Empathy and Analogy: Mindreading and Simulation

Abstract

The argument of this chapter is that mindreading, theory of mind ("mentalizing"), and simulation theory are fundamentally flawed. What is lacking is empathy. It is not that these do not mention "empathy." It occurs as both the target of explanation as well as an explanatory mechanism. The argument is rather that empathy is missing in the sense that the world of emotions, expressions of life, including the emotions, and the humanness opened up by empathy are not satisfactorily simulated by mindreading. But including empathy will not save mindreading which is a flawed and misleading approach from the start. This is a bold statement for which additional argument and evidence is provided. At its best, mindreading is a research program derivative on experiments in false belief (and a set of closely related experiments), mirror neurons, and the extension of the metaphor of simulation to the neurology of the human biocomputer. Paradoxically, this puts the approach of mindreading on the slippery slope to skepticism about the existence of other minds. Just when it seems safe to rehabilitate introspection as a form of data gathering relevant to simulation, the classic conundrum of the multiple meanings of mental concepts between the first- and third-person perspectives erupts. Several solutions are proposed by the proponents of mindreading, none entirely satisfactory. As a result of the skeptical threat, mindreading is unable to take advantage of the power of analogical thinking, since it must avoid reference to the argument from analogy, which is notorious for not being a solution to doubts about other minds. But it is precisely with analogical thinking and computational model building that the power of simulation lies.

The Cognitive Task of Mindreading has Replaced Issues about Knowing 'Other Minds'

The classical philosophical problem of how one knows other minds has largely been replaced with the cognitive task of mindreading - simulating other minds or, alternatively, discovering or inventing an account ("theory") of how one's theory of the other develops over time. Working independently Robert Gordon and Jane Heal introduced simulation theory.[1] It has been further systematized and transformed by Alvin Goldman. While "mindreading" is now one of those dynamic terms that has become, as Kant wrote in a

---

[1] R. Gordon. (1986). "Folk psychology as simulation*," Mind and Language*, 1: 150-70; Jane Heal. (1986). "Replication and functionalism" in J. Butterfield (ed.), *Language, Mind and Logic*. Cambridge, UK: Cambridge University Press. See also "Folk psychology as mental simulation," http://plato.stanford.edu/entries/folkpsych-simulation/

related matter of dialectical illusion, "a battle field of endless controversies," a specific drill down on this approach is useful:

> [S]imulation theory. . . says that ordinary people fix their targets' mental states by trying to replicate or emulate them. It says that mindreading includes a crucial role for putting oneself in other's shoes. It may even be part of the brain's design to generate mental states that match, or resonate with, states of people one is observing. Thus, mindreading is an extended form of empathy (where this term's emotive and caring connotation is bracketed).[2]

This raises several nice points. In its most straightforward form, simulation of the other individual is supposed to enable access to the other's mental life without inference or introspection, relying on "mindreading." In its lower form, simulation involves the immediate transmission of affects via a subpersonal mechanism, resembling emotion contagion, possibly implemented by mirror neurons; in its higher form, simulation requires decoding the expressions of others and the integration of these into a model of the other. At times, the simulation involved comes perilously close to the discredited argument by analogy to the access and existence of other minds. It is the position of this chapter that analogical thinking is what is most valuable about simulation; but that analogy, in its diverse forms, is most useable in the context of an already ongoing social, communal interrelation. In contrast, analogy is of limited value, indeed it is misleading, if used to try to establish the relationship between oneself and the other individual in the first place.

Now the advocates of mindreading would not necessarily dispute this assertion, claiming that no mindreader in his right mind is proposing a global engagement with the discredited, philosophical problem of other minds. However, as we shall see, the latter has an inconvenient way of reasserting itself into the mindreading approach at

---

[2] Alvin Goldman. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford, UK: Oxford University Press, 2006. Note on p. 176 (Goldman 2006) uses "simulative" and "projective" as synonymous. As will be noted, "simulation" is also used for "inference" as well as "empathy." See also R Gordon. (1995). "Simulation without introspection or inference from me to you" in T. Stone and M. Davies, eds., *Mental Simulation*, Oxford.

inopportune points in the implementation of the program. This is not accidental; and, in spite of attempts to ignore it or even cover it up, mindreading is constrained by a history it cannot escape.

Simulation Contrasted with Emulation

Additional background will further clarify the above-cited quotation and fill out the context of the conversation. As an initial matter of terminology: "emulation" differs from "simulation" in that in the former the same goal is sought as in any parallel simulation but the means or other salient aspect of the overall process is varied. The cognitive approach is front and center with the introduction of "replication or emulation." While "emulation" does not play a large role in Goldman's argument, it is never far from the discussion of "simulation" with which it overlaps. In addition to the old and rarely used meaning of endeavoring to excel others in achievement, "emulate" has now acquired a computational use. An emulator is an abstraction layer that allows one kind of software to run on a different computer for which the software was not necessarily designed by providing a real time implementation for the target environment. More on this as well as the relationship with simulation shortly. Even nicer in this quote, the folk psychology of putting oneself in the other's shoes is front and center. "Putting oneself into the other's shoes" is an everyday way of characterizing empathy, though with certain restrictions and qualification. This is as it should be. Ordinary experience and language can be imprecise and often vague; but they are the source from which philosophizing begins and the milieu to which it returns in giving meaning to our conversations, inquiries, and actions.

Ordinary People do not 'Fix on Their Targets' Mental States'

Next a few cautions are in order. Contra Goldman, it should be noted that "ordinary people" do not "fix [on] their targets' mental states." Ordinary people do not have targets – unless they are aiming with a gun, bow and arrow, or pea shooter. Ordinary people greet friends, get together for lunch, share experiences, and talk with one another about family and jobs and aspirations. Ordinary people communicate opinions, have a heart-to-heart talk, gossip shamelessly, tell stories and jokes, argue about beliefs, and express

disbelief. Ordinary people make claims, help others or offend them and apologize, engage in promises and commitments, strive to fulfill them, succeed or fail. But what ordinary people definitely do *not* do is fix their target's mental states.

I hasten to add that there is nothing wrong with "fixing a target's mental state" – and insights may fall out of this hyperintellectualiztion - but it is a clue that we are not dealing with human beings in human interrelations in any ordinary sense of the word.

Second, separating empathy into cognitive, emotive (affective), and caring dimensions, as implied by the above-cited quote from Goldman, is a useful move. "Divide and conquer" can be a powerful strategy for managing a large amount of information. In a separate discussion, I argue that what Goldman is here alluding to as the "caring" dimension of empathy already has a name, "altruism," and it is a confusion to attribute caring (or altruism) to empathy, notwithstanding an interesting relationship and dynamic between the two separate phenomena.[3]

Third, unfortunately, "bracketing" empathy's "emotive" aspects bracket's the better part of what empathy is about. Separating out the emotive dimension throws out the baby with the bath water. Of course, empathy is *not* limited to emotional experiences, either by Goldman or any author with which I am familiar. Individuals empathize with a wide variety of physical sensations, both painful and pleasureable, that others experience. People also have empathy for both diffuse and particular affects such as moods, inclinations, and desires. As far as I know, no one – I repeat, no one - limits empathy to belief – and that is not the case with Goldman. Desire – hunger, thirst, etc. - are allowed outside the bracketing in so far as they provide motives for intentional behavior that can be made the target of the predictive simulations of mindreading. Still, if my reading of Goldman is accurate, and as far as I know, then none of the examples instantiate sexual desire, the will to accumulate power or influence, or any complex conative goal. This is

---

[3] See Chapter ___ on Empathy and Altruism: From Possibility to Implementation. Note that Goldman makes caring a part of empathy, under this reading, whereas I do not.

not necessarily a show stopper, especially since one has to crawl before one walks. However, it is an indication that the level of analysis is closer to crawling than walking.

Low Level Mindreading Sneaks the Emotions Back In

Given Goldman's bracketing of the emotive dimension of empathy, it is surprising to see Goldman explicitly stating ". . . The 'empathy theory' is another label for the simulation theory" (Goldman 2006: 11). And again Goldman writes: "I turn to a principal rival of theory-theory, namely, simulation (or empathy) theory" (Goldman 2006: 17). Straining to maintain a charitable reading, this is merely a tension within the subject matter and not a flat out contradiction in the approach. Goldman's title is "Simulating Minds," and the approach is a practical and eclectic one that allows for a hybrid of simulation theory with its main rival, theory of the theory of mind (or "theory-theory" (TT) for short). Still, one cannot help but be discouraged by the amount of marching and counter-marching. The recommendation is to pick a position and stick to it. My position is that the emotions have to be bracketed by simulation theory because they cannot be simulated.

"Low level mindreading" is an attempt by simulation theory to "save the phenomenon" that cluster around the communicability of affect in the lived, embodied organism – emotional contagion, vicarious experiences, gut reactions, subpersonal changes of bodily position ("body language"). This diverse panoply of affective response is more than reflex but less the consciously intended. "Low level mindreading" is a candidate replacement mechanism that nicely maps virtually all the affective dimensions that "resonate with" others' affects including those with propositional-like content such as guilt, envy, indignation.

Yet, after initially bracketing the emotional aspects of empathy in the above-cited quote, the reader is surprised to encounter face-based emotion recognition (FaBER). "Face-based emotion recognition" forms an essential part of the discussion of "low level mindreading" – e.g., one directly perceives happiness in the smile – and does not work without it (Goldman 2006: 113f.). The face is a good candidate for an emotional "hot

spot." We pick up a lot of affective-laden information in the face. So it is also a good candidate for being a module (the nine criteria of which will be discussed below in a related context), though modularity is not a key point here.

What is a key point is that the mindreading account of FaBER immediately encounters the need to bracket the emotional dimension. The introduction of an empathy-like process of what Goldman calls "unmediated resonance, or mirroring" (2006: 131) changes the meaning of simulation substantially. This results in an "either or" dilemma for mindreading from the point of view of integrating simulation and empathy. Without simulation of the emotions, mindreading goes back to being "telepathy," that is to say, a mystery. So that horn of the dilemma is unacceptable. In this case, Goldman prefers to be impaled on the horn of the dilemma that correspond to empathy:

> Because the model posits unmediated resonance, it does not fit the traditional form of simulation in which pretend states are fed into an attributor's own cognitive equipment (e.g., a decision-making mechanism) to produce a further state. However, I do not regard the creation of pretend state, or the deployment of cognitive equipment to operate on such states, as essential to simulation. I associate that form of simulation only with high-level mindreading (Goldman 2006: 131).

In face-based emotional recognition, there is an immediate resonance with the other individual. When someone greets you with a smile, you just smile back. Let us be clear this "unmediated resonance" is not empathy, but it is the unmediated resonance on which both empathy and other forms of shared affect and vicarious feeling are based. So what has become of simulation?

The traditional form of simulation feeds a pretend state into a possible world box, an idea in Nicols and Stich (2003) that is taken over by Goldman, but that is not appropriate here because the face-based emotional recognition is supposed to be immediate and unrehearsed – and even if it is not, the best way to feign an emotion, is to reenact ("replicate") the emotion itself. So where is the simulation? Is it pushed down to a

computational implementation by means of mirror neurons? That option is not available to Goldman, and he has explicitly ruled it out:

> Let us simply mark the vital distinction between the two senses of *simulation*: *computational modeling* simulation and *replication* simulation. For purposes of the theory of mindreading, only the latter interests us (2006: 36).

In spite of ruling computational simulation out of scope, the discussion has no where else to go. But Goldman is an astute philosopher; and he is diplomatically vague about the capability of discharges of mirror neurons to give rise to empathy (2006: 136). A category mistake between the functioning of the organism and the folk psychology concepts of goals and intentions is pending. Goldman leaves it alone. He leaves it to Iacoboni (2005) to "postulate that mirror systems learn which motor acts commonly follow other acts to achieve a characteristic goal" (Goldman 2006: 140). The problem is not "learns," since that could mean just "encode" in the human biocomputer. The problem is that the chain of neurons now has a "goal." Goldman knows full well that if the goal cannot directly inhabit an individual mirror neuron – like Descartes' thinking substance inhabits the pineal gland – then not even a chain of neurons in customary conjunction will allow the mirror neurons to have the goal. The same applies to replication simulation (alternatively called "reenactment" or "interpersonal" simulation), which, however, is not brought in until "high level mindreading." A chain of mirror neurons remains a powerful computational mechanism by which other organisms may be isomorphically simulated; but they are not a replication simulation where "replication" implies an implicit "reenactment" or "interpersonal simulation" (as Goldman explicitly calls out (2006: 140).

So what about "high level mindreading"? Will that get us out of the impasse? In contrast with "low level mindreading," "high level mindreading" is an operation according to which an explicit operation of "as if" cognition - explicit putting oneself in the other's shoes - is invoked (Goldman 2006: 113f.). "Putting oneself in the other's shoes" is unpacked in terms of simulation. Here we do ascend to a full, robust sense of replication simulation, in which imaginative variations and reenactment will apply further cognitive

processing to pretend states in enactment imagination in high level mindreading (Goldman 2006: 149). We have recovered a sense of "replication simulation," but we are back where we started, leaving the emotions in favor is "an extended form of empathy (where this term's emotive [. . .] connotation is bracketed)."

The bottom line is that simulation does not survive the reintroduction of the emotions, even in the form of the self-contained face-based emotion recognition module. The neurological evidence invites the implementation of a "computational simulation mechanism." But that only gets us to unmediated resonance of the organism, not mindreading as such. The process is a mechanistic and computational one, not an interrelational or human one. Organisms are driven by other organisms to unmediated resonance by processes that invite computational modeling. Intermediate explanatory mechanisms of a functional kind such as "inner imitation" (Lipps) or "shared manifold" (Gallese) may be useful as placeholders for further inquiries. But we are no longer "reading" minds; we are relating to individuals in a community.

Mind*reading* is at its best in interweaving the framework for decoding - reading is paradigm a form of decoding - with the folk (social) psychological framework of belief, desire, and affects (if the latter are allowed). No doubt that "mindreading" as a term and as a research program has been useful in discovering, inventing, and surfacing a wide variety of results. For example, the work of Simon Baron-Cohn in the early identification and diagnosis of autism based on (lack of) eye gaze following, shared attention, and the detection of intention, is valuable and an inspiration, not only to Goldman. Specific behavioral indicators have been useful in the early diagnosis, early intervention, and amelioration of some of the worst side effects of autism spectrum disorders. For this service, we must all be grateful to mindreading. However, the suspicion is that the pendulum has swung too far in the other direction and the emotional dimension is being ignored, not because it is unimportant but, because it does not fit the mindreading approach.

Why the Emotions End up being 'Bracketed'

However, this also gives us a hint as to why the emotions must be bracketed by the mindreading approach. If one wants to make mindreading into a specific module, for example, enabling an individual to pass the false belief test, then including the emotions will blow the module to smithereens. The variety of horizontal phenomena encompassed in the emotions will rule out any simple simulationist account or at least raise the bar substantially. Even the clearly modular face-based emotional recognition routine could not handle simulation. A parallel consideration applies to empathy. Including empathy in mindreading will blow it up in that empathy takes as its primary targets precisely the diversity of expressions of life – sensations, affects, emotions – that are a superset of the emotions that already stretch mindreading.

A closely related reason that the emotions must be bracketed is that simulation, the major method of mindreading, cannot handle emotions pure-and-simple. Yes, I can pretend to be happy; and the pretence is arguably a form of simulation – in mindreading "pretence" is an input to a possible world box as part of the simulation flow chart. But this simulation only gets me pretend expressions of emotions, not real expressions of one. The emotion is the emotion in any case. In order for the pretence to succeed I marshal experiences and recollections that actually invoke the emotion itself. The result is that the attempted simulation of the emotion collapses the distinction between being the emotion and expressing the emotion – the distinction between being happy and expressing happiness. It is the latter that can be simulated; the former just is or it isn't. Indeed the best simulation of an emotion - whether basic, pretentious, irruptive, or moral – as any actor schooled in the method of Stanislavski knows – is the invocation of the emotion itself. The best methods to feign (un)happiness on the stage - or anywhere, since all the world is a stage - is to think (un)happy thoughts, put on an (un)happy face, and be (un)happy, albeit temporarily and artfully.[4]

The obvious reply to the observation from the method acting approach (in recreating an emotion through emotional memory) is precisely an early version of simulation and it

---

[4] See Chapter ___ on Empathy and Expression: Unexpressed Emotions are Incomplete.

Lou Agosta, Ph.D. (LAgosta@UChicago.edu)

deserves to be assessed in its own right. What is clear is that human beings do not need to study method acting in order to express emotions sincerely or convincingly. If anything this is a method that works on the production side of the emotions, the mysterious process of bringing forth an emotion by encoding a inchoate, emerging something = x into a proper expression of emotion, in turn, evoking the process of mindreading, which, however, has been returned to its initial transparency as something we just do without effort or understanding. What is clear is that the practical delivery of spoken lines in a theatre piece in a *mise scene* and the actor's actions and reactions in the lived moment takes the term "simulation" into a context not envisioned by either its advocates or opponents. This is where we must let it be.

Mindreading and Empathy Remain an Odd Couple

As Goldman states, there is indeed an en extended sense of empathy that applies to mindreading minus the emotional dimension. But the two cannot be equated without hyper intellectualizing empathy or making mindreading denser with affect than an analysis of false belief and everyday instrumental desires would warrant - as Goldman understands very well. Mindreading and empathy remain an odd couple.

Mindreading is a fundamentally flawed paradigm for capturing, understanding, and integrating what is most human in human interrelations; and, from the perspective of human relations, mindreading has lost more than it has gained. Mindreading does not do a better job in dealing with the cognitive impenetrability of the emotions than does any other form of belief or cognition. To deal with the cognitive impenetrability of the emotions, empathy is required.

For example, I can readily imagine what the garden looks like from where you are standing, as if I were over there. I can imagine what it is like to believe in the Democratic Party platform if I am a Republican or vice versa. I can even get a start on what it is like to feel thirsty as a parched sensation. But the emotions are tricky. When a person is in the grips of fear confronting a snake or jealous about a romantic relationship gone bad, the world is given globally as dominated by the fear or upset (etc.) constellated by the

emotion at hand. This points directly to the cognitive impenetrability of the emotions. In such situations, the imagined fear or upset – the vicarious experience that is automatically triggered by "low level simulating" - are highly attenuated trace affects of the matter itself. It is a hint or a clue. Not nothing, but not exactly the emotion as such. The cognitive impenetrability of the emotions means that what I know and believe does not influence the emotion; and, conversely, it also means that the emotion resists facile "as if" manipulation based on a superficial and easy kind belief, character traits, or elementary wants such as thirst or hunger.

It is a challenge to conjure up hard-to-control contents such as falling in unrequited love, panic, jealousy, or the need for heartfelt gratitude. I can recall such experiences – if I have actually had them – with the result that I am glad that I am not suffering through them right now. Those who would reduce an emotion to belief and desire and do so without remainder find at first that it seems relatively easy to simulate the emotion – by simulating the associated belief and desire. But then one senses that something is missing. Like translating the poem from one language to another, nothing is lost in translation, nothing except the poem itself. The experience of panic, jealousy, or falling in love is missing. What some would call the "throwness" is missing. The affectivity of the emotion is missing. One must undertake a study to recover the affect-laden quality, impact, and aspect of the emotion – once again, think method acting – in which one tries to recollect reminders of situations that have triggered such feelings, given personal, history, character, and goals, given, one's facticity. If actual experience does not work, then one tries include examples from narrative fiction, film, or the theatre that have triggered vicarious experiences of a qualitatively similar emotion.

To get traction with such a recalcitrant emotion, one must have a strong sense of unmediated resonance, vicarious feeling, and, to coopt Goldman's term, low level empathy (not simulation), with the other individual. That's a start. One must then apply further processing, cross-referencing the immediacy of the vicarious experience with further background and context. No room for simulation here. It will make sense to invoke "as if" processing in formulating an interpretation to be communicated to the

other individual as to what it might mean that he is afraid of flying. The give-and-take will provide a context of distinctions that creates a new possibility for dealing with the breakdown. Though not enough background is at hand to complete the conversation, it might be that a pattern of loss of control – as in passively sitting in the back of an airplane – is available in the individual's experience to make sense of the fear – that is, penetrate it empathically – and create a counter-experience of more control. What that would look like might be working through of the prior experiences and an appreciation that they do not fit the current situation.

But one may object: does this not contradict Kendall L. Walton, (1990) on quasi emotions? According to Walton, one obvious reason people do not run screaming from the cinema showing a horror movie is because they do not experience fear of the green slime, they experience quasi-fear (e.g., Walton: 245). Unlike Walton, the approach that argues emotions cannot be simulated does *not* reduce emotions to belief and desire. Even more importantly, much of what Walton says about "quasi fear" (e.g., Walton 1990: 245) is covered in ordinary English (and in the position) by vicarious emotions. Still, a firm point of disagreement with Walton is that I would assert vicarious emotion is a real emotion, albeit attenuated and generally less intense, whereas Walton would probably deny that the vicarious emotion, affect, and experience are real (if he used the word "vicarious," which he does not). In short, contra Walton, I line up with Freud's account of dreams (and by implication the theatre and cinema) in which the robbers are imaginary (fictional) but the fear is real.[5] I do not get up and run away from the green slime in the horror movie because my would-be action of running is interrupted by precisely that thought. No matter that I no the green slime is a fiction, I get the thrill of experiencing fear without having to worry about the consequences. The fear is real, the green slime is imaginary. Case closed.

Those who have pursued mindreading are enthralled by the punning aspect. Originally (and still in ordinary talk) "mindreading" is synonymous with "mental telepathy," the

---

[5] See Chapter ___ on Empathy and the Expression of Emotions: Unexpressed Emotions are Incomplete where a detailed drill down on this point is engaged.

transmission of propositional content ("thoughts") without the aid of conversation or other physical media, even though the other does not tell me and does not commit what is otherwise called a slip of the tongue or other parapraxis (*Fehlleistung*). This leads to an inside joke. We are all mindreaders, doing the impossible on a daily basis. In turn, this lends an "aura of mystique" to something that even grammar school kids do routinely and that we all do smoothly and easily – relate to people in a familiar and effective way on a daily basis incorporating beliefs, affects, and desires into our motives, evaluations, intentions, and struggles to survive and prosper on a daily basis. At the risk of taking the metaphor of mind reading too literally, it is not minds that we read but expressions of the mental life of persons such as emotions, affects, actions, beliefs, intentions. Naturally, once the individual has access to language, the realm of belief – and false belief – is readily leveraged. Caretakers can stop using prelinguistic forms of empathy to understand what the child wants and can ask, "Do you want orange juice or apple juice?" This aura of mystique goes straight back to the classical problem of other minds and why it should be such a wonder that the other individual fulfills or thwarts my intentions, delights or disappoints me, and supports or contradicts what I assert.

The suspicion is that a fatal misstep has already occurred with the choice of the term "mindreading." I have already missed the human being and have seized the mind, at best a subset of the whole self, as the target of an explanatory, analytic, or foundational exercise for the humanistic studies (*Geisteswissenschaften*). It does not help to substitute "mentalizing" for "mindreading." "Mentalizing" is not a word either. As I recall, when I was ten years old and in the school playground hurling insults at my ten year old buddies, we would ask, rhetorically of course, "Are you mental?" where "mental" was taken to mean "insane," a "mental case." As indicated, "mindreading" is a powerful metaphor that makes us do a double take and causes us to stop and think. How is that we understand other minds anyway? As an approach, mind reading provides us with a new and engaging way of seeing something we do seemingly effortlessly on a daily basis as worthy of a more detailed inquiry, analysis, and explanation. What was familiar and so taken for granted is put at a distance, even alienated, so that it must be brought near and made intelligible at a new level of understanding. This is a useful and important. Yet it is also

misleading in a way that, as we shall see, casts a long philosophical shadow and will come back to haunt mindreading.

Simulation Remains Computational

What is called "simulation" used to be "imagination." Simulation has gained currency in the context of computer programming where various processes are simulated in silicon prior to being implemented in their target context, for example, as when simulating the fluid dynamics of a World's Cup class yacht moving through the waves. Other kinds of simulation include testing air plane wings or the aerodynamics of an automobile in a wind tunnel. War games played by colonels and generals at the Pentagon simulate the behavior of forces in conflict on the battlefield. The most successful examples of simulations have been of physical processes. For example, during World War II computers were used to simulate the range of artillery shells give the size, force, angle, wind velocity, and distance, in advance of actual battle field conditions. The computers were, of course, mostly women left behind when all the men were drafted into the Army, ladies in green eye shades with pencil, mechanical calculators or slide rule. The results were put into tables to be referenced when used by the artillery teams. A similar method is invoked in simulating the fuel needed to complete a transatlantic flight given heard-win, tail-win, altitude (air density and resistance), and related factors.

The reason that the emotive dimension has to be bracketed by mindreading is that it cannot be simulated or, more precisely, can be "simulated" only with the utmost effort and indirection at the computational level. In either case, emotions have been encoded and decoded. Emotions have been feigned or expressed sincerely. Emotions have been understood and misunderstood. Emotions have been expressed or left unexpressed. What has not occurred is the simulation of emotion, because emotions can be expressed or not; but they cannot be simulated, at least not by a human being.

The computer game *The Sims*, on the other hand, is a *simulation* of middle class suburban life in a wide variety of its aspects, including a subset of simulated emotions. The

engaging thing about *The Sims* is the creator puts the characters in various situations and then tries to take care of them. The characters can be influenced but they tend to take on a life of their own and do not cooperate with the demigod-like creator who has defined them. They want material things, are "addicted" to possessions and become depressed when they can't have them. The program "under the hood" is complex enough to allow for random behavior, violation of expectations, and alienation that causes breakdowns in the empathy that the player forms with the family members he has created – Joe Sim, Sally Sim, Jimmy Sim.

In an ingenious implementation of the worst caricature of middle class conformity – "bourgeois" in the negative sense – to superficial norms and crass materialism, the individual Sims pout, throw tantrums, get over-stimulated, become depressed, and generally simulate a wide enough spectrum of bad behaviors and emotions – sad, happy, angry, afraid - to be entertaining from a demigod's eye perspective. Reports from game enthusiasts indicate that playing it can become so intense and demanding – like trying to manage the behavior of an irresponsible roommate - that the game player is almost relieved when the individual becomes depressed, refuses to eat, and dies. Presumably such responses do *not* extend to one's roommate. Future versions of the *Sims* game promise to encode variables in such a way that an authentic relationship between the Sims becomes a possibility. However, no matter how authentic the actions and expressions of emotions, no matter how perfectly synchronized the neurologically hard-wired micro expressions of smile lines and slants, the "emotions" of the Sims will always remain a simulation.

Rolling Simulation Up into Theory of Mind

The other approach to mindreading is to push down simulation back behind a theory of the mind that attributes a schema (or theory) to the mind. We may simulate intermittently and from time to time in putting ourselves in the shoes of others; but the simulation is derivative.

What is essential is the theory of mind that the individual brings to the interaction. At least three theories of the theory of mind ("theory theory" (TT)) exist. The following will make some simplifications and generalizations, but, by and large, aims at being an accurate presentation.

Theory of Mind as Folk Psychology

The first one, common sense or folk psychology, takes the ordinary person's usage of such terms as belief, desire, intention to be a folk psychology theory of the mind. Like much of common sense, it works well enough on a day-to-day basis, but when one looks at the details with methodological rigor, there are inaccuracies and even contradictions. Causes and reasons are routinely conflated, confabulation is common, a ghostlike homunculus shows up. One of the sources of the research program of cognitive social psychology is to build a refined and accurate theory of mental interrelations. This includes the possibility that some mental terms will not survive scientific scrutiny, though it is controversial whether everyday vernacular would be able to (or even want to) eliminate them. It is hard to imagine that "belief" or "emotion" will disappear and go the way of "phlogiston" – the discredited explanatory term for heat – since "phlogiston" was always an abstract, theoretic term and was never used to negotiate and communicate in everyday, ordinary conversation. This is the default theory of mind with which we begin and to which we inevitably return at least until such a time as mindreading research or other science succeed in changing the way we speak and act about human, beliefs, intentions, desires, and emotions.

Theory of Mind as Modularity

In the next version of theory of mind, a specific theory of mind module is postulated, the malfunctioning of which is causally connected under various scenarios with other capabilities such as eye gaze following and joint attention as the basis for diseases of mindreading – also called "diseases of empathy" - such as autism, sociopathy, and related

spectrum disorders.[6] This categorization using "mindreading" is not entirely fair to the participants, and subsequent thinkers, drawing on psychopathology, cognitive science, and early childhood research, have gone directly to the affective context of empathic relations as the matrix without which higher cognition will not even get off the ground, stopping just short of saying that it is the empathy that is the missing module in autism.[7]

Empathy fails as a module – in the "modularity of mind" sense - on several criteria. It is not domain specific, ranging across emotions, sensations, and affects such as moods and desires. Empathy is not encapsulated in that it has access to intermediate results in processing. While many of the emotions that empathy targets *are* arguably encapsulated (and so modular), empathy itself accesses a wide variety of intermediate inputs and representations extending along a spectrum from transparent to fully opaque, the former being vicarious experiences and the latter including imaginary "as if" processing. It exchanges information with other parts of the system at the cognitive, affective, and volitional levels. The operation of empathy is not mandatory in that the individual can "jump start" empathy by imaging what it is like to be the other person with the other person's character traits (when they are known). On the other hand, the speed with which empathy operates - it can be fast, very fast – supports its being categorized as a module. The breakdowns to which empathy is susceptible - precisely the point about autism, prosopanagnosia, and psychopathy – indicate modularity. Finally, the hardwired neural structure (mirror neurons are a strong candidate) suggest that empathy is modular. The depth of the outputs is a point of controversy, depending on where one draws the line between what gets coded as input and what as output, what as empathic data gathering and what as further cognitive processing. If one includes vicarious feelings as output, then the depth of empathy is relatively deep, since such feelings will be further processed

---

[6] Simon Baron-Cohen. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press, 1997. Also relevant, J. Fodor. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press, 1987: 47.
[7] P. Hobson. (2002). *The Cradle of Thought: Exploring the Origins of Thinking*, New York: Macmillan Pan, 2002.

cognitively by empathy. If not, then the output is shallow, being limited to the downstream of the information supply chain of empathy.[8]

The way to break the tie, so to speak, is to realize that empathic receptivity is a strong candidate for modularity in contrast to empathic understanding, the latter escaping the criteria of modularity:

Empathic receptivity is mandatory – the smile just shows up as a smile. Empathic understanding is voluntary – a person puts one intentionally into the shoes of the other and considers the results.

Empathic receptivity has limited access to intermediate stages of processing – the person just "get's it" – the other is angry (etc.); happiness is immediately available in the smile; sadness, in the frown. Empathic understanding can monitor and inspect intermediate results as interpretations about what the other's experience are unpacked into explicit propositions, shared with the other, and iterated in a multistage operation.

Empathic receptivity is fast. "Smile = happy" is an immediate proposition. Empathic understanding takes time as one wonders what is really going on in context.

Empathic receptivity is encapsulated in that other parts of the system do not have access. You just are aware how the other person is feeling. Empathic understanding interacts vigorously with empathic receptivity as well as other cognitive, affective, volitional, and language-oriented inputs.

The output of empathic receptivity is shallow – it is vicarious experience in its diverse forms. I am just aware what the other is feeling because I feel it too. The output of empathic understanding is deep in that it may recruit all kinds of imaginative variations in

---

[8] See Chapter ___ on Empathy and Intentionality for a drill down on the information supply chain that subserves empathic functionality.

its interaction with empathic receptivity, resulting in a sequence of complexly interrelated outputs. Nothing shallow here.

The breakdowns of empathy can occur in both areas – receptivity and understanding – but the latter seem more fundamental, causing a domino effect that also "takes down" the access to and understanding of the other individual along with it.

Mirror neurons support the functioning of empathic receptivity whereas empathic understanding relies on broad linguistic and cognitive abilities in making explicit interpretations and empathic responses. Empathic understanding is arguably prelinguistic, but by the time one wants to formulate an empathic understanding access to a form of language is essential in articulating the possibilities to be interpreted or implemented.

It should be noted that both aspects of empathy benefit from a characteristic pace and sequence of development, the sequence of which is relatively fixed in the case of empathic receptivity whereas empathic understanding follows the overall trajectory of the mastery of language and advanced cognitive ("logical") capabilities, the latter being too broad and horizontally diffuse to be modular.

Empathic receptivity fails the domain specificity as to its input, and, in that regard, does *not* satisfy one of the criteria of modularity. Empathic receptivity is broadly promiscuous in the kinds of inputs it will accept as stimuli that are able to trigger vicarious experiences. As indicated, the spectrum extends from sensations (pain and pleasure), affects such as desires, all kinds of emotions (basic, social pretences, irruptive, moral sentiments), as well as verbal contents and actions. The outputs map to all the inputs as vicarious sensations, affects, desire, emotions. However, empathic receptivity, in spite of failing the domain specificity criteria for being modular, does satisfies the other eight out of nine criteria.

Theory of Mind as the Child Scientist

In the third approach to the theory of mind, the evidence of early learning is marshaled to explain how a theory of mind develops in relation to other people, things in the world, and oneself along the lines of a scientific theory. The "child scientist," the "computational baby," incorporates counter-examples and additional evidence into its (implicit) way of representing the world ("theory") and uses these to evolve the theory in the direction of a better fit to the world of everyday people and things.[9] It is a brilliant insight that as researchers and scientists are observing the human baby in the crib, someone equally intelligent, though perhaps less verbal, is studying them in return, and using the experience to build its own view of the world.

The view of scientific theory building is a caricature of actual scientific practice and model building, though to the best of my knowledge nothing essential turns on this limitation. It does, however, closely resemble the developmental schemas proposed by Piaget with qualification from Vygotsky.[10] Piaget is famous for proposing that intelligence unfolds according to predetermined biological program(s) of assimilation and accommodation with the environment. Intelligence progresses in stages from reflexes, sensory-motor reactions, habits-like configuration, perception-action loops, to sub symbolic operations that are dependent on the milieu, finally culminating in context-free, reversible symbolic (verbal) operations. Piaget is famous for incorporating the depth of a certain version of Newtonian causality, space, time, and number into his approach. So quite a lot good, solid science is embedded in the schemas, even if the methods by which the child arrives at them are only analogically scientific. The point is that, even if the computational baby is a very anomalous scientist, and not a canonical practitioner, strong science-like schemas are available to be transformed in stepwise sequence in the direction of objective knowledge of the world as the infant and child progress through developmental stages.

---

[9] A. Gopnik, A. N. Meltzoff, P. K. Kuhl. (1999). *The Scientist in the Crib: What Early Learning Tells us about the Mind*, New York: Harper Perennial, 2001.

[10] J. Piaget. (1952). *The Origins of Intelligence in Children*, tr. M. Cook. New York: W. W. Norton, 1963. Lev S. Vygotsky. (1934). *Thought and Language*, tr. E. Hanfmann and G. Vakar. Cambrdige, MA: MIT Press, 1963.

Lou Agosta, Ph.D. (LAgosta@UChicago.edu)

Theory of Mind as Cognitive Science

Meanwhile, the proponents of diverse theories of theory of mind (or "theory-theory" (TT) for short) have themselves splintered into opposing camps – those that prefer to build a model of the mind using computational and flow charting methods familiar to the computing lab versus those committed to the evidence of child development. We see that in Goldman's reluctance to embrace computational simulation (2006: 36), the approach that seems most compelling to Metzinger (2003). These are not necessarily contradictory; and the researchers studiously try to accommodate one another's data; but the methods and framework are significantly different – computing versus child development.

Cognitivists acknowledge that nature is a tinker and that the opposition between theory-theory and simulation is ad hoc at best, thereby, enriching their theory – not to be confused with the theory of the child scientist – by means of a modified simulational mechanism. Getting from folk psychology to cognitive science now requires a journey through child development, but the route is different than the destination. Our concepts have a developmental history that they never entirely leave behind. Now surfacing the child's implicit "scientific" theory of the world is required in order to build and transform a computational model ("theory") that accounts for false belief, the possibility of being mistaken.

Under this account, folk psychology itself would be a rich body of information about the mind collected into a rough-and-ready theory readily useable by mindreading mechanisms such as simulation.[11] Cognitivists have incorporated a function of pretence, overlapping with but distinct from high level simulation, and proposed implementing a possible world box (PWB) as the fulcrum about which egocentric behavior gives way to symbolic and functional reversals in mental operations.[12]

---

[11] Shaun Nichols and Stephen P. Stich. (2003). *Mindreading: An Integrated Account of Pretence, Self-Awarness, and Understanding Other Minds*.Oxford, UK: Clarendon Press, 2003: 211. By the way, I have never met a simulationist that denied folk psychology as an input to the simulation process; however, it rarely survives the process and come out the backend unmodified.

[12] "We do, however, think that the argument justifies a strong initial presumption that accurate mindreading processes are subserved by simulation-like processes and that inaccurate ones are not" (Nichols and Stich 2003: 106). See also: "Our entire theory has been elaborated at Marr's task-analytic or computational

After a sophisticated exercise in flow charting ("boxology"), including belief, desire, pretence, possible worlds box, and supporting mechanisms, Nichols and Stich (2003) turn to address objections. They are immediately confronted with a version of the problem of other minds. Nichols and Stich assert: "The mechanisms that subserve introspection are largely independent of the mechanisms that produce beliefs about the mental states of others" (Nichols and Stich 2003: 200). The classic conundrum of the multiple meanings of mental concepts between the first- and third-person perspectives erupts. "I desire some vanilla ice cream" seems to use a different version of "desire" than "You desire some vanilla ice cream." The first person monitoring mechanism fragments away from the desire detection mechanisms targeting the other, the third person. Unity of consciousness is at stake and possibly lost. Nichols and Stich regard this dilemma seriously. Three solutions are proposed, perhaps indicating that the authors are none-too-sure of the solution.

The first is a preestablished harmony between first and third person identity. This is not as implausible as it sounds at first. In its charitable interpretation, it relies on generalizing from content identity (my desire, his desire) to concept identity (desire pure and simple) – an individual's desire for ice cream is a token of thetype desire for ice cream. The sense organs of different individuals are similar, though perhaps not even qualitatively identical. This apparent misfit between my sensing and desire and that of a candidate third person is, contra Locke, that I can see squareness without feeling squareness. Vision and touch are already connected in a common map of the psychophysical organism, the basis of the ancient notion of a common sensorium. Unfortunately, the skeptic will have a field day with this one, since the preestablished match between the first and third person requires the hypothesis of a common sense that is possibly non-existent and raises the bar on

---

level, and we've said nothing about algorithms or physical implementations" (Nichols and Stich 2003: 210).

A second approach dispenses with a priori agreement in advance and proposes conceptual change. The concept of desire starts out being a first person one, and then changes to accommodate the third person. The other individual has confirming and disconfirming experiences that enlarge the one's concept to include that of the other. How this conceptual change occurs without presupposing the disconfirming experiences on which it is based is left a puzzle, which the authors acknowledge: ". . . We're afraid we can't be of much help since we share the puzzlement" (Nichols and Stich 2003: 205).

The third solution is an argument from analogy:

> One obvious route to explore would be a suitably psychologized version of the venerable argument from analogy (Nichols and Stich 2003: 204)

While analogical thinking can be useful in the context of particular doubts about what someone believes, wants, or feels, it is no solution to global doubts about the availability of minds to read in the first place. Goldman tries to cover his bases again: "It must be recalled, though, the overall theory I endorse is a simulation-theory hybrid. It is no embarrassment to such a theory that some mindreading is done by theorizing" (Goldman 2006: 83).

> "It is not hard to find special first-person methods being invoked even by leading exponents of the child-scientist approach . . . . [I]nfants use imitation as a 'discovery procedure' in mentalizing. When a child observes a creature that imitates the self and hence is behaviorally 'like me,' this prompts an inference to the conclusion that the other is also mentally 'like me,' using an analogical argument from one's own mental states to those of the other" (Goldman 2006: 90)

At this point, Goldman cites several examples of analogical reasoning from Meltzhoff with which he (Goldman) finds no objection from at ST perspective, while maintaining it cannot be integrated into the child scientist viewpoint (Goldman 2006: 90).

The position of the extreme simulationist is that we are simulation machines all the way down. It should be noted that there are no extreme simulationists and even Goldman acknowledges: ". . . We advance an ST-TT [Simulation Theory – Theory Theory] hybrid, with emphasis on the simulation component" (2006: 21, 23). Even so, it is useful to stretch the approach to see where it breaks. Goldman clearly distinguishes between a computer simulation and the resemblance ("copy") approach to (replicationally) simulating processes that he pursues. When a computer is performing a mathematical calculation it is not simulating mathematics, it is doing mathematics. When the computer is performing the same calculations as part of the process of predicting the weather, then it is simulating the weather using the calculations. Goldman allows that (some) mental processes are computational processes. But adds the caveat: "All I am saying here is that being computational does not make a mental process a simulation in the sense relevant to ST [simulation theory]" (Goldman 2006: 36). This is an amazing statement considering that what Goldman calls "low level mindreading" draws extensively on the evidence of cognitive neuroscience, especially mirror neurons (Goldman 2006: 119).

At every level of mindreading – neurological, functional, informational (computational), representational, phenomenological (awareness) – simulations of the environment, including both other individuals and things – are occurring.[13]

Neurological: In about 1997, Victorio Gallese and his colleagues discovered a class of premotor neurons in the macaque monkey brain that were activated not only when the monkey executed purposeful actions such as grasping objects, but also when observing other individuals (monkeys or humans) discharging similar behaviors. Gallese called them 'mirror neurons':

> "Interestingly, on the functional level of description, they are only visually activated when another agent is observed acting in a purposeful way upon objects, for example, with his hand or mouth. In many mirror neurons we find a strict congruence between the particular action that is being observed and the executed

---

[13] These levels are mostly analytic, not compositional, and are pervasive in Metzinger (2003), though I do not use them exactly as he defined them.

action effective in driving the same motor response that has been observed. Mirror neurons, therefore, have been hypothesized to constitute the neural correlate of an observation/execution matching system… [14]

This research was extended to the human brain, obviously using less invasive methods than monitoring individual neurons, and human mirror neurons were discovered in a sector of the posterior parietal cortex connected with the area F5.[15] As indicated in the section on the philosophical significance of neurology, these results were further extended to humans using various neuroimaging methods such as magnetic transcranial stimulation (TMS), functional magnetic resonance imaging (*f*MRI), and electro encephalogram (EEG) cartography.[16] Such technology enabled many of the results to be extended to humans. We humans have mirror neurons too, which are activated both when we take basic actions as well as when we observe the same kind of action in others. For example, the individual whose brain is being monitored by the *f*MRI machine is viewing photographs of hands and feet being painfully impacted by a door or a thumb being sliced by a knife (along with the cucumber being diced). These areas correspond closely to those parts of the brain recruited and activated when individuals actually experience pain.[17]

A number of qualifications are needed to prevent jumping to conclusions about what this all might mean in terms of other minds, mind reading, and empathy. Indeed the interesting issue is precisely what does it all really *mean*. The most value neutral description that I can find is that these experiments point to a low level neurological mirroring of the two organisms where "mirroring" means "isomorphically mapping." The

---

[14] G. Rizzolatti and M. Gentilucci. (1998) "Motor and visual-motor functions of the premotor cortex" in P. Rakic and W. Singer, ed., *Neurobiology of Neocortex*. New York, Wiley cited in T. Metzinger, (2003) *Being No One: The Self Model Theory of Subjectivity*. Cambridge, MA: MIT Press, 2003: 366-7.

[15] V. Gallese, (2007) "The Shared Manifold Hypothesis: embodied simulation and its role in empathy and social cognition," in *Empathy and Mental Illness*, eds. T. Farrow and P. Woodruff, Cambridge, UK: Cambridge University Press, 2007: 452.

[16] See Chapter ___ on The Philosophical Significance of Neurology for Empathy: The Philosopher's Cerebroscope where a detailed drill down is performed.

[17] Philip L. Jackson, Andrew N. Meltzoff, and Jean Decety. (2005) "How do we perceive the pain of others? A window into the neural processes involved in empathy." *Neuroimage* 24 (2005): 771-779. See also J. Decety & P.L. Jackson (2004) "The Functional Architecture of Human Empathy" in *Behavioral and Cognitive Neuroscience Reviews*, Vol 3, No. 2, June 2004, 71-100.

discharge of neurons and activation of cerebral areas are isomorphic across two organisms, the one organism performing and the other merely observing.

Goldman, a major proponent of mindreading as simulation, is careful to distance himself from the conjecture that mirror neurons [MNs] are the foundation of a simulation heuristic that give rise to social understanding, empathy, or mind reading, none of which are to be taken as synonymous (i.e., they are all different, though perhaps overlapping in some details). Mirror neurons are not necessarily the basis of a "simulation heuristic":

> It should be emphasized that the hypothesis being advanced here is not that MNs themselves constitute a full-scale realization of the simulation heuristic. In particular, we do not make this conjecture for MNs in monkeys. Our conjecture is only that MNs represent a primitive version, or possibly a precursor, of a simulation heuristic. . . [18]

The suggestion is that what is neurological mechanism in the macaques can be further elaborated in other species to provide additional functionality. A brute force mechanism mirroring the corresponding neural discharges in the observer and actor provides as much information to the macaque monkey as it needs to accommodate itself in relating to macaque conspecifics. However, in other species, this mechanism could be recruited for other purposes in other organisms such as humans that had the capability to elaborate the information along additional trajectories in affective, volitional, and cognitive distinctions.

At least one scientist regards mirror neurons as sufficient for empathy, social cognition, and related affectivity. Both theory of mind (TOM) and simulationist approaches are refuted—or at least put in their place—by existential neuroscience according to M. Iacoboni.[19] We do not get to the other by means of an inference "Thus, it seemed that mirror neurons also provide a mechanism for understanding the intentions of others"

---

[18] V. Gallese and A. I. Goldman. (1998) "Mirror neurons and the simulation theory of mindreading trends" in *Cognitive Sciences* 2: 493-501: 498 cited in A. I. Goldman.(2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press, 2006: 136.
[19] M. Iacoboni. (2007). "Existential empathy: the intimacy of self and other" in *Empathy and Mental Illness*, eds. Tom Farrow and Peter Woodruff, Cambridge, UK: Cambridge University Press, 2007: 310-21.

(2007: 318). The gist of it is: mirror neurons do *not* support simulation – you must be in interaction in order for the individual to see the action of the other person and have his mirror neurons be activated: "The existence of mirror neurons only makes sense for agents fully interacting with other people and their environment, in which the basic dichotomies inherited by classical cognitivism melt down completely [. . . including] subject/world [and] [. . .] inner/outer (Iacoboni 2007: 317). Iacoboni regards what he calls "existential empathy" as implementing rich, human social relations as described by Heidegger and other existential thinkers directly in mirror neurons: "Thus intention understanding is implemented in human mirror neurons by activating chains of mirror neurons" (Iacoboni 2007: 319). If the intention cannot directly inhabit an individual mirror neuron – like Descartes' thinking substance inhabits the pineal gland – then perhaps a chain of them will allow enable the intention. Without making too much of a simple category mistake across physical and mental attributes that are incompatible with existential empathy, let us note the something is lost in translation between the two levels of neurological and folk (social) psychological discourse.[20]

Describing the mechanism of mirror neurons as causally effective in the communication of emotion is far from being the only possible description. The activation of mirror neurons could also be epiphenomenal—an idle wheel—while the communication of affect is effected by recognition of the emotion in the face, voice, gesture, of the other person through visual and auditory perceptions. The activation of the mirror neurons could also be described as simulation of one ore more aspects of the viewed emotional state.

At this key juncture, simulation becomes for Goldman what it is all along, an eclectic, catch all word used to stitch together disparate research experiments from an engaging perspective. Goldman deserves credit for avoiding the howler committed by Iacoboni. Yet Goldman argues against making the mirror system the unifying basis of social

---

[20] See Chapter ___ on The Philosophical Significance of Neurology for Empathy: The Philosopher's Cerebroscope.

cognition (e.g., as Iacoboni proposed above and Gallese proposes elsewhere[21]), urging a more modest and inconclusive result that the motor system provides a basis that will have to be further specified. Goldman: "Because the model posits unmediated resonance, it does not fit the traditional form of simulation in which pretend states are fed into an attributor's own cognitive equipment (e.g., a decision-making mechanism) to produce a further state. However, I do not regard the creation of pretend states, or the development of cognitive equipment to operate on such states, as essential to simulation. I associate that form of simulation only with high-level mindreading" (Goldman 2006: 131).

Informational (computational): From an informational (computational) perspective, it is provocative to redescribe the neuronal mirroring as a "common coding" mechanism whereby the organisms share a code in which the premotor neuron are activated.[22] This is also elaborated by Metzinger as the principle of substrate sharing (Metzinger 2003: 84).

Here "simulation" comes into its own as the intermediate level of analysis between the neurological (mirror neurons) and the functional (inner imitation), the latter engaged in the next section (Functional). Simulations are *off-line* when they are being computed by way of an explicit plan to travel to an unfamiliar location in town, or trying to imagine how one would feel *if* one were in his shoes, consciously cogitating as one might say. The simulations are *on-line* when spontaneously occurring due to computations of which we are aware only of the results such as day dreams, reveries, or active imagination. A background or reference model is required against which to simulate – i.e., compute - other possible worlds in planning for the future.

Here the notion of simulating an emotion and the capture of an expression of emotion in empathy by mean of an information processing (computational) model does initially get traction. Of course it should, simulation is part of a computational model. But it is not

---

[21] "What do mirror neurons mean?" Various papers: http://www.interdisciplines.org/mirror/papers (site checked on 2008/10/14)

[22] W. Prinz. (1990). "A common coding approach to perception and action" in O. Neuman and W. Prinz, eds., *Relationships between Perception and Action: Current Approaches*. Berlin: Springer-Verlag: 167-201.

quite what it seemed at first when simulation was supposed directly to encode and decode emotions by isomorphic discharge of mirror neurons.

The way in which all the developmental history of experiences with significant others, parents, teachers, peers, in all manner of contexts get associated together using contiguity of space and time, resemblance, and causal connections in the 18th century Humean sense of associationist psychology links up with the modern model of a connectionist neural network. The individual can then try to make sense out of this entangled complex of associations by applying further cognitive and verbal distinctions from a serial processor – conscious thinking and problem solving - that accesses the results of the parallel network. Whether the access to the developmental history is by means of an associationist psychology or a connectionist network model, however, the target phenomenon, the expression of emotion that is empathically received, remains a complex tangle of interrelated emotional vectors – a causal thicket if you will (Wimsatt 1994) – that is neither more nor less explanatorily clear than the target phenomenon itself.

Thus, to give the simulationist approach its due credit, computationally, empathy is a serial processor on top of a neural network mechanism. These processors work in parallel and dynamically interact. The functional microstructure of empathy is that of a parallel distributed processing network. Visual inputs such as bodily posture, facial expression, the tone accompanying the speech, perlocutionary force, content, even smell, for example, the smell of fear, are integrated in the parallel processor. These, in turn, evoke memories and feelings that will be relevant (or not) to the situation. These inputs form a pattern of informational holism. The serial processor then projects a holistic meaning onto the network. It is able to handle "a huge palette of shades of meaning" (Clark 1988: 109) and cross reference them as in a connectionist network. What gets semantically interpreted by empathy will be patterns of activation of such units as facial expression, tone of voice, statements, bodily posture, aspects of behavior as integrated in context. The emotion – e.g., anger – will emerge from a continuum of activation patterns. The nucleus of the emotion is not directly projectable onto the connectionist unit whose interrelated activation constitutes the emotion. The quality of the emotion remains

ineffable, extending as it does "down" into the complex of connectionist nodes. The activation pattern of the processing structure encodes the identify of the particular emotion (Clark 1988: 135), leading to a possibly intractably large number of qualitative shades of affectivity. In that sense, the emotion is not semantically transparent. A virtual machine – a serial processor on top of the parallel processor - is deployed by humans to deal with the meaning of such emotions. "The boss is angry again." But the virtual machine is operating serially and with conscious rules on a parallel mass of networked connections that encode a vast array of overlapping experiences – particular incidents with teachers in school – for example, who would not even be recognized as an authority figure without the rule interpreter. Full, adult mature empathy has one foot in each camp – the connectionist network and the serial processor riding atop it.

Functional: A series of functions of empathic information processing emerge as a useful decomposition of the results of empathy as the "higher" level, fundamental intention of human relatedness. "Function" is used in the ordinary information processing sense as mapping input to output, possibly including transformations in the operation. Usually designers employ functions in order to compose artifacts that implement higher level operations. For example, I am thinking of building a computer program to navigate the shortest distance between two points out of elementary, basic assembler language statements (functions) such as move, calculate, join, disjoin, if-then, load, add, change, delete. In a sense, a function has as its purpose the mapping of input to output, possibly in the process of transforming it and can be used synonymously with "purpose" or "by design." However, functional-mechanistic descriptions usually occur at a different level than intentional ones; and that is the case here. Since I cannot figure out—possibly no one can figure it out--how to reduce or eliminate the one in favor of the other, both will be used (see next subsection under "Representation") and accompanied by the necessary conditions and qualifications.[23]

---

[23] This paragraph is quoted verbatim in Chapter ___ Empathy and Intentionality – one instance should be deleted.

Functionally speaking, input perception on the part of the organism causes an imitative behavior that is visible in neonates and in certain forms of pathology (echopraxis and compulsory imitative behavior) but which is usually inhibited, resulting in nondisplayed ("inner") imitation.[24]

> Famously, Meltzoff and Moore [. . .] discovered that infants as young as 42 minutes can imitate certain facial gestures such as tongue protrusion. . . . Meltzoff proceeds to argue that the infant's ability to interpret the bodily acts of others in terms of their own acts - and similarly for experiences – provides leverage on the problem of other minds (Goldman 2006: 194).

In this task, Meltzoff was guided by the early results of Jean Piaget who carefully recorded the sounds, head movements, and hand gestures of infants at one month, two month, three months, etc. Piaget carefully distinguishes imitative reflexes that coincidently converge with the actions of others in the environment from the accommodation of the neonate's sensorimotor schema to observed actions – head nods, hand grasps, smiles – in the neonate's interest in making these actions continue by responding imitatively to them with "behavior designed to make interesting sights last." At the risk of oversimplification, such sensorimotor schemas as waving, grasping, vocalizing syllables, become the basis for holding a spoon, feeding oneself, holding a crayon, drawing, as well as uttering one's first words. These and many more like them are a step in the direction of elaborating sensorimotor schemas in the direction of conceptual intelligence and language, though the former are never completely eliminated and remain the basis for further intellectual achievements.

No one is maintaining that infants who are 42 minutes old are consciously mimicking anyone else's gestures. The neonate comes out with an imitation program that causes its organism to mimic what it perceives. Fairly early on, the infant learns to inhibit this constantly running imitating mechanism, resulting in the "inner imitation" that occasionally rises to the threshold of conscious awareness in T. Lipps (1903) and related aesthetic discussions, for example, in Suzanne K. Langer:.

---

[24] S.L. Hurley. (2002). *Consciousness in Action*, Cambridge, MA: Harvard University Press, 2002.

> The perception of emotional expression leads the observer to an unconscious, merely incipient imitation of the fleeting act, and . . . the resulting faint tensions involve an equally faint feeling by which he understands what is passing in the other person (Langer 1967: 176).[25]

Inner imitation is a candidate for the functional basis of empathy. Working from the outside inward, the subtle and subpersonal imitation of "smile lines" and "grief muscles" and related postures activated throughout the entire body at the physiological level gives rise to qualitatively similar emotions that occasioned them in the other individual. This provides the basis for emotional contagion and diverse forms of vicarious experience. It is a further question whether these actions are based in mirror neurons at the previous level, and the suggestion is that they are indeed so based, making this an example of explanatory decomposition of inner imitation into the connectionist information network into mirror neurons.

In contrast to inner imitation, an alternative operation on which functionally to base empathy is identification. This can take two forms. The first in which the individual accommodates itself to the other, which is isomorphic to imitation; the second in which the individual assimilates the other to itself, substitutes itself in the role of the other in a given context, e.g., role playing. Of course, identifications of the form "I am you" and "You are me" are logically absurd in that they physically identify two psychophysical entities that are distinct. This logical absurdity that no amount of reasoning can completely eliminate - we are distinct psychophysical organism that are always and inevitably fellow travelers in community - is perhaps the ultimate basis for the skeptic's obsessive unease that one can really get to know another.

Therefore, empathy from a functional approach starts out being built on a semantico-physical absurdity. The DNA of the two organisms is definitely distinct, though the National Geographic and IBM Geonomic Project has provided evidence that all human beings on the planet earth right now (2008) are all related, that is, share overlapping DNA

---

[25] S. K. Langer. (1967). *Mind: An Essay on Human Feeling*. Vol. 1. Baltimore: Johns Hopkins Press cited in Anna Aragno, "The language of empathy: an analysis of its constitution, development, and role in psychoanalytic listening," *Journal of the American Psychoanalytic Association*, 2008; 56: 713.

from a single (or small number of) common birth mother(s).[26] The idea of global identification does indeed occur in the reproductive process where the chromosomes from two organisms perform a dance in which the genes are mixed and matched in a process of mitosis; but it ends with the process of birth when the baby becomes a physically separate entity, having been a part of one physical system for nine months.

Therefore, this cannot be the "is" of identity, an "is" that is symmetrical (as in 7 + 5 is 12). It is a partial identification, a trial identification based on commonly shared properties, traits, and attributes, in which one individual is substituted in a network of associations to which it does not initially belong. In a compelling, concise study, Ted Cohen calls this "metaphorical identification," making the point that seeing X as Y is generally different than seeing Y as X. Seeing the Lord as my shepherd is quite different than seeing the shepherd as my Lord. Thus:

> [. . .] The creation, expression, and comprehension of metaphors must involve speaking and thinking of one thing as another [. . .] Understanding one another involves thinking of oneself as another, and thus the talent for doing this must be related to the talent for thinking of one thing as another; and it may be the same talent differently deployed (Cohen 2008: 9).[27]

In engaging with art and literature, the substitution of one for another happens spontaneously. In those instances where the reader (in this case myself) is experience-distant from the events at hand – e.g., growing up female in an Islamic family in Afghanistan – it may be useful for the reader to ask himself, "How would I feel if that happened to me?" However, even then it is the mark of an accomplished master of the art that no explicit effort on the part of the reader or audience to cause the identification is required. It is also a mark of the master that the audience seems to be both addressed individually yet is coextensive with humanity itself. Still, if one were to make explicit what is happening in such a personal identification, then it would have the form of

---

[26] National Geographic. https://www3.nationalgeographic.com/genographic/Geonomic project. It is particularly compelling to reflect on this scientific result when you meet a person, who, for whatever reason, initially rubs you the wrong way. You and this individual are relatives.
[27] Ted Cohen. (2008). *Thinking of Others: On the Talent for Metaphor*. Princeton, NJ: Princeton University Press, 2008. I am not sure Ted would agree with my use of his insight, which, in any case, contains numerous conditions and qualifications that I cannot discuss here.

putting oneself in the place of the other. The distinction between I and you is tentatively cancelled.

In extreme situations – without art or aesthetics being involved – identification with the aggressor can enhance one's ability to survive, since it enables one to figure out what not to do in order to avoid further negative consequences, punishment, though the identification can become maladaptive unless interrupted by differentiation from the aggressor. Thus, a global mirroring of the behavior and affects of two persons does occur. Other special situations occur that invite the activation of a partial identification including such ones as the preparation of a role in the theatre, undercover police or espionage work, and unconscious imitation in the mirror transference in psychoanalysis, in which repetition of roles and behavior pattern substitutes for recollection. The global identification is an intrapsychic one in which conscious pretending is put out of circuit - the actor reacts spontaneously as the target behavior is enacted based on a transient global identification with the source.

Representational: In simulation of the representational (intentional) relation the other individual is simulated as a second-person perspective on the subject-object relation with the object component being another individual (person) with intentionality of its own. Yet there are significant short comings with this definition of representation as simulation when it comes to mindreading as mindreading. As individuals and as a species, human beings seem to suffer from a hyperactive intentionality detector.[28] Individuals are quick to attribute intentions to others that may or may not be accurate. We are quick to give meaning to the statements and incipient gestures of others that may or may not be accurate. This is what individuals in fact do, whether it is the result of intelligent design or variation and natural selection that occurred among our hominid ancestors, dwelling in the valleys and on the plains of present day east Africa. It is easy to invent a just so story that failure to attribute bad (hostile, murderous) intentions to others ("conspecifics") can quickly prove fatal whereas the false attributions of good (friendly, altruistic) intentions

---

[28] This idea occurred to me many years ago; but I have since discovered it published and available in Metzinger (2003) where he attributes it in a reference to a third contributor; but I cannot find the particular page.

is relatively indifferent—in the former case my enemy will already have killed me before I realize he was the enemy—there is a perverse security in attributing base motives. This attribution of meaning and intentions can and does occur with near hallucinatory intensity and force, bordering on clinical paranoia, though no physical objects are being visualized in the "hallucination." In short, we go around giving meaning to behavior and situations; and often our meaning attributions are wildly off the mark. We ought to attribute some motive, but we are lame when it comes to doing that accurately. This results in various deformities in logic and probability that have exercised thinkers in debating how deep is the rationality of rational economic individuals. Having a small profit in hand, we are risk averse in seeking a potentially higher payoff. Having a small loss in hand, we exaggeratedly welcome risk, even at the cost of a larger loss.

It is useful, indeed highly adaptive, to be able rapidly to predict another individual organism's behavior – if its next action is going to be attack, or offering food, or an invitation to mating behavior, "you would do well to anticipate this quickly."[29] We give these actions meaning. We make sense of them in terms of the intentions of the other. We do this well enough to survive individually, for the most part, and to have survived as a species. In many ways, we have prospered as a species.

In an unrelated example of hyperactive intentionality attribution, I knew a boss who was considered by his employees to be an extremely demanding and unforgiving task master. Others found him highly critical and fault finding. Of course, he was also humorless, grouchy, and seemingly short tempered. People usually came away from an interaction thinking he was angry due to their own short comings or inadequacies. Not so. He suffered from arthritis. He kept it a secret from his coworkers because he did not want sympathy and due to privacy concerns about health. It was his health and that was his decision. He was short tempered because he was in pain. He was rather happy with the performance of the team, who he found to be hard working and responsive, albeit a tad timid and shy when they were in his presence (which he found to be annoying). In that

---

[29] S. Baron-Cohen, (1995), *Mindblindness: An Essay on Autism and Theory of Mind*, Baron-Cohen, Simon, MIT Press, 1995: 12. [Editorial: the following was inspired by the above and references it at a couple of points, but really does seem to be original (6/2/2008)]

sense, the skeptic about other minds is onto something, even though he subsequently gets carried away in hasty over generalization. We cannot tell whether the boss frown is due to annoyance at our work or due to indigestion. Sure we can ask, and successful workers do. But that takes the uncertainty up a level, for does he really want to share that information? Is he dissembling? Whose business is it anyway? We should not be too certain that we can see into and "read" the intentions of other individuals. We mostly do a very sloppy job of it, though we can improve with practice and follow up conversations.

Empathy is not the simulation of another person's intention. Empathy is the intention that another and one form a community of which the other and oneself are a part. It is precisely empathy that acts as a corrective on our overhasty attributions of specific friendly or hostile intentions and beliefs to others. Our receptivity to the affections and emotions of others is what gives us the initial clue as to the value(s) that are the goals of the intentions. The mind gets us information – we also need as assessment of that information in the form of an unmediated resonance as to how the situation feels – an elementary value judgment, if you will, whether the entire milieu is in order or something does not fit. In primate groups it may be who wins higher status (Baron-Cohen 1995: 15), thus getting better tidbits of food or the attention of a preferred mate. Yes, we are adapted to apprehend and use the joint attention mechanism and to attribute intentions to others; but like so many mechanisms the intentionality detection mechanism is a kludge.

Phenomenological: In general, I do not simulate the experience of the other person. I experience the other person. We have a conversation over a cup of coffee. We do not simulate a conversation; we talk.

Here "phenomenological" refers to the experience of a given individual in their conscious (i.e., non-dreaming, waking) awareness. The way in which simulating others occur in everyday awareness is by "putting oneself in the other person's shoes." When those arguing in favor of the skeptical position about other minds are treated unfairly and

wronged, then they suddenly find a powerful argument for otherness in righteous indignation as they ask: "How would you feel if this happened to you?" [30]

When one individual cannot get a feel for the other person, whether due to cultural or profound personal differences, then the one individual tries putting her- or himself in the other's shoes. If that does not immediately succeed in bringing experience-near what was remote, the individual uses additional tools – similar to the "as if" - such as building an analogical comparison.

The "as if" processing invites an iterative approach. One individual uses the "as if" operation if he does not know the other very well or is having trouble appreciating the other's behavior or statements. But if I immediately attribute my own character traits and dispositions and history to the other, then I am likely to miss the mark. If the other were like me in these ways, then we would probably already understand one another. If the other were like me in thee ways (or vice versa), then we would not be disagreeing and, positively expressed, would be in interrelational harmony with one another to start with. And even if we did not understand one another, we would be able to agree on our disagreements, reach an accommodation about our differences, and so reach an understanding. But, by hypothesis, in this example, that is not working. We are at loggerheads. We just do not get one another's position, preference, or point of view. So egocentrically attributing my position onto his as the default procedure is not going to work. Indeed it will make things worse, deepening the misunderstanding.

So, in this example, my thinking of the other has got to be informed by something more than egocentric attribution. I have got to have a clue, a single. Sometimes a person can get started with a lucky guess. Using intuition, one can find a difference or salient convergence amidst divergence. Thus, I realized that the new boss really found gambling in LasVegas to be a vacation (whereas I had a negative reaction to the glitz and perceived artificiality). He was not at all worried about the uneven odds. He was thrilled by the

---

[30] See Chapter ___ Empathy and Altruism: From Possibility to Implementation, including a discussion of this question which is a key lynch pin of the argument in T. Nagel's *The Possibility of Altruism* (1970).

uncertainty, like others enjoy a roller coaster ride, and found the entertainment value a fair compensation for his (modest) losses. I knew all this like a fact reported in a newsletter; and it did not really hit home until I experienced firsthand how relaxed and refreshed he was by the entire experience. I got a sense of refreshment that I would take away from a quite different milieu such as a thrilling piece of theatre; but this was a "bottom up" experience that was affectively disclosed listening to him recount what fun he had, and by quarantining my own (negative) reaction to gaming and focusing on his expression of enthusiasm and delight. Only then, when I was able to use this trace affect as input to my empathic understanding ("as if"), was I able to identify – whether metaphorically or empathically is a matter of terminology – with him in his experience of his vacation. This identification in turn generated additional empathic results, including some experience where we actually shared values around fine wines and cooking. (He was not impressed by the cooking in Las Vegas.)

The Example of the Life of Others

Let's consider another example where the first impression is off putting and challenges our empathy in a specific way. The intense and emotionally empty stare the communist party apparantik, HMW, in the film the *Life of Others* (*Das Leben des Anderen*) is affectively alienating. This suggests an exercise and a comparison based on what would have had to happen to me in order to stare vacantly into space without a trace of affectivity or emotional warmth. This individual (HMW) is a true believer in the validity of the East German model of communism, and he is good at what he does. That is, he is a master at breaking down the defenses of so-called counter-revolutionaries – people with normal middle class values and a desire to express their own non-conforming opinions. HMW is definitely *not* a likeable or sympathetic person (in the ordinary sense of the word "sympathy").

Yet as HMW spends hundreds of hours eaves dropping on the lives of the well-regarded party playwright (the other protagonist) in order to discover compromising information about him, so that his boss can seduce the playwright's girlfriend, HMW gets access to

his own humanness through the emotional turmoil of the relationship between the playwright and his girl as well as the unknown troubles of being the target of a politically motivated "witch hunt" (of which at least the playwright is unaware).

Without excusing the behavior of HMW, the audience starts to experience hope and fear for HMW's fate amidst the complexity of the situation. HMW's own life is empty, his own sex life portrayed pathetically, stereotypically, and exaggeratedly, to make the point, with a Walküre-like East German Brunhilde. He starts liking vicariously through the people he is wiretapping. Without intending to funny, HMW spins out a ludicrous and elaborate lie in order to protect the playwright and his girl. In order to delay and derail the planned depredations of HMW's boss, HMW pretends that he (the playwright) is a totally loyal party member writing a play about Lenin's heroics. In order to pretend that the play is being written by the playwright, HMW must actually write large parts of the play itself (since he has to include them in the transcript). The best pretence is to create the target of the pretnece. As regards, HMW's emerging emotional life, the film makes the viewer think, "If I were such a person, then I must have been brought up in an emotionally empty environment and then further buried what was lost, creating even more of a second order forgetting." What happens when experience directly reminds me of the emptiness of the value system I was brought up to cherish and defend? Something like an identity crisis or redefinition of the self looms large.[31] When my immediate experience is inadequate for getting a feel for what is going on with the other person, then I try building a model, deploying intellectual and conceptual tools as opposed to immediate affective openness. Based on the model, I run a simulation; but the simulation optimally gets going when informed from the bottom up by a sensory or affective input, in this case, by a sense of emotional deadness, that reveals the world of the other in its immediate and qualitative impact.

---

[31] Of course, as a loyal East German, who actually believes in the system, HMW is heading for "world collapse" as the Berlin Wall comes down in 1989 and a third of the population was shown to be working for the Stasi (State Police) in a rotten and corrupt system. On "world collapse" see Jonathan Lear. (2008). *Radical Hope*. Cambrdige, MA: Harvard Unvieristy Press.

When my empathy breaks down, then I deploy an argument by analogy. Then I try simulating the experiences of the other using imaginative variations, wondering what I would feel in such a situation. What would I feel knowing what I know about the other person, that he is different from me in these particulars, that he feels this way in this circumstance. This is likely to be disconcerting to the proponents of the simulationist approach to understanding others, since it contrasts empathy with simulation whereas they want to assimilate the two. Simulation is a useful method along with analogy, inference, recentering, imagining, direct receptivity, experiences recollected in tranquility, judgment, transforming, listening, understanding language when one's empathy breakdown or is inaccessible for any arbitrary reason.

As an ontological or even epistemological foundation, simulation comes too late. That does not mean it is inaccurate or wrong. It means it is tactical, not strategic. Phenomenally, simulation theory considers how a person comes to know what another person believes (e.g., Goldman 2006: 181). If I know what John believes, how would I predict further conclusions from what he believes? Goldman calls this "inference prediction" and points out that John's beliefs logically entail other beliefs. Here "simulation" is functionally equivalent to "inference." I can deploy the usual mechanisms of logic to make immediate and syllogistic inferences. I can make probabilistic and inductive inferences. As soon as I succeed in recognizing the first belief to the other person, I can make many related inferences. John believes that the room around him is a space in three dimensions. This implies belief in the truths of three dimension geometry. The challenge for simulation theory is to get to the first belief. Presumably logic will be useful at getting to that first belief, but are perilously close to the slope that will send us in the direction of an analogical argument from my experience to that of the other in order to establish the latter's accessibility and existence in the first place.

Once we get to the discussion of simulating propositional attitudes, then the entire structure and mechanism of language is at our disposal, then simulation is functionally equivalent to communicating with the other person and can be replaced by it while preserving the truth. Goldman suggests that our "default procedure" (2006: 176) is to

simulate, but this seems misleading and incomplete. Our default practice is to relate to friends, colleagues, family, co workers, not to mindread targets. We relate; we don't read minds. We ask "What do you think?" - we don't mentalize. This may seem to be hair splitting, yet, the choice of the approach and its defining description of "mind reading" has sent us down a particular path that over-intellectualizes others as minds they are so much more (and sometimes less). When our default procedure breaks down and we are unable to related directly, then we deploy a simulation, then we look for an argument form analogy in the local context of our relationship, not globally for all other minds at large.

Bringing Experience-Distant Phenomena Nearer via Model Building

At this point it is nearly to revisit what it's like to be a bat—the famous article by Thomas Nagel, that is, as well as the "battiness."[32] Have we learned anything that will make a difference in addressing the puzzle provided by this paradigm case? One of the many things that make this example so maddening and insight generating is that it picks an example that is experience-distant yet not totally inaccessible. I can make progress with the matter by building a model. I can make progress by imagining what it is like stepwise for me to be a bat by morphing myself into one. Yet I come up short when I work from the outside in—imagining myself able to fly, hang upside down, and use echolocation. That is still me doing these things. I am a comic book character—Batman. ON the other hand, I also come up short when I work from the inside out—if I had the neurophysiologic constitution of a bat, the awareness of battiness from the inside of the bat's perspective, then I lose the ability to articulate the immediate quality of the bat's experience, and the bat's experience escapes again.

Recall that Nagel calls for "an objective phenomenology not dependent on empathy or the imagination" (Nagel 1974: 402). So the example is relevant to a discussion of empathy. Yet the demand itself on the part of Nagel is problematic and worthy of comment. The requirement to operate independently of empathy or the imagination may

---

[32] T. Nagel. (1974). "On what it's like to be a bat" in *The Mind's I*, eds. D. R. Hofstadter & D. C. Dennett, New York: Bantam Books, 1981: 391f.

turn out to be inconsistent with Nagel's commitment to finite human understanding. Even if one wishes to dispense with empathy, imagination is one of the fundamental capabilities of the human intellect.

The risk of dispensing with either or both of these – empathy or imagination - is that the bat's experience becomes a Kantian *ding an sich*, inaccessible as a matter of the way we have posed the question—"humanly inaccessible facts" (1974 Nagel: 396) – unless we possess and access a god-like, divine omniscience. We do not – especially after Kant – and still we would still like to answer the question.

As an argument that human knowledge is finite, subject to revision, that there are things we humans will never know, it is a brilliant tour de force. The value lies in asking the question at all and posing various replies, each of which shows another aspect of our human limitations and contingencies. We have the kind of body that we have; and inhabit it the way we do. Based on what we humans have, we would still like to answer the question.

The question – the game - goes to a new level when Nagel says that the question is not what it would be like for me to behave as a bat behaves rather: "I want to know what it is like for a *bat* to be a bat" (Nagel 1974: 394). Based on this work on empathy, I propose three answers.

First, the bat does not know what it is like to be a bat. The bat is aware in the immediacy of its experience what it is like to be a bat. The bat is present in the immediately of its experience what it is like to be a bat. However, the bat does not *know* what it is like to be a bat. So why should I (or any human) know? And if I do know, then I still know more than the bat will ever know. In a footnote, Nagel emphasizes that his point is about the limitation of our imagination, not knowledge. In order to form a conception of what it's like to be a bat one must take up the bat's point of view, and one's rough conception is too rough to count as anything worth having (Nagel 1974: 397ftnt.). The qualification and footnote come too late. The objection is still the same—the bat has no conception, or

at most a very degenerate, limited one, of what it's like to be a bat. The question is rejected as unfair. The bat does not know what it's like to be a bat so how can anyone else?

Still, it is a very powerful question. It deserves another pass, perhaps a paraphrase. The point is not to have a debate about whether a bat has any concept of experience so that it can have knowledge. The issue will be reproduced even if we paraphrase "know" out of the equation by a proper phenomenological expression such as "be aware." Nagel (and I) want to be aware of what it is like for a bat to be a bat in its own awareness." That is actually a question with which we can get traction, though it will not result in knowledge in the formal sense.

Second, the bat's protoconception of what it's like to be a bat is readily accessible to us human by tuning down – dumbing down if you will - our higher cognitive facilities and imaginatively enabling our non-existent capabilities for echo-location and mouse-like instincts. With apologies to the bat, our human knowledge of its battiness includes the insight that the bat does not know what it's like to be a bat as well as we humans do. Friedrich Schliermacher would be delighted, at least initially, since we understand the bat better than it understands itself. However, understanding the bat better than it understands itself is still not understanding what it is like to be the bat. Unfortunately, the result is that we fail to understand—to have a concept of--what it's like to be a bat.

The third answer is that Nagel under-estimates human resources when he says: "Yet if I try to imagine this [what it's like to be a bat], I am restricted to the resources of my own mind, and those resources are inadequate to the task" (Nagel 1974: 394). Not so. Nagel under-estimates our uses of model building to bring experience-distant worlds, forms of life, near to our own experience without, however, disregarding the distinctions between our own experience and that being modeled. The model would look something like this. Begin by turning out the lights. Enable echolocation so that objects in the environment "appear" as a field of contrasting sounds, some of which bounce back establishing a form of permanence, and others which simply vanish into space. In addition, enable perception

of incoming chirps or tones that are audible to the subject and that signal the presence of other sources of spontaneous echoing—other bats. Presumably when the other bat got nearer, then additional sensory factors would be enabled such as smell, enabling male/female, friend/foe distinctions to be made. The end result would start to look like the audible, sound only world that Strawson starts to construct in *Individuals*, albeit with difference in detail. How would we know if the model approximated the bat's experience? In addition to being able to predict the bat's behavior, we would ask the bat. That is, we would perform scientific experiments, consistent with the ethical treatment of laboratory subjects, that would use different sound absorbing surfaces to explore the bat's choices of food, mates, escape routes. The point is that we would get to know what it's like to be a bat much more intimately and at a level of awareness than we (or Nagel) ever thought possible. Would our knowledge be absolute? Of course not. Would it be subject to revision and improvement? Naturally. The philosophical point is that for creatures such as ourselves who initially imagine that we cannot access the battiness of the bat, we can nevertheless bring experience-distant closer through model building.

A further nice point about bringing experience-distant things experience-near is raised by the example of Michael May, who had his sight restored after basically being blind from birth. [33] Recall that Mike lost his vision at an early age and, after it was restored, it turned out that he was unable to regain his depth perception. His vision was restored, but even with extensive training, he was unable to dispense with the blind person's white cane. His depth vision never returned. Curbs on sidewalks appeared as two different shades of shadow on a two dimension surface. He still needed to use the tapping, tapping, tapping – a kind of echolocation – to navigate curbs. Do not under-estimate the resourcefulness of the human being, individually or collectively. The point is that something that was experience distant is brought closer to human experience. Of course, we do not have sonar; but Mike was simulating echo location with the tapping of his cane.

---

[33] This example is considered from another perspective than that of simulating the experience of the bat in Chapter___ on Empathy and Intentionality. Robert Kurson. (2007). *Crashing Through: A True Story of Risk, Adventure and the Man Who Dared to See.* New York: Random House, 2007.

This is a related approach to that taken by Akins (1993: 264) where a full rich experience of battiness is required to be deployed by means of a network of interrelated qualities ("properties") that capture the bat's entire environment.[34] Akins tries to find what is sensible about Nagel's analogy, focusing on the experience-distant rather than the creation of an experience-near model. Akins describes a visual translation of the bat's sonar perceptions similar to the representation of a film of a sonar representation of the bat's sonar. What is missing, according to Akins, are the "sympathetic sensations" that go with swooping around a cave. Once again, do not underestimate the resources of the human being. To capture those sympathetic sensations of swooping and flying Akins would have to upgrade from film to a full immersion in a virtual reality milieu such as a flight simulator.

But, one may object, a bat is at least a warm-blooded mammal. Perhaps a bit creepy in close up photos in *National Geographic* and due to the aura of horror movies. But still creatures demonstrably like ourselves. What about really experience-distant "animals" such as insects? I'm glad you asked.

In some cases, analogical thinking must initiate the opening of empathic receptivity and guide it from start to finish. We encountered an obvious candidate for this angle of intersecting between empathy and analogical thinking in reconstructing the visual world of the red ant of the north Sahara. It is questionable whether the ant's sensitivity to polarized ultraviolet light at wavelengths beyond the bounds of visible light can ever be said to be experience-near to man. Nevertheless, the scientist is able to make sense out of the ant's experience by following question, "How does the ant find its way back to home base?" The task is not so much to see the world visually as the ant does, but to reconstruct an account of how the ant navigates. The model that the scientist constructs of the analysis of data that occurs in the ant's retina and related synapses has little, if anything, to do with how the ant experiences the world. The application of this model is a reconstructive interpretation that permits the scientist to make sense out of how the ant

---

[34] K. Akins, (1993), "A bat without qualities." In M. Davies and G. W. Humphreys, eds., (1993). *Consciousness*. Oxford: Blackwell, 1993.

can find its way back home. So in a way, the scientist has a better understanding of how the ant gets home than it does. This extra understanding is just a sign of how far we still are from being able to have an experience that is constitutive of the ant's world *as* the ant "understands" it. The initial and final states of the scientist's regulative model and the initial and final states of the ant's position may be exactly alike, but the functions by which they are reached are different, as for example in the computer simulation of arithmetical calculation compared with the same calculation as arrived at by a man himself.

In the case of opening up the world of the ant to empathic receptivity it is not accurate to speak of a breakdown of empathy, for empathy does not even get started until an elaborate model has been constructed. However, it does make sense to say that empathy which has been blocked by the radical difference in experience is effectively released by the proper analogy. A polarizing filter on a camera lens can give the investigator a good sense of how light can be decomposed into bright and dark bands. Granted that this analogy is only approximate; for the ant literally sees wavelengths to which men are not receptive. The introduction of this specific picture analogue is still the best available for finding a sensible manifold to which to apply our receptivity. So, analogical thinking entailing both the (re)construction of a model for interpreting as well as a more literal picture guides the empathic receptivity of the subject into the world of the ant.

I think it is worth noting how well this use of analogy to guide empathic receptivity back to experience that would otherwise be lost corresponds to Kant's. Recall how he says that in philosophy analogies signify something very different from what they represent in mathematics.[35] In the latter, equality between two quantitative relations is constitutive of experience so that if three members of the equation are given, the fourth is likewise given. On the other hand, in the case of philosophical analogy, it is not constitutive but only regulative of experience. We obtain understanding of a relation to an aspect of experience, not necessarily the experience. We have "a rule for seeking the fourth

---

[35] I. Kant. (1781/1787). *The Critique of Pure Reason*, tr. N.K. Smith. New York: Macmillan, 1928: 211: A179/B222.

Lou Agosta, Ph.D. (LAgosta@UChicago.edu)

member in experience, and a mark whereby it can be detected." In terms of our above example, the picture analogy provided by the polarizing filter on a camera gives us a way of detecting the form of the ant's experience, not the experience as the ant experiences it. That is to say, the scientist's model of data processing in the ant is not constitutive of the ant's experience. It does not completely capture how the ant experiences the world. Rather we are given a set of rules which tell the scientist how a coherent and unified whole can be made from the ant's perceptions. But the way in which our model makes a unified whole of these data is different from how the ant makes a unified whole of them. It is the latter of which we remain ignorant. A limit is set to our receptivity in that our model is not constitutive of the ant's experience, but only regulative of it.

Naturally, there is no question of communication between species that are so radically different as man and ant. So the situation, strictly speaking, is not interhuman except is no far as science is a collective enterprise of the human species. What is significant is how the radically different experience of the world can be opened up through the use of an analogue model and picture, which guide empathy to understanding where there would otherwise not be any.