and the ideas that comprise them, are modes in this sense.

In this sense, modes are not universals. They instantiate universals, that is, modes have properties and stand in relations of various sorts. The dent may be oval; the ripple may diminish in size as time passes; the file may be small when it begins to exist, and then grow large. Here the second use of `state' is useful. In engineering and science, the state of a system is constituted by its local properties, those it has in virtue of the intrinsic properties of its parts, and the ways they are related. A state can be total or partial. And we can speak, in either sense of states of the whole system in which we are interested, or states of its subsytems, or parts of the system. The state of an engine is constituted in part by the states of its crankshaft and camshaft, for example. The state of the solar system, at a certain time, is constituted by the way the planets are oriented to the Sun and to each other, where they are in their rotations and their orbits, and things like that. The solar system may have other important properties at a given time; perhaps it is now coming closer to a Black Hole that will eventually suck it up and destroy it. That's an interesting fact about the solar system, but not a part of the state of the solar system.

The states of a system, in this sense, do not exhaust its properties, for a system will have many properties in virtue of relations of things inside the system to things outside the system. The lake is in the state of having a ripple, but it came to be in that state because a boy threw a rock into it; that is a property of the lake, a fact about the lake, but not a state it is in. The car has lost value, not simply in virtue of the dent, but because people prefer cars without dents. The solar system will soon be annihilated, not because of the state it is in, but of a property it has to things outside of itself, namely because of its proximity to a black hole. An idea may be of a certain property, or thing, because of the

circumstances in which it was formed, its association with certain words, and the way it influences action. A belief may be true or false because of the way the external object its ideas are of are related to one another.

Individual beliefs are parts of a system of beliefs and other cognitions and the properties of the cognitions and their relations constitute a person's *cognitive state* at a given time. The states of individual beliefs and other cognitions make *incremental* contributions to the state of the whole cognitive system. We are quite adept at characterizing these contributions, and predicting what sort of difference they will make to the whole system, and the activities of the person to which it belongs, in spite of our relative ignorance of their intrinsic nature. For example, a normal human being comes into a bookstore, and the clerk asks what she is looking for. If she says ``Mysteries,'' he will lead her to one part of the store, if she says, ``Philosophy,'' he will lead her to another part of the store. The choice of one word rather than another indicates the presence of one desire rather than another, and this difference, which one supposes must be a very tiny physical difference in the whole system, gives the clerk good evidence about where he is likely to make a sale.

When we make such inferences, from bits of language or other small behaviors, we manage to do so in spite of our ignorance of the intrinsic nature of cognitive states. We deal with such states indirection, in terms of a rather remarkable conceptual apparatus, our system of *contents*.

Contents are a system of abstract objects, usually propositions, that philosophers and others have developed to model and systemetize our practice of describing cognitive states by the conditions in which they are true, in the case of beliefs and other doxastic attitudes, or satisfied, in the case of desires and other pro-attitudes. A system of propositions will encode such conditions, given a scheme of interpretation. For example, the ordered pair of the property of being an orator and Cicero, might constitute the proposition that Cicero is an orator, where the scheme of intepretation is that the proposition is true if the second object instantiates the first. Propositions encodes the truth and satisfaction

conditions of attitudes that are reported with utterances like ``She believes that Cicero was a Roman Orator," or ``He hopes that Cicero was a Roman orator (because that's what he said on his exam)" and the like.

I say that this is how we describe and classify beliefs states, but these little perturbations of the brain clearly do not have their contents intrinsically. Their contents depend on their relation to other perturbations, the way the whole system functions, and its historical and practical connection to the external world. Properties like *believing that Santa Cruz is east of Berkeley,* or *wanting to help relieve one's arthritis*, are not fully determined by the cognitive states of the person of whom they are properties, but also involve circumstances that the person is in, that relate him to the things his thoughts are about, and to other thinkers and language users.

Returning to Clinton, imagine him to be at a barbecue. He sees a man at a grill in the distance flipping burgers. His brain changes in a certain way. He acquires a belief, which he might express with ``That man is making hamburgers." What he believes, the content of his belief, is that a certain man, the one he sees, is making hamburgers. The intrinsic nature of his cognitive system doesn't determine which man the belief is about and which proposition serves as its content. That depends on who it is whom Clinton sees. The little change in his brain makes an incremental contribution to the state of Clinton's cognitive system, which itself occurs in a certain brain, of a certain person, on a certain patio, with eyes pointed in a certain direction, and a certain man flipping hamburgers in the distance. When we use a singular proposition to classify Clinton's cognitive properties, we are not directly describing his cognitive states, but describing them given external factors, such as whom he is looking at.

Thus the little change that Clinton's new cognition makes combines with all sorts of other facts, to make it the case that Clinton believes what he does, namely, that a certain man, the one he in fact sees, is making hamburgers. So we have here a state of Clinton's brain, caused by perception, and causing various

things, perhaps a smile at the anticipation of eating a hamburger.[3] The content of the belief depends on the relations of this state to other things inside of his brain and a lot outside of his brain. It is a belief about hamburgers because it involves the Clinton's idea of a hamburger; it's about a certain chef because he is the person Clinton perceived, and so on. The content of Clinton's belief is not in his head, although it is the content of something that is in his head. This I take it is the lesson we should learn from ``externalism".[4]

To emphasize a possibly counterintuitive aspect of the view I advocate, the contents of ideas and beliefs are *not* intrinsic to them. Cognitions, such as beliefs and desires, have the contents they do in virtue of the causal roles that belong to states in virtue of their types, and in virtue of particular connections they have to objects outside of the system of which they are a part. My belief that water is a liquid is true because H_20 is a liquid; it has the content that H_20 is a liquid. If things were as Hilary Putnam imagined, and I was on twin earth, I could be in the same state, but its content would be that XYZ is a liquid (setting aside problems that result from Putnam's choosing water, one of the main things of which I am composed, as an example.) If my basic ``operating system" were different, then the same state might have different causes and effects in my brain, and have a different content.

So, I take that beliefs and other cognitions are modes, particular states of brains, parts of a system of cognitions, each with intrinsic properties and systematic connections to other cognitions. The system as a whole and its parts have extrinsic properties, most importantly those which derive from the particular chain of events that causes them, the particular chains of effects that they cause, and the effects they would have in various counterfactual situations. Content properties give us a powerful, if indirect, way of describing beliefs and other cognitions largely based on their extrinsic properties.

---

[3] This example is apparently out of date, since, according the *The National Enquirer*, Clinton has become a vegan.

[4] See [Kripke 1980], [Putnam 1975], [Burge 1979].

INFORMATION GAMES

I seek an account that tells us three things about self-beliefs: their causal role, their content, and the connection between the two. That is, I want an account that relates a self-belief's incremental contribution to causal properties of the cognitive system of which it is a part, and its incremental contribution to the contents of that cognitive system of which it is a part, and explains how and why these fit together.

Let's start by thinking about the function of beliefs and cognitions in general. It seems pretty clear that an important function of perception and belief formation is to help guide action. That is, if things go right, we pick up information in perception that tells us what we have to do, to act successfully. What counts as success depends on our beliefs, desires, and needs.

So, for example, a chicken needs to eat seeds and bugs and similar things to survive, and if it hasn't had any food for a while, it will be hungry and desire to get some. It has eyes that are very good at spotting seeds, and the ability to peck at the spot where it sees a seed. When it sees a seed its cognitive system, such as it is, changes, and as a result of this change, it pecks at that spot, and gets nutrition.

This is an example of what I call an *information-game*. In an information game, a system gets information, typically about the world around it, and then uses that information, in the sense that it acts in ways that will be successful, given the information. What characterizes different information games is the relations between the agent, time, and place where the information is picked up, and the agent, time and place where it is used. The chicken's game I call the ``straight-through" information game; the agent, time, and place are pretty much the same for the pick-up and the use. The chicken learned *that there was a seed in front of it*, and it engaged in an action what was successful, for the purpose of getting nutrition, in virtue of the fact *that there was a seed in front of it*. I think the concept of an information game helps us see how different sorts of doxastic states fit into human activities, in a way that illuminates the structure they have, and

our rather amazing ability to indirectly characterize them, with our system of contents, in useful ways. The system depends on attributing a certain local rationality, even to chickens. We think that the contents of their inner states will mesh with their causal roles. Basically, we attribute pro-attitudes and doxastic states in a way that makes sense. We see the chicken as motivated by pro-attitudes and doxastic attitudes of a primitive sort. The effect of a pro-attitude and a doxastic attitude --- the actions the chicken performs --- will promote the satisfaction of the pro-attitude, if the doxastic attitude is true. The chicken will get something to eat, if there is a kernel of corn or a grain of millet at the spot where it perceives one to be.

People engage in far more complex information games than this one. Like many animals, including chickens I suppose, we track objects, and accumulate information about them. A thief watches me take money from the ATM; he picks up the information that I have cash on me. Then he tracks me, looking for an opportune moment to mug me and take the money. As he tracks me, he may accumulate more useful information; where I put the money; how fast and strong I am (or am not), whether I am carrying any weapons. He accumulates information about the object he is tracking, and when he strikes all of this information, obtained at different times, may be brought to bear on how he acts.

A third kind of information game I call ``detach and recognize''. It is of great importance in human life, although probably not uniquely human. Perhaps rather than tracking me, the thief merely notes some identifying characteristics about me: what I look like and what I am wearing. Later in the day, in a dark and secluded spot, he recognizes me as the person who took cash from the ATM; he infers that since I had a lot of money, I likely still have some, and this information motivates his attack. He picks up information about me at one time; he breaks off his perception of me --- that is what `detach' means--- then later he uses some of the information he picked up about me to recognize me, and other bits of information to guide his action; because I had cash, I probably still have cash, and am worth mugging.

Finally, an even more characteristically human information game that depends on the detach and recognize game: communication. The thief sees me walking from the bank down Hamilton Avenue towards my home; he phones a confederate lurking at the corner of Hamilton and Guinda, and tells him what I look like and that I am loaded with cash; the confederate waits until he sees a shabby, grey-haired, tweed-jacketed pedagogue, and then mugs me. Information picked up at one time, by one person, is *applied* --- that is, guides the action of --- another person at another time.

In all of these information games we can distinguish between what I call the *source* and the *applicandum* of a game. The source is the object information is picked up about. The applicandum is the object the action guided by the information is directed at. In these stories, I was both the source and applicandum; I was the object the original thief picked up information about, and the object that was eventually mugged by one thief or the other.

But clearly, things don't always go as they are supposed to; the source isn't always the applicandum. Some other grey-haired, tweed-coated shabby pedagogue who hasn't been to an ATM and is in fact quite penniless may make it to the corner of Guinda and Hamilton before I do; the second thief may mug this fellow and come up penniless. In this case, I was the source; the other poor fellow was the applicandum. In the detach and recognize information game, our detached beliefs --- the ones not tied to any perception --- have two duties. First they should help us to recognize the source --- ideally insuring that the source is the applicandum. This went wrong in the present example; the second thief's identifying information the second thief had wasn't detailed enough. The second duty is to provide information that helps the agent choose among various actions. The original thug sees what I look like, which will aid in recognition; he sees that I have cash, which provides a motive for mugging me.

Desires are a species of a wider class of ``pro-attitudes,'' including wants, hopes, sudden urges, yens, and the like. Beliefs are also a species of a wider class of doxastic (pro-) attitudes, including perceptions, memories, conjectures,

knowings-how, and the like. The word ``belief'' is sometimes used as a portmanteau phrase for all of these, but I won't use it that way. I'll use belief for the sorts of states humans get ourselves in, that have truth-conditions that depend on what the world is like, but are *not* tied tightly to our position in the world at a given time; they may be caused by perceptions, but are typically retained for later use. Beliefs are connected to our detach and recognize information game, and especially to communication. So chickens have doxastic attitudes but not beliefs in this narrow sense; they see that a seed is in a certain relative position; but (as far as I know) they don't store this information for later use; they act on it or not, and then it is gone. On the other hand, suppose I give my dog a bone and she buries it. She will return to the same spot days later, when the bone is properly aged and cured through the effects of bacteria in the dirt, dig it up and enjoy it. It seems that she has a sort of a belief about where she buried the bone, that remains after the perceptions of the buried bone and the tree it is under are gone.

NOTIONS AND IDEAS

All of our information games are based on what I call *epistemic roles* and *pragmatic roles.* An object plays an epistemic role in our life, when it stands in a relation to us that affords picking up information about it; most obviously, when we can see it, hear it, touch it or in other ways perceive it. An object plays a pragmatic role in our lives, when it stands in a relation to us that affords acting upon it, that is, acting in ways that effect it, or at least in ways whose success depends on facts about it. Basic epistemic roles and basic pragmatic roles are typically linked by natural facts; the things I can see and touch are likely to be things I can have some effect on, or things whose properties may effect my actions. If a chicken sees a grain of millet it can peck it; if it sees a fox, it can try to avoid it.

Nature provides us with basic epistemic and pragmatic relations; technology, in the broad sense, provides us with many more, and in these cases the epistemic and pragmatic are not always closely linked. By knowing a name,

we often have access to pools of information[5] about a person or thing, stored in books, other people's minds, and now files around the world accessible through the internet. For example, if I look up your name in ``Who's Who" I can find out things about you, due to a complex relation I have with you, involving the mediation of language, books, libraries, publishing companies and the like. This relation will enable me to think about you and say things about you, but it may not put me in a position to effect you in any more significant way. I can find out a lot about Aristotle, for example, by consulting various works, but there is not much I can do to him. Modern weaponry is a depressing example of how we can be in pragmatic relations to objects without being able to learn much about them. A sniper that has me in his sights has the ability to end my life, although his knowledge is limited to what he gets from a fleeting glance. Modern world leaders can unleash weapons that will destroy thousands of lives, in spite of abject ignorance about the people who live them.

Embedded deeply in folk psychology and cognitive science is the concept that our beliefs, desires, and other cognitions have structure; that we have ideas, that are involved in a variety of combinations, in many different cognitions. I have beliefs about Stanford, desires involving Stanford, I make conjectures about Stanford, and experience emotions that concern Stanford. My idea of Stanford is a common element of all of these cognitions. Similarly I believe that Harvard is a university, Princeton is a university, MIT is a university, the University of California at Riverside is a university, and so on. My idea of the property of being a university is a common element of all of these. My belief that Stanford is a university has a structure; it involves these two ideas.

For my purposes in this essay a rather crude taxonomy of ideas will suffice. I distinguish between *notions*, ideas of things, and ideas of properties and relations, which I'll just call ``ideas". Among notions, I'll distinguish between

[5] In this essay I use `information' loosely, so some misinformation counts as information. I am using `information' the way David Israel and I use `informational content' in [Israel and Perry 19xxx].

buffers and *detached notions*. Being a buffer is a status a notion has, when it being affected by perception, and shaping volition, in virtue of being associated with epistemic or pragmatic relations. I notice familiar looking person at a party, and form a buffer of them. I notice more things about them, and eventually recognize that it is, say, David Rosenthal, looking a bit older, greyer and wiser than last time I saw him. My notion of David Rosenthal, until then a detached notion, becomes attached to the perceptions I am having; it ceases to be detached and becomes a buffer. If I am certain of my identification, the original buffer ceases to have a separate existence; To ask whether it is merged with my old Rosenthal notion, or uploads its information and disappears, or uploads and remains for reassignment, is to press my theory, (or model, or metaphor, or fantasy) for more details than are provided by the mixture of folk psychology, reflective common sense, and phenomenology that motivates it. (I realized this when François Recanati noted inconsistencies in different presentations of the theory, and pressed the questions.) There is a pretty clear difference, however, between cases where the recognition is tentative and reversible, and the buffer and detached notion co-exist for a while, and cases where it is quite definite (although perhaps, nevertheless, incorrect) and only one notion survives.

Cognitions like beliefs involve notions being associated with various ideas and other notions. The cognitions are about the objects the notions are of. It possible that a single notion, perhaps developed through a series of misidentifications (imagine Rosenthal has a twin), is of two objects, or, to put it more carefully, is related to two objects in ways that would suffice to be of either of them, in the absence of the other. Then we have what a I call a "mess", a technical concept that is developed further in [Perry, 2001].

The basic picture is this. Our beliefs, and other cognitions, provide us with two connected databases, which correspond to the attached and detached distinction alluded to in the phrase ``detach and recognize''. Buffers are perceptual; they are attached; they are of objects that we perceive, or are at least actively keeping track of, so we know where they just were and have expectations

about where they will be found again.[6] As we navigate around the world, we are constantly filling up buffers with information about objects we perceive; objects we see and hear, things we pick up and handle, and so forth. A second database involves detached notions, those we have of objects that we don't perceive at a given time, notions that are typically the results of previous encounters. Links are made between the databases, as we recognize the objects of which we have formed buffers, as the very objects of which we already have detached notions --- or mistakenly think we do. Here is a striking individual at my office door; an intelligent face; a kind demeanor; a strong athletic build; a well-clipped moustache; who could it be? Maybe Omar Sharif? Or is John Fischer? What would Omar Sharif be doing at my office? And wouldn't he look older by now? It must be John Fischer. Information flows from my detached file to my buffer; this fellow I see is the one I know from the past, the one named ``John Fischer''; the one who is chair of the department, the one who seems obsessed with moral responsibility and Stanford football. And information is piped up from buffer to notion: this fellow looks young and energetic. The combined information leads to inferences: being chair, being devoted to Stanford football, and obsessing about moral responsibility, aren't aging Fischer as much as one might have expected.

An important part of the pictures, or pictures, is the concept of *know-how*. We know how to do things involving objects, to achieve effects, conditional on

---

[6] As Recanati has pointed out, I actually employ (at least) two pictures, or metaphors, or models, which aren't entirely consistent. Sometimes I think of detached notions as *file-folders*. Here it is natural to think of the file-folders as having two different uses, as buffers and as detached files. When the student is in your office and the folder is in your hand and you are making notes, or giving advice, it is functioning as a buffer. When it is put away in the file cabinet it is functioning as a detached notion. The second is of a three-story house, with file folders on the third floor, and in-boxes and out-boxes on the first, and the second full of wiring to connect and disconnect the first-floor boxes and the third floor file-folders, when detachment, recognition, or misrecognition occur. On the second metaphor, buffers are not phases of notions but separate entities connected or unconnected to them. Both metaphors emphasize the key distinction, between notions (or phases or notions) that are tied to various epistemic and pragmatic roles (introduced below) and those that are not. For full generality, the file-folder limps badly when we get to relations. The second metaphor is more at home with the idea of linked databases.

what the objects are like. I know how to shake hands, or startle, or irritate, or find out more about, the object in front of me, or to my right, or in the next room. The chicken has know-how; it knows how to peck the grain of millet in front of it, or flee the fox it sees. We have more a more complex repertoire of doxastic states, in virtue of the ability to detach notions and store information, which enable us to play more complicated information games.

SELF-BELIEF

I distinguish three levels of self-knowledge; one can have knowledge of the same fact in all three ways. I've mentioned two of them. Knowledge about the person one happens to be is the sort that Clinton had in the POTUS example, when he knew that POTUS was going to have lunch with the Queen, but before Hillary explained how the term `POTUS' was used. When he realized that he* was POTUS, he had self-knowledge of the second kind of the same fact. The third kind of self-knowledge is really the most basic sort. I call it knowledge from the perspective of a self. This is the sort of self-knowledge that a chicken has about itself.[7] Each time the chicken learns through perception that there is a kernel of corn in front of it, it is learning something about itself, for the chicken is the "it" that the kernel is in front of. But the chicken doesn't need a representation, file, or notion of itself, because it doesn't need to keep track of which chicken the information it picks up is about. All of the information it picks up is about itself and how things are related to it. I say that the chicken is "unarticulated constituent" of the things it knows about itself.

The chicken might have notions of all sorts of things without needing a notion of itself. It might recognize the barn, the farmer, and even a rooster or two. But it doesn't need to think "the farmers coming towards me with an ax; it just needs to think "the farmers coming towards with an ax."

---

[7] Since I said that selves are persons, and that humans are paradigmatic persons, and I don't believe that chickens are persons, talking of the self-knowledge of chickens is mildly inconsistent, but, I think, not irreparably so.

In thinking about this most basic level of self-representation, it is important to keep the distinction between a representational state, and the content of a representational state, clearly in mind. I'm not sure how sophisticated chickens are, but lets suppose that they can distinguish between corn and millet, and prefer corn. Our imaginary chicken goes into state M when it perceives a kernel of millet, and into state C when it perceives a kernel of corn. In state M, it will peck at the millet, unless it is also in state C. In state C, it will peck at the corn whether or not it is also in state M.

Given these causal roles for state M and C, what account of content would mesh? We might plausibly suppose that each chicken has a fixed desire to eat when it is not sleeping, and to prefer corn to millet. Given that, can't we simply say that M represents millet and C represents corn? Or better, M represents millet-in-front, and C represents corn-in-front? But this won't quite do. We want the content of each state to capture the conditions under which the act caused will succeed in satisfying the motivating desire. For the pecking to succeed, it is not enough that there is a kernel millet in front of some chicken or other; there has to be a kernel of millet in front of the very chicken that is doing the pecking, the very chicken that is in state M. If we are talking about Bertha the Leghorn doing the pecking, then the condition under which that pecking will succeed is that there is a grain of millet in front of Bertha the Leghorn.

Here is how I think of it. The content is an abstract object that *encodes* truth-conditions or success-conditions. An object can be a constituent of the content of a state or event for two different reasons. There may be a part of the representation that has the job of representing that that object, rather than some other, is the one that has to meet certain conditions for the representation to be true, or to be successful. Or it may be that architecture of the representational system insures that this object belongs in the content, without an explicit part of the representation having that job. In the former case, I say the representation is *about* the object, and it is an articulated constituent of its content. In the latter

case, I say that the representation *concerns* the object, which is an *unarticulated constituent* of the content.[8]

A simple example of what I am getting at is the needle on a speedometer on the dashboard of a car. It has the job of indicating the miles per hour that the very car in whose dashboard it sits is traveling. The speedometer is set up to discriminate among different speeds, but not among different cars that are going that speed. But if the needle points to `60', for the speedometer to be correct the very car in which it sits must be going 60 mph. It's content, its truth or veridicality conditions, are not simply that *some* car is going that fast. So the car is an unarticulated constituent of the content of the speedometer; the speedometer provides us with representations that *concern* the car.

What makes it the case that a particular car is the one the speedometer concerns, and tells us about, is a relational fact, that the speedometer is mounted in the car, and connected in the usual way to the car's drive train. That doesn't mean that the speedometer somehow contains a representation of the car in question as the car that meets that condition.

But we of course do have a concept of ourselves as ourselves, the same being at different times and in different places. We have an idea, or notion of ourselves. Our thoughts do not merely *concern* ourselves, but are often fully *about* ourselves; we are unarticulated constituent of our thoughts --- at least some of the time. We have a notion of ourselves, which I call "the self-notion." We have thoughts in which we are explicitly represented, by our self-notions, and not just thoughts that concern ourselves like chickens and many other animals and I suppose children up to a certain age. This is the second kind of self-knowledge.

Let's return to Clinton. The first thing to observe, is that Clinton does play a very important role in his own life, the role we call ``self''. One's neighbor is the person who stands in the relation *living next store* to one. One's mother is the person who stands in the relation *female parent of* to one. These are important

[8] See [Perry, 1986]

roles. One's self is the person who stands in the relation *being identical with* to one. This is an even more important role.

Identity is an epistemic and a pragmatic relation, and self and epistemic and pragmatic role. There are certain ways of knowing about the person who is one's self, that don't normally work for anyone else; I call them *normally self-informative*. And there are certain ways of doing things to and with yourself, that don't work for anyone else; I call these *normally self-effecting*.[9] The chicken's method of finding out about kernels of corn is normally self-informative. What makes the chicken's perception veridical is that there is a kernel of corn in front of a certain chicken, and whenever it has such a veridical perception, the chicken the kernel is in front of is the chicken doing the perceiving.

I call these methods "*normally* self-informative'' because in abnormal circumstances, particularly of the sort philosophers like to imagine, one may pick up information about someone else in spite of using them. Imagine Clinton looking to see if he has spilled ketchup on his shirt. Compare this with the method Hillary might use to see if Bill has spilled ketchup on his shirt. Clinton lowers his head, the way one examines one's own shirt, the normally self-informative procedure for seeing whether there is a ketchup stain there. Hillary looks at Bill's shirt, more or less at her eye-level, which is normally a way of finding out if the person in front of you has spilled ketchup on his shirt.

But philosophers can invent examples where methods of knowing that are normally self-informative actually provide information about someone else. Perhaps a trick mirror is set up, so that when a person looks into it they don't see themselves, but the person standing to their right. If that person was chosen to resemble the person doing the looking, he might be fooled. Clinton might think he has seen ketchup on his own shirt, when he has really seen it on the shirt of his double. Arguably, some methods are not just normally but necessarily self-

[9] `Normally self-effecting' is shorthand for something like `normally self-effecting, or at any rate productive of effects whose nature depends on one's own properties and relations."

informative, as one's methods for finding out whether one is in pain, or whether one likes hamburgers better than salmon burgers, or, at least, whether one thinks one likes hamburgers better than salmon burgers. When I say a method is normally self-informative (or self-effecting) I don't mean to rule out it's being necessarily so.

Now my suggestion about self-belief is this. We each have what I call a *self-buffer*, that is the repository of information picked up in normally self-informative ways, and that is also the motivator of actions done in normally self-effecting or self-involving ways. Self-belief is belief involving these buffers, which we can call ``self-notions''. A self-notion will always be associated with normally self-informative methods of knowing and normally self-effecting ways of acting, and so it will always be a buffer.[10]

One of the reasons we need a self-buffer is that in addition to normally self-informative ways of getting information about ourselves, we have other methods. Basically these were involved in what I earlier called "knowledge about the person one happens to be," like Clinton had before he knew what `POTUS' meant. If we pick up knowledge about the person we happen to be, and recognize that we are that person, such knowledge will be integrated into our self-notion, and be an instance of the second kind of self-knowledge. (In such a case I say that the person still has knowledge of the person he happens to be, but no longer *mere* knowledge of that sort.)

Perhaps Clinton ran his office like this. Each morning a schedule was printed out for Clinton and his senior staff, a grid with the names on the left and the hours of the day across the top. Along about 4 p.m. Clinton would glance at his copy to see where he was supposed to eat dinner. He looked for his name --- or perhaps, once he has learned the convention, ``POTUS''--- in the same way he

[10] The self-buffer won't always be associated with the same methods of knowing and acting. Injuries or philosophical imaginations may deprive a real or imagined agent of many of the normal methods, but not, as far as I have been able to imagine, all of them at once, without simply putting the agent in a state where obtaining knowledge or acting in any way is impossible.

might have looked for someone else's name, if he wanted to find out where they would be in the early evening. If he saw that Clinton was to attend a barbecue at some rich person's house, sharing a limo with some other staff at 5:30, he realized that that means that he* was to attend the barbecue, and he* needed to get to the limo. He didn't find out this fact about himself in a normally self-informative way; he found out about it in the same way any member of his staff would have, by finding the name `Clinton' and looking the events along the row to the right. Another example: suppose Clinton misplaced his cell-phone, and decided to call himself on another phone to make his cell phone ring. If he hadn't memorized his cell-number, he could have looked it up on the staff directory the same way he would have looked up anyone else's phone number. In these sorts of cases there is no use of normally self-informative methods; we find information about ourselves in the same way we find out information about others. Still, we associate this information with our self-notions. Once he learned his cell-phone number, Clinton thought of it as his* cell phone number.

A second reason for having a self-buffer is when we pick up information about ourselves by observing other things like us. I see that other people are removing their shoes before going through airport security. I generalize that everyone must remove their shoes. I need something like a self buffer, to instantiate this generalization to myself.

So far then, the idea is that we have self-notions, which are buffers, in the sense that they are tied a epistemic/pragmatic role (self) that a certain individual plays in our lives, and allow us to think *of* the person who plays that role in a certain way. And we also have information about ourselves, that was acquired in the same way we acquire information about others, from publically accessible sources of information, such as our names, social security numbers, telephone numbers, addresses, office numbers and the like.

Typically, all or most of the second kind of information will also be associated with our self-notion; most of us know who we are, and recognize, at least most of the time, when we are reading about ourselves or hearing about

ourselves. There are exceptions, cases of amnesia and dementia, for example. In such cases, someone might be wrong about who he is. Some people think they are Napoleon. I admire John Searle a lot; I often fantasize about being him, and approaching life and philosophy with the intelligence, confidence and energy that he does; perhaps, as I slip into senility, these fantasies will give way to delusion, and I'll come to believe I am John Searle, and start showing up for lectures he is scheduled to give. And there are other sorts of examples, where we have been assigned names, or something like names, that we don't necessarily know. A bank employee maybe looking over a list of seriously overdue accounts, and see his own account number without recognizing it, and form the belief that *that* person had better pay their bill, without believing that he* was delinquent.

But most of us have it right about who we are. And herein lies a great difference between self-notions and most other buffers. Unlike most notions, self-buffers don't need to be detached from the epistemic and pragmatic roles they are at associated with, as time passes and circumstances change. I can accumulate information about the person I'm looking at, as I track her through a lecture; but when the lecture ends and we go our separate ways, the information in my buffer need to be detached, either uploaded into a pre-existing notion of the person, if I recognize her, or used as the basis for a new notion. But I don't have to do that with my self-notion. It can remain tied to the epistemic/pragmatic role of *self*. Once you have the connection right, it's good for life.

In the general case, I need ways of recognizing the objects I am currently in epistemic and pragmatic relations to, in order to get the new information I pick up to the right notion, and apply the information associated with that notion to the objects. I need to have someway of getting the person I see at my office door associated with my `John Fischer' notion rather than my `Omar Sharif' notion. But in the case where the buffer serves the purposes of the detached notion, as a place to accumulate knowledge about the same thing, the need for such modes of recognition vanishes; this is the basis for the ``immunity to misidentification'' that we consider in the next section.

The self-notion isn't unique in this respect. I have a buffer for the planet I live on, the one I find out about by looking around. I have a file for the Earth; I can read about it in the Encyclopedia. The buffer and the file are one, and can safely remain so, until interplanetary travel becomes not only possible but affordable. And then there is my ``this universe" file, and my ``this possible world file". As Ken Taylor once put it, if anything is ``immune to error through misidentification," it's the actual world.

ROLE-LINKING AND IMMUNITY

In "Self-Reference and Self-Awareness" Sydney Shoemaker argues for the thesis of *immunity to error through misidentification relative to the first person pronoun*. [1968/1994]. If, while wrestling, I perceive a bloody foot and think "I am bleeding," I may be wrong, because although someone is bleeding, the someone is not me. In contrast, if I see a canary, and, on the basis of visual experience say ``I see a canary," it will not turn out that I was right that someone was seeing a canary but wrong that it was me doing the seeing. In the bleeding case we don't have first person immunity, in the seeing case we do.

On my account of self-belief, when I see a canary, I find out , in a normally self-informative way, that that someone is seeing a canary and incorporate the condition of being one who is seeing a canary into my self-notion. I don't have to figure out who is seeing the canary, in any sense that would involve matching features of the person seeing the canary with some identifying information associated with one of my notions. It is the way that I know that someone is seeing a canary, and not features I notice of that someone, that ensures that it is me.

I might think it *possible* that someone else is seeing a canary. Perhaps my visual centers have been wired to someone else's eyes, so I have the visual experiences that person has; he is the one who is seeing a canary; I am merely seeming to see a canary. In the throes of such doubts, I might restrain myself from an otherwise automatic reaction, and not believe that the someone I have

discovered to be seeing a canary is me. Given such doubts, I won't use criteria of identification to look for sure signs that it is me, but rather check to see if there is any evidence of someone playing tricks on me---feel my scalp for new electrodes, perhaps.

Even if I can't eliminate the possibility of such skullduggery, I will believe that I am *seeming* to see a canary. It seems that my experience is not simply a normally self-informative way of finding out that someone seems to see a canary, but a *necessarily* self-informative way of doing so.

We seem to have three sorts of situation. If I see someone else looking at a canary, my perception by itself only gives me a role-based way of identifying the person: the person I'm now seeing. To get the information about the person associated with some detached notion, I'm going to have to figure out who it is. I'll need some method of identification, some criteria, in order to do this. If I figure out that it's my neighbor Dick Grote who is seeing the canary, it will be because of some set of more or less identifying conditions I associate with Dick Grote, and see exemplified by the canary seer.

In the second sort of situation, I see a canary. So I know the normally self-informative way, that someone sees a canary. Here I don't need to identify who it is, or at least to do so I don't need anything like a set of identifying conditions to connect my notion of myself to something I notice about the canary-perceiver. The only person who's canary seeing I find out about them that way, at least in anything like normal circumstances, is me. If I'm worried about it, I'll check to see the circumstances are normal, rather than searching for some telltale sign of me.

The third sort of situation is when we have necessarily self-informative methods. If I find out that someone has a headache, or seems to see a canary, or is thinking about the futility of existence, in the normal ways that one finds out about oneself, there really seems to be no possibility that one is mistaken that it is oneself that is having the headache, or seeming to see a canary, or thinking.

In the last two cases we have something suitably called "immunity to misidentification." It's absolute in the last case and conditional in the second case. But in neither case does one need an identifying description or anything like it to assign the property to its rightful owner, oneself.

There is nothing mystical going on here. The *primary* way we identify objects is by the role or roles they play relative to us, in virtue of the fact that they affect us in ways appropriate to object in those relations. If there is more than one object that can play that role in our lives, and if it matters to us which one it is, *then* there is an issue of further identification, and the possibility of a mistake. The chicken doesn't care which grain of millet it sees and pecks, and has all the information it needs in order to peck it, simply in virtue of seeing it. I care a lot about who is having the headache, but if I know about the headache in the way that people who have headaches know of them, the issue is settled. I am the only one whose headaches I can find out about in that way.

Rosenthal says,

> When I assert that you are in pain, I must identify the self I ascribe pain to in some way independent of your simply being in pain. But there is no similar need, according to Shoemaker, to identify the individual I take to be in pain when I assert that I am in pain or am aware of myself as being in pain. But it is far from clear that this is so. For one thing, when I am aware of myself as being in pain, I identify the individual I take to be in pain as the individual that is aware, in a distinctively first-person way, of being in pain. Such identification is very thin, and doesn't take one far. But if I am aware of myself as being in pain, it does distinguish the individual I take to be in pain from you, as well as any other person. (21)

I think this is confusing in a couple of ways. First, we need to distinguish between asserting and being aware, and the relevant concepts of identification. I don't think it is part of Shoemaker's concept of immunity that I don't need to identify the individual I take to be in pain, when I assert that I am in pain; certainly in the normal case of communicating with someone I'll have to do so,

most often by using the word `I'. With assertion, as Rosenthal says, I need to identify myself when I'm talking about me, just as I need to identify you when I'm talking about you, if I want to be understood. Here identifying means making known to my interlocutor.

Nor does Shoemaker say that when I recognize that I am in pain, I don't identify myself. The point is that the identification is immune to error, not that doesn't happen. Unlike chickens, we have self-notions, and when we are in pain we usually believe that we are in pain, and not someone else. We know that the person whose pain we are aware of in a certain way is ourselves. What we don't need to do is ascertain this by checking the features of the person known to be in pain against some description or set of critiera, which would create the possibility of mistaken identity.

When I am driving, and look at the speedometer, I identify which the speedometer is telling me the speed of; that is, the car I am driving. This is a normally car-one-is-in method of ascertaining how fast some car is going. And normally, the identification is automatic; it's a nice feature of cars that the speedometers in them tell us about the speed of the cars they are in; it's a feature we are very used to, and don't need to consciously worry about. Now for certain purposes, I might need to *further* identify the car. Maybe I have a bet with my wife that my 96 Mazda Protégé can't go over 70 miles per hour. I see from the speedometer, much to my surprise, that I am going 75. To be sure I've lost the bet, I need to double check that the car I'm in is the Mazda, and not my wife's newer are perkier Hyundai.

Similarly, if you and I have a best that the oldest person in the seminar can't figure out the square of 37 in his head, and I somehow manage to do this, I can be sure that it is me that has done it. I might be wrong that my solution was correct, but I can't be wrong that I was the one who arrived at it. But I won't thereby know that I have won the bet, until I check that I am the oldest. But that's why Shoemaker says, immunity to misidentification relative to the first person. My thought, ``I have figured out the square root of 37 in my head,'' is immune to

an identicatory error, when its basis is being aware of just having done so, but my thought, ``The oldest person in the seminar has figured it out,'' is not immune.

Suppose I hurt my hand and have a sensation of pain. I also have a higher-order thought, or HOT, about this sensation. Now it seems that a sensation is an event, which can be referred to in a variety of ways. Call my sensation `SJP'. Now that I've named it and told you about it, you can think about SJP. You may wonder what caused it: Did he go to sleep? Did he hit his hand with a hammer? Just how painful was SJP? And I can also think of my sensation by this name. Maybe years from now I'll read this paper, and have some thoughts about SJP: was that the sensation I had when I accidentally drove a staple into my thumb? Or was it the time I was eating too fast and nearly bit off the end of my middle finger?

These are not the kind of HOTs that Rosenthal is thinking of, of course. There's nothing about such thoughts that could confer special status, like being conscious, on my sensation. The kind of thoughts Rosenthal has in mind are the sorts of thoughts we can have about a sensation or something else going on in our own minds at the time it occurs, and only if we are the very same person that has it. I think Rosenthal is correct, that this sort of thought is what he calls ``essentially indexical,'' which mean simply, I think, that it is what I call a ``self-thought,'' a bit of self-belief. And I think he is correct that this requirement does not present a telling objection to his theory of HOTs. But I don't see that the account of the fact that HOTs are self-thoughts requires one to abandon Shoemaker's doctrine of immunity to misidentification.

SELF-REFERENCE

So let's go back to Clinton and the hamburgers. Perhaps the episode begins with a vague aroma of a hamburger cooking on a grill a ways away. As he walks through the crowd shaking hands, he find himself heading in the general direction of the grill, even though he is, in some sense, not conscious that he is smelling of the hamburger, and that it is affecting him. As he approaches the

grill, the aroma becomes stronger, and he becomes aware that he is sensing it, and that he has been edging closer to its source, and that he is hungry, and would like a hamburger. He says to the chef, ``I would like a hamburger, please.''

Here's how I see this unfolding. First, Clinton has a certain kind of sensation, appropriate to a hamburger being grilled someplace upwind, a bit off to the right of the direction he has been heading. I think it is like something for Clinton to have the sensation, even if he is not conscious of it, but some may doubt this, and think that it's being like something to be in the state he is in only happens when he is conscious of the state. At any rate, the sensation, or perhaps potential sensation, has an incremental effect on him; combined with being slightly hungry, and liking hamburgers, and various other auxiliary beliefs, it causes him to change direction. Clinton, I would say, is *attuned* to the fact that this state is a sign of food in a certain direction, and changes direction without the need of conscious thought. In order for someone else's mental state, or brain state, or sensation, to affect Clinton, it seems he would have to have some sort of thought about it; Hillary's desire for a hamburger, for example, would not directly affect Bill, rather he would notice something about Hillary's behavior, or hear what she said, conclude that she would like a hamburger, and go order one for her. But his own brain states do not have to affect him by being represented elsewhere in his brain, for they are already in his brain, and can do so directly.

At some point, being conscious and reflective, Clinton has a HOT about his sensation. I'm not sure that Rosenthal is right that consciousness is a matter of higher-order thought, but this does seem more plausible than thinking that he somehow has a sensation of his sensation, so let's assume that part of his account is basically correct. Clinton has a HOT about his own sensation. This is what Rosenthal calls an ``essentially indexical'' thought, it is one he would express with the first-person; he takes himself to have the sensation, as well as the thought about the sensation.

A brief aside about the phrase ``essential indexical''. I doubt that I was the first to use this phrase, although my use of it seems to have put it into the

philosophical vocabulary. As noted above Castañeda had written extensively about the topic before my article, ``The Problem of the Essential Indexical'' appeared [Perry, 1979]. The term has come to be used by some people for the view that having a self-belief requires the word `I' or some other first-person expression, most notably by Ruth Millikan, who calls this view ``The Myth of the Essential Indexical''[Milliken 1990]. However, that was not my view. The *problem* of the essential indexical was why indexicals *seemed* to be essential for expressing first person thoughts, and how an account of propositional attitudes might deal with that fact. My solution then was that belief-states were more directly associated with *roles* than they were with propositions. Roles were formally analogous to the *characters* that Kaplan had shown were associated with sentences; characters only determined propositions relative to context. I didn't claim, didn't believe, and don't believe that first-person pronouns are essentially involved in having self-beliefs. While my theory of roles and belief states has become quite a bit more complex, for better or worse, that's still the basic idea behind it.

Rosenthal talks about the mental analogue of the first-person. I think that can be rather misleading. Self-notions are similar to uses of the word `I', in that self-notions are notions *of* their *possessor*, while uses of the first-person are references *to* the *speaker*. But self-notions are mental structures that are components of belief-states, while utterances of `I' are bits of purposeful activity. Self-notions are about their possessors because of their causal and informational role, not because of convention, as is the case with `I'. `I' is a basically tool for communication, but the self-notion is a tool for organizing information and guiding action.

Back to the main thread. While Rosenthal thinks the HOT must contain something like a descriptive identification of their possessor, in order to be about the possessor, I think that is dubious. In general, if we believe in something like the principle of sufficient reason, and don't think that aboutness and reference are built into the basic structure of the world, we must suppose that there is some

relation between a mental event or structure and the object it is about, and between an utterances of a term and the object it is about. Accepting this bit of common sense is not the same as accepting a descriptive account of that relation. A descriptive account holds that the relation is mediated by some set of beliefs, and fact about who or what fits these beliefs. Attunement to the relation that our self-notions have to ourselves, or our perceptions have to the object they are of, does not require belief or thought about the relation; it requires know-how, and not knowledge-that. This is not the way that descriptions work. The relation is not mediated by our beliefs about who we are. It might be correct to say that when I see the cup, and reach out for it to get a drink, I identify the cup as being in front of *me*, and reachable by *me.* But this sort of identification isn't like the case in which I identify a certain person as the oldest man in the seminar, or a certain cup as being the one I bought at the Grand Canyon. And, in particular, it seems immune to misidentification in the way Shoemkaer says. If I see the cup, and take it to be me that is seeing it, I can't be wrong about that.

To get back to Clinton. We have him now in front of the grill, having sensations of hunger and hamburgers cooking, aware of these sensations, thinking about them and what to do about them, and this all leads him to his utterance, ``I'd like a hamburger." Clinton knows English, and knows that if he utters this sentence, it will be true if and only if he would like a hamburger. And, like most of us most of the time, Clinton knows how to produce the utterance. He knows that an utterance he produces intentionally will be his utterance, and that a use of `I' in it will refer to him. Normally, he will be aware of the characteristic changes in his mind and mouth involved in producing an utterance, and know that it is his mind and his mouth that are giving rise to it. Using the first person is a normally self-effecting way of acting, and so under the control of our self-notion. So Clinton knows how to produce an utterance, using the first person, that expresses his beliefs and desires about himself.

Why, one might ask, does he say ``I would like a hamburger," rather than ``Bill Clinton would like a hamburger"? According to referentialist theories of

indexicals and names, the two utterances would express the same proposition, that Bill Clinton would like a hamburger. Granting this, still, the plans that he would have in mind, for getting the hamburger, would be rather different.[11] In the one case, he assumes that the chef will hear the utterance, recognize that the person in front of him is the speaker, and so the utterance is true iff the person in front of him would like a hamburger. He knows how to give a hamburger to the person in front of him, and will do so. If Clinton had said ``Bill Clinton would like a hamburger,'' Clinton's plan would have put an additional cognitive burden on the chef: he will hear the utterance, realize it is true iff the person the speaker is referring to with ``Bill Clinton'' would like a hamburger, recognize that the person in front of him is the person named ``Bill Clinton'' that is most likely being referred to, and hand him a hamburger. This second plan is less modest than the first, in that it assumed the chef will recognize Clinton.

Of course, the chef likely will recognize Clinton. Still, it seems a bit pompous to refer to oneself with one's name, when the fact of who one is, and what one's name is, is irrelevant to one's goals, and Clinton, at least for a politician, isn't very pompous.

Thus our theory enables us to account for an insight of Ellen Goodman's.[12] At a political rally in Philadelphia, while campaigning against Clinton in 1996, Bob Dole said,

> If something happened along the route and you had to leave your children with Bob Dole or Bill Clinton, I think you would probably leave them with Bob Dole.

Goodman commented,

> I am not at all sure that I'd want to leave my children with someone who talks about himself in the third person

[11] For (a lot) more about such plans, see [Korta and Perry forthcoming] and [Perry 2001]..
[12] Ellen Goodman, *Boston Globe*, 1996.

REFERENCES

[Burge 1979] Burge, Tyler. 1979. Individualism and the Mental. *Midwest Studies in Philosophy* 4: 73-121.

[Castñeda 1966]. Castañeda, Hector-Neri, 1966. He: A Study in the Logic of Self-Consciousness. *Ratio* vol. 8 (1966): 130--57.

[Giberman, 2010] Giberman, Dan. 2010. *Glop Theory: A New Trope Ontology* Stanford Doctoral Dissertation.

[Korta and Perry forthcoming] Korta, Kepa and John Perry, *Critical Pragmatics*, Cambridge University Press, forthcoming,

[Kripke 1980] Kripke, Saul. 1980. *Naming and Necessity*. Cambridge: Harvard University Press.

[Kaplan 1989] Kaplan, David. 1989. Demonstratives. In Almog, Joseph, John Perry and Howard Wettstein, *Themes from Kaplan*, New York: Oxford University Press, 1989.

[Locke 1694] Locke, John, 1694. *Essay Concerning Human Understanding*, 2nd edition, London.

[Milliken 1990] Milliken, Ruth. The Myth of the Essential Indexical. *Nous* 24.5 (1990) pp. 723-734.

[Perry 1979] The Problem of the Essential Indexical. Noûs 13: 3-21. Reprinted in (3).

[Perry 1986] Perry, John. 1986. Thought Without Representation. *Supplementary Proceedings of the Aristotelian Society* 60: 263-83. Reprinted in (3).

[Perry 1998] Perry, John, 1998. "Myself and I." In Marcelo Stamm (ed.), *Philosophie in Synthetisher Absicht*. Stuttgart: Klett-Cotta: 83--103. Reprinted in Perry, 2000.

[Perry 2000] Perry, John. 2000. *The Problem of the Essential Indexical and Other*

Essays. 2nd, Enlarged edition. Stanford: CSLI Publications, 2000.

[Perry John] Perry, John. 2001. *Reference and Reflexivity*. Stanford: CSLI Publications. 2nd edition forthcoming, 2011.

[Perry 2002] Perry, John. 2002. *Identity, Personal Identity and the Self*. Indianapolis: Hackett Publishing.

[Perry forthcoming] Perry, John, 2011. ``On Knowing One's Self. Forthcoming.

[Putnam, 1975] Putnam, Hilary, 1975. The Meaning of `Meaning'. In *Mind, Language and Reality*, *Philosophical Papers,* vol. 2. Cambridge: Cambridge University Press: 215--271.

[Rosenthal forthcoming] Rosenthal, David. Awareness and Identification of Self. This volume.

[Shoemaker 1968] Shoemaker, Sydney. Self-Reference and Self-Awareness. *Journal of Philosophy* (1968): 555-67.

[Spinoza 1677] Spinoza, Benedictus de. 1677. *Ethics.*