Norbert Wiley, Sociology Professor Emeritus, University of Illinois, email: norbert@redshift.com

Theory of Mind: A Pragmatist Approach*

Abstract. Cognitive science has been investigating the theory of mind for several decades, and the discipline of sociology seems to have neglected this field. But pragmatist sociology has actually been pursuing this issue for a long time, though it has been calling it the sociology of mind or the sociology of the self. In this paper I sketch a pragmatist theory of mind, using the concepts of reflexivity, role-taking, dialogue and interaction. So, to the existing three theories of mind (the theory theory, simulation theory and phenomenology) I am pointing to a fourth, that of pragmatism.

The theory of mind refers to the way people can identify and understand their own mental states and those of others. These states include thoughts, desires, motives and feelings among other things. For a long time this ability was considered

*This paper will be presented at the annual meetings of the American Sociological Association, August, 2003, New York City. Thanks to Michael Flaherty and Robert Perinbanayagam for advice on this paper and scholarly support.
ordinary commonsense or simply being a human being. It came with being alive and was hardly noticed.

*Introduction*

But when academics starting using Turing’s analogy between computers and humans, and as this analogy became more popular, it became obvious that computers lacked the theory of mind. Computers themselves have no mental states, nor can they recognize the mental states of others. To put it another way, the computer analogy has produced two opposed insights: how computers *are* like people and the how they *are not* like people. The differences are as illuminating as the similarities.

The theory of mind was perhaps first identified in Premack and Woodruff’s “Does the Chimpanzee have a Theory of Mind?” in 1978. This theory was also discussed in some detail in Nicholas Humphrey’s 1983, *Consciousness Regained: Chapters in the Development of Mind*, which drew on papers as old as the early 1970s. Another notable statement was Simon Baron-
Cohen’s 1995, *Mindblindness: An Essay on Autism and the Theory of Mind*. The author showed that autistic children have little or no theory of mind, i.e. little understanding of the mental states of others. Baron-Cohen’s writings on this issue began in the early 1980s. So the absence of a theory of mind in computers and, to some extent, in autistic children drew scholarly attention to this previously ignored skill of human psychology.

Another historical development that popularized the theory of mind was the gradual appreciation and upgrading of “folk psychology” (Horgan, Terence and James Woodward, 1985, Davies and Stone, 1995; Hutto, 2008). This psychology was originally, in the 1960s and 1970s, considered mere common sense, in the negative meaning of that term. For example a common put-down was to say that common sense had claimed the world is flat, but science has shown it to be round. (Also heliocentrism.) Folk psychology or the theory of mind was originally thought of as a crude picture of psychological reality, consisting of superficial platitudes.

But gradually it became clear, at least to some thinkers, that
the theory of mind was not only a relatively powerful kind of knowledge, it was also unlikely that scientific psychology would ever discover a more accurate or richer theory. It is the process of mindreading that makes folk psychology interesting. The actual product -- that so and so believes this or that -- might not always be accurate but the process of being able to discern other minds, is uncanny. In fact it became apparent that psychology would probably never produce a first person theory of mental states at all, since these experiences can be understood only by introspection in the self and empathy with others. Again, not all theorists hold this opinion, but a significant number do, including some central to cognitive science.

Explanations of the Theory of Mind

In the field of cognitive science the theory of mind was initially explained with something called the “theory theory.” This was the idea that people used a cognitive scheme, akin to a scientific theory, to discern each other’s mental states and to predict each other’s behavior. This scheme enabled people to understand each other’s “intentionality” (in Edmund Husserl’s sense), that is, beliefs, desires purposes, etc.
To make this argument, the idea of folk psychology was elevated from a pre-scientific and casual interpretive scheme into one which was amazingly discerning and far better than anything cognitive science could produce. To put it another way, what the computer couldn’t do, read minds, the ordinary person could do. The problem was in figuring out how people did this, or explicating the theory people used in exercising this process. (cites).

The theory theory commanded significant attention, but it also drew negative comment. For one thing the methodological individualism has been questioned. This individualism implies that the understanding of the other is achieved by the individual actor, reasoning in a solitary way, and not by two or more people conversing and interacting. It is claimed to be a psychological rather than a social process. A second criticism concerns the extreme mind-body distinction presumed by this theory. Approaches which envision a stronger connection between the mind and body, such as those taken by people influenced by Merleau Ponty (1962), would argue for a more intimate relationship between the two. In addition to these
two problems the theory theory has never explained how mindreading actually works.

Theory of the mind was gradually confronted by another explanation called the “simulation theory.” This was the idea that people imagined or simulated what was in other’s minds. With this psychological device they were able to figure out the mental states of others. In a sense the missing theory turned out to be the simulation process. But it proved impossible to explain convincingly how people could simulate each other’s mental states, and this approach, like the theory theory, was unsatisfactory. The one new strength of simulation was the discovery that mirror neurons “lit up” in the brain when people simulated each other’s behavior, suggesting that simulation was neurologically based. However the mirror data was subject to more than one interpretation. Among other difficulties, it was not clear whether the mirror process was cause or effect of simulation. This theory also had the same two methodological weaknesses as the theory theory.
The phenomenologists, who also have a theory of mind, have an interesting critique of both the theory theory and the simulation theory. Froese and Gallagher state this critique in two parts (Froese and Gallagher, 2012, p. 437). One, as mentioned earlier, is that the two theories of mind ignore social or interactive causes and rely on methodological individualism. The second is that the two theories rely exclusively on sub-personal or unconscious mental processes. Froese and Gallagher show that social processes, particularly face-to-face interaction, can lead to a theory of mind, and that these forces do so, to a large extent, as conscious rather than unconscious processes.

It is especially in social interaction that people open their minds to each other. In particular people’s gestures, intonations, facial attitudes and speech help show their minds. We know people’s mental states, not through a theory or through simulation, but by the way they reveal themselves to us (Gallagher, 2004).

I have now laid out the three approaches to the theory of mind: (a) the theory theory, (b) the simulation theory, and (c) the phenomenological approach. Given that all three of these
have their strengths and weaknesses, I want to introduce the social psychology of American pragmatism as a fourth approach. This discipline -- based primarily on the thought of John Dewey, Charles Sanders Peirce, William James and George Herbert Mead -- has a rich set of concepts. These ideas were especially formed for the analysis of mental states (Franks, 2010, Hopcroft, 2013).

Neo-pragmatists have been developing this theory for a long time, a relatively early statement being Randall Collins’ “Toward a Neo-Meadian Sociology of Mind” in 1989 (with fifteen comments). Earlier there were Blumer, 1969, Cicourel, 1974, Garfinkel, 1967, Berger and Luckmann, 1966, Rochberg-Halton 1986, Shalin, 1984, and Perinbanayagam 1985 among others. Both before and after Collins’ ground-breaking paper there has been continuous research on pragmatism’s theory of mind. But this work has been done with little reference to cognitive science. My paper is, in part, a re-naming of this research. In addition to being a “sociology of mind” pragmatism is also a “theory of mind.”
Pragmatist Social Psychology as a New Approach

I will first point out how cognitive science’s computer analogy looks from a pragmatist point of view. Perhaps the main criticism would be that, although the computer can handle syntax and similar formalizations, it cannot handle semantics or meaning or understanding (Horst, 2011, pp. 123-144, Searle, 1980, Sayre, 1987, Leudar and Costall, 2009, p. 24). In addition the computer lacks emotions and imagination. The inaccessibility of semantics and meaning is merely another way of saying the computer lacks a theory of mind.

The strength of a pragmatist theory of mind is that it can handle meaning, imagination and also emotion (Hochschild, 1979). At present cognitive science cannot reach what might be called the psychological interiority of human life. This is the same realm as Husserl’s intentionality and the pragmatists’ sphere of meaning. There can be no satisfactory theory of mind without an explanation of this psychological interiority. The thesis of this paper, then, is that pragmatist social
psychology has the best means for explicating this interiority and the theory of mind.

Outline of the pragmatist theory

This paper is meant to be only a sketch, a series of themes that I think would be useful for a theory of mind. I will not try to fully use the substantial pragmatist literature that already exists on this topic. My sketch will draw on some of these resources but it is only a first stab, and it will take some time, effort and collaboration to organize the pragmatist theory of mind.

A theory of mind needs to explain two things. How humans know their own mental states and how they know those of others. I will show, as this discussion proceeds, how pragmatism answers these two questions. My outline of the pragmatist theory will rest on four themes: reflexivity, role-taking, dialogue and interaction sui generis. Reflexivity and role-taking are both based on Mead, although many scholars have further developed these ideas. Dialogue will add Mikhail
Bakhtin’s addressivity to Mead’s role-taking, a move first suggested by Robert Perinbanayagam see (1991, pp. 6-7). The fourth concept, interaction as a distinct ontological level, was first introduced by Erving Goffman (1967, see also 1964). Later, his somewhat programmatic insight was clarified and put to use by Anne Rawls (1987). This batch of ideas is only one way of constructing a pragmatist theory of mind. Other ideas in the wings are Dewey’s notion of embodiment (Solymosi, 2011), James’ concept of self-feeling and Peirce’s theory of semiotic.

Before I present the theory of mind as such, I should mention a “two step flow” that is entailed in this theory. The first step is to choose an ontology or theory of being. The major options are materialism, idealism and dualism. The dominant ontology in philosophy today is materialism, this being the position allegedly implied by science, but there are several competing versions of materialism. Actually any ontology can be used to interpret science, since science does not have any ontological implications. But materialism is the one that most people think is implied by science.
Idealism, in contrast, is now quite out of fashion, even though it has certain strengths. It was the dominating philosophical position in the United States in the late nineteenth and early twentieth centuries.

Dualism itself has two major varieties. Substantive dualism means that there are both material and ideal substances. This was the classic position of Aristotelian and Scholastic philosophy, that there are both bodies and souls. But this position too is out of fashion. Instead “property dualism” seems to be the most common dualism today. This is the position that, although all entities may be material, some material things have both material and non-material properties. In particular the human body has material properties and the mind has non-material properties. These non-material properties are consciousness itself and such traits as thought, imagination, choice and feelings. John Heil, the noted epistemologist, thinks that while most philosophers are materialists, most are also property dualists. Heil refers to non-reductive physicalism (i.e. property dualism) as “today’s orthodoxy in the philosophy of mind.” (Heil, 2013, p. 183).
The choice of property dualism, then, is the first step in pragmatist theory. Having done that, the next step is to examine the psychological processes that explain the theory of mind. There is more than one way to do this, given the interpretive play in these concepts. But since I am merely attempting to state a satisfactory pragmatist theory of mind, I think my approach, even if not accepted by all pragmatists, will do the job. I will use the concepts of reflexivity, role-taking, dialogue and interaction,

*Reflexivity.*

The term reflexivity has more than one meaning in social psychology (Archer, 2010), but I want to single out the usage most relevant to this paper. Although Mead sometimes meant reflexivity as inner speech or the internal conversation (Wiley, in process) he usually meant self-awareness or psychological contact with oneself. He defined the self as the reflexive entity (Mead, 1934, p. 136; Blumer, 1969. Pp. 62-64 ). The flexing here, to use a spatial metaphor, is an act which goes out of the self and then bends back in a kind of “U-turn” and returns to
the self. If this were a visual act it would go out to some kind of reflecting device such as a mirror and then go back to the self, thereby visualizing the self or getting into contact with oneself.

Another analogy is that we look out to another person, take their role, and then look back at our self from the standpoint of the other. This is in some ways a better analogy because it is closer to the way, in child development, the baby attains reflexivity. Baby first identifies with the close caretaker, usually the mother, and has something of a merger with her. This merger gradually goes through stages of separation as the baby moves toward autonomy and independence from the caretaker (Mahler, Pine and Bergmann. 1975). When the baby can take the role of the mother and look back at him- or herself through the eyes of the mother, baby has attained an early form of self-awareness or reflexivity. I am skipping over the empirical findings of child development research because I want to describe this process in bold strokes. There are actually several sub-stages and complexities, but these are not relevant to the present, summary description.

The notion of introspection, which means self-awareness or self-observation, is the same as the way I am using reflexivity.
When we have reflexive awareness of, say, a thought or an emotion, we engage in ordinary or first order introspection. If we want to take a closer look at one of our mental states we might take an outside or “meta” position and observe our reflexivity there-from. This might be called second order introspection, because it is distinct from and added to the first order. It is also less than precise. As Sartre said, we can imagine the Greek Pantheon but we cannot count the columns.

The intermittent, historical controversy in psychology over introspection has been concerned primarily with second order introspection. But our knowledge of our own mental states, which is a by-product of having these states at all, is based on reflexivity and first order introspection. If these two meanings of introspection are kept distinct a lot of arguments can be avoided.

Second order introspection is also, as I am defining it, the same as second order reflexivity. In first order reflexivity the self is split onto two parts: a reflecting part (referred to, by Mead as the “I”) and a reflected part (referred to by Mead as the “me”). But in second order reflexivity the first order self is not split. It is whole. Now the two “parts” are the second
order self or “I” at the meta level and the first order self, taken in its entirety, at the first order level.

I have now answered the first question, how we understand our own mental states. We know these states because, as Kant pointed out, knowledge of them is part of our having them at all. When we have a thought or any other mental state we engage in reflexivity or self-awareness.

In addition, having a mental state is a psychological act and therefore a form of agency, a thing we do. The major pragmatists all recognized some kind of voluntary quality in our acts. We may sometimes act habitually or automatically, but associated with these habits is an acceptance or permission. In the case of acts that are not habitual, which we deliberately enact, the voluntary quality is overt and explicit. This volition entails knowledge of the mental acts. As we choose, we know what we are choosing. Given the two processes of reflexivity and volition, we are, to repeat the response to the first question, routinely and continuously aware of our own mental states.
This analysis of reflexivity is important partly because sometimes scholars have claimed that the thermostat or the computer, both of which can reflect on themselves, have the same reflexive powers as the self. In saying this, these scholars are over-reaching the computer analogy. These mechanical devices can never reflect on their entirety. There are always the two parts: the reflecting part, e.g. the physical thermostat or the recursive language of the computer, and the reflected part. The self at the second order can reflect on its first order entirety.

In other words the first order self is not split into two parts, even though the second order self is so split. In contrast the mechanically reflexive devices can reflect on only part of themselves. Unlike the self, they are always split at the first order. Therefore the notion that mechanical reflexivity is the same as psychological reflexivity is an error. The idea that the thermostat or computer is self-aware and is therefore the same as a human being is a fallacious notion of self-awareness.

*Role-taking.*
Role-taking is also one of Mead’s key ideas. It designates a process that is central to the theory of mind. To take the role of another is to empathize in a cognitive manner with them or to step into their mental shoes. Role-taking is essential to the communication process, for it designates how we can establish a meaningful relationship with someone else (Shook, 2013).

It is sometimes thought that humans first become selves and then learn to take the role of others. This two-stage idea is behind many of the excessively autonomous or asocial notions of human nature. Social contract theories and most versions of classical economics are of this character. But for Mead the self is social first and autonomous later. The first acts of reflexivity or self-awareness, as I suggested above, are based on role taking.

When we are infants, our caretaker addresses us with hugs, loving gaze, billing and cooing and all-encompassing feelings. We identify with that caretaker and his or her role-taking of us. In this way we learn to relate to ourselves in the same manner.
We learn to reflect on ourselves by identifying with and role-taking our mothers or mothering one. Initially our role-taking of the mother and our reflexing on ourselves are a single act. Gradually we learn to separate out the way mother role-takes us, the way we engage in reflexivity on ourselves and the way we role-take people other than the mother. It is in this relational matrix that we become selves. The self is inherently social because it is created by a social relation with our caretaker. In addition its internal structure – the dialogue of what Mead called the I and the me and Peirce called the I and the you -- is based on the relation to the caretaker. She (Mom) gradually becomes “me,” that is, the partner in our internal dialogue.

*Dialogue.*

The notion of role-taking can be enhanced by adding Mikhail Bakhtin’s notion of dialogue. For Bakhtin dialogue is not something selves do; it is what selves are. As he put it

*Life by its very nature is dialogic. To live means to*
participate in dialogue: to ask questions, to heed, to respond, to agree, and so forth. In this dialogue a person participates wholly and throughout his whole life: with his eyes, lips, hands, soul, spirit, with his whole body and deeds. He invests his entire self in discourse, and this discourse enters into the dialogic fabric of human life, into the world symposium (Bakhtin, 1984, p. 293)

Bakhtin regards dialogue as intimate, such that we are closely attentive to the psychological traits of each other (Peringanayagam 1991, pp. 6-7). For Bakhtin, when we address another we take into account their entire personality and all the qualities we want to speak to. We talk not only to them but to all of their “states of mind.” In other words mind reading, to Bakhtin, is a normal and routine part of dialogue.

Bakhtin’s analysis of inner speech, along with that of Vygotsky, also has troublesome implications for Chomsky’s linguistics. Chomsky does not attempt to explain ordinary interpersonal speech or speech acts, but the language we use when we talk to ourselves. As he puts it
Language is not properly regarded as a system of communication. It is a system for expressing thought, something quite different... language use is largely to oneself: “inner speech” for adults, monologue for children (2002, p. 76-77).

But Bakhtin and Vygotsky have shown that inner speech has semantic and syntactic peculiarities that resist formalization. Chomsky actually formalizes ideal speech, even though he says he is formalizing inner speech. In addition inner speech is thoroughly dialogical, and Chomsky’s language, insofar as it is speech at all, is monological (Wiley, 2013).

Interaction sui generis.

Erving Goffman worked with role-taking in his notion of interaction as a kind of reality (Goffman, 1967, Scheff, 2006). For him when we communicate we encounter the other in a unique field. When Anne Rawls (1987) called Goffman’s idea “interaction sui generis” she was designating an ontological level, (Wiley, 1994, pp. 137-141) mediating between the individual and the social structure. This level, to return to theory of mind, has emergent properties. And it is these
properties that permit the dialogue and mental closeness between the participants. In addition to the obvious cues, such as words, bodily gestures, dress, facial expressions, changes in voice, etc. the doors can be opened between the two selves. This opening is a characteristic of this level, and it helps explain how two people can understand each others’ mental states (De Jaegher, et. al, 2010)

An analogy is a ritual or celebrating crowd in which everyone shares intense solidarity and knows what is in each others’ mind. It is the crowd’s consciousness and shared beliefs that is in these minds. Goffman thought that all interaction had a ritual quality, which could create a kind of charisma between the participants. A ritual bond is an extreme case of interaction sui generis, but it designates the dynamic of the mind reading process.

What I am saying then, to return to the second question, is that people discover each other’s mental states in the process of interaction. The closer the interaction and the more frequent its occurrence, the more likely this process will work
well. Once interaction sui generis is attained, there is a legitimacy to rendering and expecting some self disclosure. The semiotics or cues of this state already begin this disclosure, and the flow of interaction will usually enhance it.

**Conclusion.**

In this paper I described a theory of mind, a specialty which began in cognitive science several decades ago. This specialty has also been pursued in pragmatist social psychology for a long time, though it has gone on under other names, such as sociology of mind or the theory of the self. I used major concepts from pragmatism to sketch my version of this theory. This is obviously not meant to be a finished theory, nor can I claim to be identifying and using all the conceptual resources of pragmatism. Still I did give answers to sociology of mind’s two major questions. These answers draw on the unique resources of pragmatist theory, so I think it is reasonable to say they make a contribution to the ongoing discussion concerning the theory of mind.

**REFERENCES**


Davies, Martin and Tony Stone, eds. 1995. *Folk Psychology: Readings in Mind and Language*. Oxford: Blackwells


______. In process. Inner Speech: An Unexplored Continent. Currently under review by a press.