

Neurological Reductionist Theory of Self

Atreyee Mukherjee
Assistant Professor in Philosophy
West Bengal State University

Abstract

My primary concern in this paper will be regarding the nature of the entities we call *selves* or *persons*. This is a metaphysical or ontological question. However, this metaphysical question is intimately connected with the phenomenological question. In the first section of this paper, I would like to explore the different ontological stances one may take regarding self. The second section includes exploring the different grounds for retaining the notion of self in bigger metaphysical picture. In the concluding section of my paper, I would like to collate all the fragments of arguments to get a total picture of the main structure of the argument. I would also like to mention at the end the ontic status, I would like to confer to the self and also the theory of self that would stem on being conferred upon such an ontic status.

While arguing for such a theory I would include on the one hand the positive arguments in favour of the thesis I would like to uphold and on the other hand show how the proponents of contemporary theories of self could explain the phenomenon which could not otherwise be explained by the traditional theories of self. Not only evidences from current research disprove some of the traditional functions and attributes assigned to self like epistemic transparency, etc. but also these (traditional) self-theories cannot accommodate many of the functions which currently neurosciences ascribe to the self.

Section 1

Varieties of Reductionism

Several attempts have been made over the years to answer the metaphysical question relating to self. However, Galen Strawson recommends that the best take off point on any discussions regarding this issue is by considering the phenomenological question or what many have called a naïve view (simple or intuitive view) on self. So let us rush through what Galen Strawson has to say about such a naïve view.

“I propose that it is (at least) the sense that people have of themselves as being, specifically, a mental presence, a mental someone, a single mental thing that is a conscious subject of experience, that has a certain character or personality, and that is in some sense distinct from all its particular experiences, thoughts, and so on, and indeed from all other things.”ⁱ

To begin with it is a fairly handy and accurate description of what mostly people think of when they reflect on the concept of self. However, Strawson himself by the end of his phenomenological investigation reaches the conclusion that several elements of this definition like diachronic identity, agency and personality can be eliminated. That is to say, a deeper analysis will reflect that these attributes really do not form the core sense of our self. However, there are many philosophers who would really not agree with him claiming that a person “is same thinking thing, in different places in different times.”ⁱⁱ

To understand the matter clearer, let me begin with the traditional debate amongst theorists of self, where one group sides with the theory that metaphysically a real *given* self exists over and above the body, while the other group claims that there is no real ontological self apart from the brain-body complex, the notion of self is merely a *construction*. In fact, almost all versions of Given-self theories³ consider diachronic identity as a prime attribute of the self. In fact, their whole theory of personal identity is construed in these terms. But again there are these Constructed-self theorists⁴ who are in agreement with Strawson on this issue. They go to the extent of saying that philosophers who conceive the problem of personal identity in these terms already start with the presupposition that there are persons and they do have an identity. So fundamental questions about whether at all the issue of personal identity over considerable period of time exists, is systematically ruled out. Consequently, a theory that espouses the existence of a self but not of personal identity over a large temporal frame of time is marginalized or rarely considered. This also results due to the failure to understand the difference between eliminativism and reductionism. Traditional philosophy is always inclined towards reaching to absolute answers to questions regarding the existence of the self. By absolute answers, I mean all-or-nothing type replies, that is to say, either nod in affirmation or negation to the reply to the question ‘do selves exist?’. So, inviting in indeterminate replies or perspectival approaches to such ultimate questions is rarely entertained. However, the vast array of puzzles that the thought-experiments on personal identity pose clearly indicates that we cannot have such all-or-nothing type replies to issues on personal identity. A deeper analysis of the several thought experiments, relevant and pertinent to this issue discussed by Parfit, Dennett, Buddhists and the likes, clearly shows how the problem of personal identity cannot be tackled by a simple ‘yes’ or ‘no’ type answerⁱⁱⁱ. So before any further diversion let me go straight into the issue on the ontological status of selves.

Initially philosophers thought that reply to the question ‘whether selves can be included in the ultimate ontology of our discourse?’ is to be attempted in either of the ways – Non-reductionism or Reductionism (here reductionism is used in the sense of eliminativism).

The simplest formulation of ontological non-reductionism claimed that self is an independently or separately existing entity, which is distinct from the person’s brain and the body on the one hand, and his experiences on the other hand. This self is substantive in nature and it forms the core or essence of the person^{iv}. While ontological reductionism was thought to be a straightaway denial of such a substantive core self, they subscribed to an anti-realistic attitude towards the self. Though Parfit too talked about a dichotomous taxonomy with respect to self, yet his form of reductionism can be interpreted in a somewhat different manner.

However Hume^v, Siderits^{vi}, James Giles^{vii} and many others have insisted upon a trichotomous taxonomy while dealing with the ontological status of any entity. According to them, one can be a non-reductionist about self (or for any sort of entity say K) claiming that the self is indeed an extra add-on and should be distinguished from the psychophysical complex that constitutes the human individual and the existence of the human individual consists in the continued existence of a self. An eliminativist about the self would recommend the self to be simply excluded from the ultimate ontology of our discourse, since, according to them, ‘such entities are utter fabrications’. A reductionist about self would, according to these theorists, hold that selves may be in some sense said to exist, but their existence just consists in the existence of things which are of a more basic sort (the psycho-physical complex), things of which selves are composed such that selves cannot belong to our final ontology. Now this third alternative may not apparently seem to be a genuinely different theory, however a little scrutiny may reveal its ingenuity as a third possibility (Parfit’s Reductionism).

Let us first examine how ontological reductionism is different from ontological non-reductionism.

The Non-Reductionist’s case: A non-reductionist about the selves would claim that the selves are neither mere fictions (Dennett *et al*) nor mental constructions but are ultimately real. That is to say, while charting out the ultimate ontology if we exclude selves, then our ontology will be incomplete. Again non-reductionists by countering that the selves are not mental construction seek to claim that the selves are mind-independent reality. Siderits cites an example of a corporation to explain this.

A non-reductionist would argue that corporations be allowed to be included in the ‘ontological inner circle’ or ultimate ontology ‘since it has genuinely autonomous causal and explanatory powers that cannot be reduced to the causal and

explanatory power of its members' and moreover a corporation is not identical to its members or reduced to it without the members and the furniture, etc. (though, there is no existence of corporation) because 'a corporation can continue to exist through the replacement of all its original members and property by new members and property (provided the replacement process is gradual enough'.^{viii} A reductionist however would challenge the non-reductionists on the issue of lack of evidence for their ontological claim of a corporation since no one observes anything other than its members and the properties when one perceives a corporation. Facing this charge, the non-reductionists are left with no grounds to accept the ontology of corporation other than the claim that corporations have genuinely autonomous, causal and explanatory powers.

The Eliminativist's Case: Siderits again cites the example of an inter-theoretic reduction in science and now shows how reductionism stands off better than cut throat eliminativism. Let us consider the theory of covalent bonds in organic chemistry and bridge laws of quantum mechanics. 'It has been seen that talk of covalent bonds can be symmetrically replaced (through employment of bridge-laws) by talk of certain quantum mechanical states... of the formers. The straightforward identification of the covalent bond with quantum phenomena does not seem to be an option. On the other hand, neither does outright elimination of one theory or the other seem to be in the cards. Since, however, quantum mechanics is a well confirmed theory that has a wide variety of applications far beyond the case of the covalent bonds, it appears to have the upper hand here. Given this, two questions arise: (a) Why not simply declare organic chemistry to be superfluous and eliminate it in favour of quantum physics? and (b) Why not "lighten" our ontology? The successful reduction of organic chemistry to quantum physics shows the former to be in principle dispensable, but showing this is not the same thing as giving a positive reason for eliminating the first theory. For one thing, there is not the marked incompatibility between statements of the two theories that we saw in the case of the demonic possession and microbial-infection theories of disease. While it seems to us incomprehensible how a multitude of germs could have malicious intent, we think we can see how the quantum shift could underlie some of the properties we attribute to the covalent bond. Given that the relevant bridge laws render the truths of organic chemistry deducible from those of quantum mechanics, the entities and states referred to by organic chemistry turn out to lack autonomous explanatory and causal power. Since we can explain the facts of this entire domain and more with just the entities and states referred to by quantum physics, it would violate the principle of lightness to include the covalent bond in our ultimate ontology. Here the nature of our reasons for retaining and not eliminating the theory of organic chemistry gives us some guidance. These reasons were thoroughly pragmatic. The covalent bond is not, then, ultimately real. Are we to be eliminativist about the covalent bond after all? Not precisely,

says the reductionist. It is not, they say, wholly false to claim that covalent bonds exist although, they hasten to add, this is just because of the way that we talk and think. The covalent bond is not in our final ontology, but it is a posit of a theory that, while in principle dispensable, is in practice indispensable (for creatures like us). This is what reductionism about the covalent bond comes to.’^{ix}

The distinction between reductionism and eliminativism has a semantic dimension. Semantically speaking non-reductionists about selves will claim that the selves be mentioned in the ultimate ontology or ‘our final theory about the ultimate reality’^x. Both reductionists and eliminativists will disagree with this since none of them will consider the self as an extra entity to be added in our final ontology. While the eliminativists will straight away eliminate all such talks about selves, be it our final ontology or our conventional way of speaking, the reductionists about self will declare that from the pragmatic point of view we need to carry on with our self-talk. However, in principle, we can eliminate such self-talk from our language.

It is interesting to note that the Buddhist technique of level distinction (conventional and ultimate) sprouts out because they embrace a similar form of reductionism. The difference between the Parfitian type and the Buddhist type reductionism being while the Buddhists claim that wholes are unreal (since they are conceptual constructions), Parfit does not opt for any explicit commitment on the same. While Parfit employed the sortie argument to rule out that selves are wholes (having independent ontological status), the Buddhist opted for a rigorous mereological reductionism. The ontological stance Dennett would like to adopt regarding the issue of self comes much closer to the Buddhists who deny the ontological existence of any entity called the self at least in the ultimate level. Like the Buddhists he would retain self-talk at the conventional level since self-talk is instrumental in making our functioning of day-to-day chores better.

However, a fourth view can be still talked about that the selves non-reductively supervene on the more basic or particular entities, mainly the psychophysical elements. Now what is exactly meant by non-reductive supervenience? According to the definition of supervenience, for every difference in the psychological sphere there is a corresponding difference in the physical sphere. That is to say, there cannot be any difference between two mental states, unless there is a difference in the physical states corresponding to those mental states. Non-reductive supervenience engages in a type of determination without any explanation. All the facts about the supervening level are determined by the facts about the base level, yet there is no room for explanation of the link between the base level and the supervening level. Consequently, the entities that non-reductively supervene, in these case selves, have to be bestowed with autonomous explanatory power. This view upholds that when psychophysical elements interact in a sufficiently complex way, they give rise to supervening

properties (person-properties), which could not have been predicted, from the knowledge of the constituent psychophysical elements alone. This view is closer to non-reductionism in the sense that it still believes selves to be really existent. It also respects our commonsense since, according, to this theory, the selves are bestowed with explanatory power, and thus it raises the status of selves from being mere instrument or pragmatic tool.

Several charges have been leveled against the Buddhist theory of mereological reductionism. This gave rise to yet another possible view – non-reductive mereological supervenience. These theorists claim that if it is accepted that not all wholes are non-identical with their constituent parts or have an independent or rather distinct ontological status apart from their constituent parts, a fifth possibility can be considered – at least some wholes are neither identical with their constituent parts nor are they ultimately distinct from them. That is to say, these wholes can be supposed to possess really novel properties, which cannot be reduced to the properties of its parts. Again, the wholes cannot be said to be distinct from the parts since their relation is that of asymmetrical dependence. In fact, if we call them identical then the wholes and the parts have to be placed on equal ontological footing. So wholes are said to non-reductively and mereologically supervene on its parts. This completes almost a comprehensive account on the various shades of reductionist theses on the self.

Section 2

Exploring the Grounds for Accepting ‘Self’

The notion of self (especially its ontic status) has undergone a renaissance with the advent of contemporary cognitive science. Thus Damasio remarked ‘of late, the concept of mind [or self] has moved from the ethereal nowhere place it occupied in the seventeenth century to its current residence in or around the brain’^{xi}. The current trend in both philosophy and cognitive science is to discard any ontic immaterial entity referred to as self. Even those theorists who like to retain the ontic self, consider it nothing over and above the body or brain (which itself is like any other part of the body).

Though theorists across the board today concede that activities in the brain are responsible for giving rise to our sense or feeling of self-hood (that is to say, the feeling that we are selves or we possess selves), yet each theorist has his own different story to tell about the self. While a group of cognitive scientists like Gazzaniga^{xii} have attempted to pin down an ontic self much like what Descartes did in a specific part of the brain and assign to it the traditional functions of self, others like Joseph Le Doux^{xiii}, Patricia Churchland^{xiv}, Damasio have rephrased the self-talk in terms of self-representational capacities. Still there are other cognitive scientists like Dennett who would support in no form the project of

localizing the ontic self. He would prefer to call the self a Centre of Narrative Gravity^{xv}, that is, the central character or protagonist of the narrative we generate (autobiography). This protagonist gives arguments to prove the existence of a coordinator from the existence of coordination or unity in our thought and expression. This line of argument may be regarded as the ‘coordinator’ argument. The ‘coordinator’ argument may appear to be similar to the design argument for the existence of God. In the latter argument one posits the existence of a designer from the existence of a design in the world. Dennett contends that the same attitude prevails while we posit the existence of a coordinator from the fact that there is coordination in our thought and expression thereof through language. So according to Dennett, the self is nothing but a theoretical posit underlying our linguistic convention.

As the story of self-changes being retold by different cognitive scientists, so also the metaphors used to express the notion of self-changes. Thus we have a catalogue of strikingly diverse and contradictory metaphors related with self-talk. Interestingly, the diverse metaphors suggested by these cognitive-scientists actually represent the multifarious functional capacities they thought central to the notion of self and hence assigned to it. However, these metaphors do not give us an exhaustive list of functions that the self purportedly performs. It leaves out especially some functions that contemporary neurobiological accounts assign to self.

There can be theories, which accept the view that the self-reveals itself without mediation of anything else^{xvi}. On these views we need no further proofs to establish the existence of a self. Apart from these theories, all other theories require proof to establish the existence of the self.

Now what kind of things may count as a proof for the existence of the self? The most suitable candidates that come to our mind as grounds for accepting the self are presumably the functions and the properties of the self.

Irrespective of whether it is SAG theory or a SAC theory, it appeals to be incumbent upon one to articulate the functions of the self. The self-talk as such hinges, to a great extent, if not solely, on the existence and articulation of the properties and the functions that we tend to ascribe to the purported self. So let us first briefly recapitulate the different grounds traditionally considered for admitting self and then simultaneously present their counterpart in the current debate.

The grounds are as follows:

Biological Ground

Inextricably linked with the notion of 'self' is the notion of 'other'. The very thought of 'self' is conceived as something which is distinct from the 'other'. This distinction has to be grasped by all life forms since it is the key to self-preservation. Fundamental to self-preservation is demarcating or defining identity of the respective organism so that it can distinguish itself from all the 'others' and then devise mechanism to defend itself from its predators. So *self-recognition* and *resisting self-destruction* constitute the basic biological principle of self-preservation.

Contemporary Scenario

This brings in the notion of boundary, which is the primary strategy of demarcating 'self' from 'other', 'me' from 'rest of the world'. In human context much work has been done by Paul Rozin while Dawkins^{xvii} did a lot of work with the lowest form of animal life in his famous book *The Extended Phenotypes*^{xviii}.

Now this biological perspective ascribes self not only to the lowest forms of life (amoeba, hermit crab, etc.) but also to cells and tissues. Several other papers show how the science of immunology proposed the 'self-other' dynamics as the key working strategy of the host cells and the foreign cells.

Interestingly though this view gains support from secular human thought^{xix}, it is incompatible with Christian theology. Not only Christian theology, confining the presence of self just to the higher animals like human beings was the slogan of the Cartesian doctrine too. Descartes in fact ascribed the status of mere automata to all animals excepting human beings. Confining selfhood only to human beings was conceived in the modern age perhaps because the philosophers of those days thought of self as the human counterpart of the divine notion of soul. However, progress in cognitive ethology today disproves such anthropocentric views. Cognitive ethologists have shown from a number of case studies that animals are not bereft of emotion and hence exhibit mental properties in the sense human beings do. Evolutionary biology also stands in support of the view that a primordial notion of self is experienced by all life forms (from the lowest one to the highest animals).

Semantic Ground

The self is unanimously regarded as the referent of the personal pronoun 'I'. Whether one is a believer or non-believer in the existence of a self apart from the body, each of us uses this word while referring to herself or himself. Those who identify self with the body find it difficult to explain expressions like 'my body' again those who accept self apart from the body cannot explain expressions like 'I am fat'.

Contemporary scenario

Current cognitive scientists disavow the claim that self is the referent of the personal pronoun 'I'. They say that it is just a linguistic convention that we use 'I' so as to make all our feelings, desires, beliefs and thoughts intelligible and comprehensible (communicable) to others. If we briefly recapitulate Damasio's view on neural self, we will see how the self emerges or works without any interference of the language. According to Damasio, it is the continuous reactivation of the two sets of representations that is responsible for our sense of self-hood. The first set of representations is about the key events of an individual's autobiography (what we do, whom we like, our recent activities, etc.), while the second set of representations constitutes the primordial representations of an individual's body (what our body is like in general, what the body has been like just before and after the processing of perception of any particular object, the background bodily states and emotional states). Now let us imagine a third set of representations, which is different from both the first and the second set yet 'reciprocally interconnected with both [the sets]'^{xx}. These representations are all held in brain's working memory. Subjectivity emerges during this third step when the brain is producing the third kind of image, that of an organism in the act of perceiving and responding to an object^{xxi}. (We can compare this third level representation with the Nyāya view of *anuvyavasāya* or *after-perception*).

Again, it is assumed that 'I' has a particular meaning in all languages - that is of being a single bounded entity. However, there are several languages where 'I' is hyphenated with many others. For example, in Nagaland, a state situated in the northeastern part of our country, the natives use a language where 'I' is hyphenated with spirits of their ancestors. So also in Africa there exists a language where 'I' is relational in nature.

Unity of Consciousness Argument

It is often claimed that positing an enduring self over and above the body is essential for otherwise we will not be able to account for the multifarious experiences as belonging to one unified subject. Whether it is Gautama^{xxii} or William James^{xxiii} both consider that only self can explain the *unity of consciousness* phenomenon. A tagged claim that comes with the unity of consciousness argument is that as our experiences belong to one unified subject this subject also possesses one body, that is, to say each self or subject can possess one corresponding body, and vice versa. Though it can receive stimulus from different sense organs and be conscious of those respective sensations (which includes visual sensation, auditory sensation, tactual sensation, etc.), still these sensations are received in the same body and hence belong to a singular self. This is to say that there exists a one to one relation between the self and the body. So the unity of consciousness argument ascribes to self the fundamental

function of binding and unifying all our experiences (sensations, feelings etc) into a harmony.

However, such a theory cannot accommodate cases where the subjects' actions, feelings and thoughts are not harmoniously unified, for instance, the cases of dissociated personalities or what we call Multiple Personality Disorder. Such disorders, we have seen, are mostly reported in cases of children who have been exposed to prolonged sexual abuse. To overcome their traumatic experience and to survive psychologically they construct more than one personal narrative. Each story though unified from the perspective of one personal narrative, it is totally different or discontinuous with the experiences mentioned in the other narratives. MPD cases also challenge the traditional assumption of one to one correspondence between body and self^{xxiv}.

Contemporary Scenario

Current research has proved again and again how all our experiences are not unified and continuous. Rather they are often random and discontinuous. We have a general tendency to unite all our experiences which itself has an evolutionary endorsement. However, this continuity often breaks down at smaller time scale. Max Wertheimer's experiment on *Phi phenomenon* conducted in the year 1912 proved this fact. If we recall the puzzling questions that the Phi phenomenon posed, we will see that subjects reported or rather constructed the first spot, say, the red spot moving towards the location of the second spot, say, the green spot and in the middle of this illusory passage the red spot abruptly changes into the green spot.

Dennett, in this context, discussed about two models to explain this strange phenomenon – the *Orwellian Revisionism Model* and the *Stalinesque Revisionism Model*. Both these models indicate or suggest that all our experiences are received as discrete representations initially. They are all held in brain's memory library or brain's editing room, till all the images are received. And finally they are projected into consciousness creating effect as if all experiences are continuous and unified. So it is wrong to suppose that all our conscious experiences are unified in the first place. Moreover, brain activities are capable enough to unify our experiences. Secondly, there is no need to posit an additional ontic self apart from body to explain the phenomenon of consciousness

We can analyze the traditional given self theorists' unity of consciousness argument, into two theses – The strong version of the thesis claims that wherever there is personhood, there is unity of consciousness, while the weak version of the thesis would claim that personhood is generally accompanied by unity of consciousness. MPD cases would counter the strong version of the thesis, while the experiments that show that our experiences are actually discrete, it is totally brain's activity that it produces an illusion that all our experiences are unified

especially when they are projected in the consciousness disproves the weak version of the thesis.

Experiencer Argument

The signals and stimuli that come from the external world are sometimes pleasurable, sometimes painful. Introspectible mental states like pleasure, pain etc., entail that there must be some experiencer. Now the question arises who is the experiencer – the self or the body? The traditional reply favours the former. Their argument runs as follows:

It is true that without body we cannot experience, yet the body is merely considered as one of the instruments of experience. The body being inert or natural like stone, table, chair, etc., cannot feel pleasure or pain. Therefore, it cannot be the experiencer. So there must be an immaterial entity that is the experiencer and this experiencer is the self.

Contemporary Debate

Contemporary researchers in general disagree with two important assumptions of the traditional Experiencer argument. First, they do not believe that consciousness cannot emerge from any physical or material cluster. Secondly, the experiencer argument does not consider the possibility that network or system where the neural activity occurs itself experiences. Rather contemporary theorists like Damasio, Patricia Churchland and others are moving towards a consensus that experience is nothing but representation and experiencer is the one whom the representation affects.

Now let us consider or identify three different cases and try to see how we can attribute consciousness to one while not to the other two.

Case A: Dead body

Case B: Living body

Case C: Watch

Behavioral Response Test

Let us try to observe behavioral response against a known set of stimuli, say, very low temperature in all these three cases and see how each reacts:

Case A: There will be no response in case of the dead man.

Case B: The living body will withdraw its organs.

Case C: The watch will stop working.

So, if we use this method for discriminating then we will not be able to discriminate or distinguish between the status of watch and the living body. Since both have shown some sort of behavioral response to stimuli in this case.

Self- Experience Test

Now if we depend on self-report of one's own experience, then we will find there exists two groups of entity – one group who can report (they are experiencers) and the other group who cannot report (they are not experiencers).

Now how can we be assured that the entities who cannot report cannot be experiencers?

For instance, pre-linguistic infants cannot report their experiences but we all know they do have experiences. This is evident from the EEG reports. However, one may raise objections here that EEG response itself is also a type of behavioral response but it does reflect a cognitive activity.

Moreover, animals possessing brains (even rudimentary brains) or nervous systems will respond to EEG or something like EEG.

Now the question arises what is EEG or EEG-like technology?

EEG is a technique, which measures activity/ procedure in the brain using highly sensitive recording. In other words, any EEG-like technology is a way of measuring voltage fluctuations of electrical activities. Now, any electric circuits can have voltage fluctuations and hence we can have an EEG-like graph for both a living system as well as, say, a computer or any system having an electric circuit. So what initially seemed to be a promising criterion for finding out whether a system is conscious or can have experience now turned out to be suffering from the charge of liberalism. Current researchers are now trying to find out a delimiting criterion that could distinguish between different kinds of EEG patterns. They are trying to chalk out a computational model that will help us to distinguish between two kinds of voltage plots – one plot will perhaps reveal that the system generating it has experience while the other do not.

However, it is a different issue whether current research can find a qualitative difference between the two types of voltage plots by working a mathematical value, which would determine that one is generated by a living system while the other by a non-living one. In fact, if we propose instead of a robust sense of self (like the traditional theorists) a minimal sense of self then we can associate it with any system - living or non-living based on its ability of (a) self-recognition, that is, identifying its respective boundary and thus in the process recognizing its 'others' and (b) devising techniques to resist all sorts of self-destructions. Many cognitive scientists like Dennett have thought in this line and argued that if we abandon the traditional anthropocentric notion of self and be a little liberal then we can not only ascribe self to all lower forms of life but also extend it to robots and other systems like it. Incidentally Dennett strongly believes that human beings are zombies.

The Self As Knower

Our thoughts, knowledge, etc. imply the existence of a knower or a thinker. Doubting or denial of a self is impossible without admitting a self. The traditional theorists can extend this line of argument in case of any activity, say, narration and say that there can be no narration without a narrator. However, Dennett has described a mechanism for generating narration without a narrator.

Contemporary Scenario

Dennett indeed conceived of one such situation, which breeds a narrative without a narrator. Dennett, we have discussed, offers a thought experiment consisting of a psychoanalyst who identifies the seeds of a story by asking a few questions having 'yes' or 'no' answers; but there is really no narrator of the story. Dennett extends this view to the notion of self thereby asserting that there can be 'selves' or the illusion of identity produced in individuals without accepting any foundational notion of self.

Such an attempt has been made by Oliver Selfridge in 1950 in the *Pandemonium Program* where he has conceived of computer programs, which can perform its functions without falling back on central authority for its decisions.

Again, the cases of anthills, termite colonies stand as evidence to the fact as to how creatures accomplish all the projects without positing any President Ant or Queen termite. Again Dennett's alternative way of explaining Pim Levelt's *Conceptualizer Model* bears witness to the fact how our systems can work efficiently without possessing a central authority.

Though all the above arguments come from the contemporary camp, similar line of arguments were forwarded by the Buddhist philosophers centuries ago while they argued in support of their anatta thesis (No-Soul doctrine).

The Self As Agent

The self in the traditional scheme is thought of as the agent who has knowledge of the materials and the means for attaining some end. So he also has goals, beliefs and desires and hence is conscious. Both western and Indian philosophical schools entertain such a conception of self. Thus we find in an older Nyaya text '*jnana janya bhabhet iccha - iccha janya bhabhet kriti- kriti janya bhabhet cestā janya bhabhet kriya*'. That is to say, cognition leads to desire, desire to will, will to physical effort and physical effort to action. However, amongst Indian schools of thoughts, there exists the *Sāṅkhya* school, which parts with this type of analysis. According to them, *Purusā*, the self, considered as Pure Consciousness, can never be an agent.

Again, agency is thought to be the correlate of responsibility. We attribute beliefs, desires, sensations and intentional actions to the self together with responsibility for those actions. Thus the self is not just an agent but is also a

moral agent who is liable to be blamed or praised for his actions. Objection was raised as to why moral responsibility is ascribed to the self when body is the performer of all our actions. First, the traditional theorists believed that body could not have beliefs, desires, etc. since it is inert. Moreover, the body is never free. It is bound and often performs actions under biological constraints. Unlike the body, the self has autonomy over what it does.

Contemporary Debate

Contra such views, current research tells us how we are always under the influence of culture and natural selection. All our decisions that traditional theorists thought to be taken by the self are actually determined by our genes and memes. A host of cognitive scientists believe that all our decisions are wired by heredity and experience. Human actions are nothing but a complex interplay between the memes^{xxv} and the genes. Hence there is no need to call for an extra self to explain such experiences. It is an illusion that we think we have an inner 'me', the conscious persistent self, who is the initiator of all our actions and does the deliberations. Now, different cognitive scientists hold different views regarding the notion of free will. Some like Susan Blackmore^{xxvi} supports the view that the notion of free will is totally false while others like Dennett accepts some version of free will and ascribes it to the illusory self. Blackmore asserts that it is the memes, which manipulate all our activities ushering in 'cultural and mental design' (opposed to free will).

Here the contemporary theorists anticipate the much raged objection that memetheory would reduce us to mere artifacts. To this they reply that our outlook towards the memes should not be like 'the dangerous memes'. In fact we should not devalue our meme-immunological system, since it is our 'better-than-chance' habit of selecting memes that has helped us not only to rise above our animal heritage but also increased our survival possibility.

Memory Argument

Another strong argument for accepting an enduring entity as self is for explaining memory. Memory, being a quality, must reside in a substratum. Now, neither sense organs nor the body can be regarded as a suitable substratum for memory. No sense organ can sense all sorts of objects. While the body undergoing constant change from infancy to youth to old age cannot also be regarded as the substratum, for then something perceived in one's youth cannot be remembered in one's old age. Since memory needs a permanent substance, the material body or the brain cannot be its locus.

Contemporary Debate

Now this whole memory argument proposed by the traditional theorists assumes that memory cannot be stored in any material medium. If one ceases to uphold

such an assumption, then the traditional argument for self becomes less acceptable. It is also clear from the memory argument that self presupposes memory. Autobiographical memory is necessary for self-representation. However, a few case studies reveal that autobiographical memory is really not essential. One can have self-experience and represent this experience but still is devoid of his autobiographical memory. It is reported in Damasio lab (in Iowa City) that a patient who is called R.B suffered from Retrograde Amnesia. Due to attack of *herpes simplex encephalitis*, R.B suffered massive destruction of both temporal lobes including the overlying cortical areas, as well as the deep structures including the amygdala and the hippocampus. The resulting effect was that he lost all his memories of the past (whether he was married or not, whether he had any children or not, etc.). In short, he had lost considerable amount of his autobiographical memory. Interestingly, R.B retained some of the important features of self-representation. The strongest evidence in support of such a claim is that he would without any extra effort refer to himself with the personal pronoun 'I'.

The leading cognitive neuroscientist Michael Gazzaniga talks extensively about false memories. We come to know from his work how our brains interpret or process data long before we come to know about it. What we see now is not what our brains see now. Our brains are much ahead of us as far as experiencing is concerned. Our brain mechanism controls our cognition to such an extent that 'By the time we think we know something – it is part of our conscious experience - the brain has already done its work. It is old news to the brain, but fresh to us'.^{xxvii} This also reveals that our memories are nothing but reconstructions to fill out the gaps in the narratives our brain generates. In a way Gazzaniga questions the validity of our memories.

Memory Argument: Buddhist View

Apart from the evidences of the contemporary research, the Buddhists had a series of arguments to refute the claims of the memory argument. They would argue that the evidence of memory could not be invoked or imported to suffice for the proof of the self because an enduring self and the phenomenon of memory involve each other. Memory can be explained on the supposition of an enduring self and proof of the self is explained by the fact of memory. So there is *anyonyāśrayatva* or mutual involvement.

Here ends briefly my discussion on the various grounds and arguments both in traditional and contemporary times in the different quarters of philosophy regarding the existence/non-existence of self.

Section.3

Neurological reductionism

After discussing the grounds and arguments for accepting a special version of the self, let us go straight to the theory of self we have tried to develop or more humbly the theory of self that actually emerges by collating these arguments.

As I have mentioned in the first section of this paper that I subscribe to a particular variety of reductionism. My commitment to cognitive science has restrained me from accepting an extra add-on entity in any form called self apart from the brain body complex, thereby supporting no version of non-reductionism. Again, eliminativism has never been a hot favourite with either the cognitive scientists or the philosophers. So reductionism is really talk of the day. From evolutionary biology to cognitive neuroscience, importance of self-talk is recognized. A reductionist can indulge in self-talk without succumbing to an ontic self, though initially reductionism was seen with much suspicion as a possible third alternative stance one can take with respect to the ontological status of selves.

Now that philosophers have started taking interest in exploring this possibility we should enquire into the different senses of reductionism endorsed by different philosophers. The self-theory as propounded in our project will be a type of *neurological reductionist theory*. What I mean by *neurological reductionism* is the view, which rejects or denies the existence of a higher order reality apart from the facts about our neural representations with regard to selves. In the subsequent section we will explore how this program could be carried out without falling back on any other concept of self traditionally associated with the term in its ordinary employment. Such a stance would come closer to that of Dennett, Damasio and many others. Committing oneself to this particular variety of reductionism, it becomes necessary to explain the reason why human beings construct such a concept as self apart from the brain-body complex. The rest of my self-theory would cover two dominant issues in hand – the reason behind our construction of such a notion of self-hood, and the several factors responsible for construction of our identity.

The cognitive scientists are now visualizing self-talk as a natural phenomenon, which has a strong evolutionary backing. However, we should never overlook the social aspect in the sense that social factors too contribute to the phenomenon of self-talk. It is recognized widely in both social sciences and cognitive neuroscience that the self and its social world are reciprocally determined.

Initially it was thought that this whole ‘naturalization project’ was impossible. The whole phenomenon of conscious self-hood emerging from a series of causal chain of events in the physical world was not entertained until the last couple of decades. Huge body of neuroscientific data collected by researchers working in

brain science all point to the fact that the existence of an autonomous given self is an illusion created by the physical processes in neuronal networks in our brain. Today a host of brain scientists like Damasio *et al* charts out a detailed representational, functional and computational structure of the self-hood illusion or rather the neural, functional or computational basis of self.

At this juncture, we have to distinguish between three different levels, which bear upon the self, and the self-talk, viz.

- A) the level of neural representation
- B) the level of the self
- C) the level of the self talk

Now here I must give a note of caution that mere taxonomy of the different levels of representations of the self as it is stated above does not necessarily commit us to the reality of any one of them.

Since I am taking a reductive stance in my thesis, some of the levels will be reduced to the others. It is clear from what I have already said that my contention is not to eliminate the self-talk as the eliminativist does.

My basic aim being to explain the possibility of self-talk without accepting an ontic self underlying all ascriptions of self as the centre of narrative gravity, I will concentrate only on the reduction of the second level to the third level. And with regard to the qualia and other mental representations, I mention only in passing that they bear strong correlation with the first level, the level of neural representation. So I would rather take a reductionist stance and say that the mental representations like qualia, etc. can be also reduced to the levels of neural representations.

This discussion is important because it brings out why I am uneasy with the Buddhist thesis of mereological reductionism. Let us ask the Buddhists – whether they accept the reality of the *nāma-skandha-s* like *samskāra*, etc., without accepting their correlation (if not dependence) on neural representations? If they answer affirmatively, then their thesis will run head long with the reports of the current research done in brain science. On the other hand, if they answer negatively, that is to say, these *nāma-skandha-s* depend on the neural representations for their existence, then the Buddhists will be committed to some sort of dualism. Here it is maintained that the *nāma-skandha-s* like *samskāra*, etc. are ‘mental’ in some non-physicalist sense of the term. On the top of that if the Buddhists accept a physical basis of these mental states, then they are committed to a form of dualism. And hence all the charges that are raised against the add-on theories (non-reductionism/ dualism) will be raised against them. Therefore, the Buddhist rejection of the add-on thesis (self as a psychological construct) remains half-hearted.

Arguing against the mereological reductionism of the Buddhist sort, I have shown why I am inclined not to accept the reality of mental representations. What emerges from these current researches is that thought, experiences, actions, etc. all exist as neural representations and though there is thought, experiences, actions, etc., there is no thinker, experiencer, actor, etc.

References

1. Strawson, Galen., "The Self, Models of the Self", (ed.) Shear Jonathan, Imprint Academy, 2000, United Kingdom, p.6
2. Locke, John., "Essay", (ed.) John Perry Berkeley, University of California Press, 1975, California, p.2
3. These theorists believe in the real existence of a self as an entity over and above the body. Here after these theorists will be referred to as SAG theorists.
4. These theorists believe that no such entities as self exist and such entities are fictions, mere constructions. Here after these theorists will be referred to as SAC theorists.
5. For relevant discussion on Parfitian notion of self see *Reasons and Persons*.
6. However non-reductionist theses on self are not of a homogeneous variety.
7. Hume, David. "A Treatise of Human Nature", (ed) L.A. Selby Bigge, Clarendon Press /OUP, 1978, Oxford.
8. Siderits, Mark., "Personal Identity and Buddhist Philosophy: Empty Persons", Ashgate Publishing Company, 2003, Vermont.
9. Giles, James., "The No-Self Theory: Hume, Buddhism and Personal Identity, Philosophy East and West", Vol. 43, No. 2, University of Hawaii Press,1993, Hawaii.
10. Ibid 8 pp.3-4
11. Ibid 8
12. Ibid 8 pp.82-83
13. Damasio, Antonio., "Descartes Error", Harper Collins Publishers, 2000, New York, p.225
14. Gazzaniga, M., "The Mind's Past", University of California Press, 1998, California
15. Le Doux, Joseph., "The Synaptic Self : How our Brains Become Who We Are", Penguin Putnam Inc, 2002, New York.
16. Churchland, Patricia., "Self-Representation in Nervous Systems", *Science*, Vol 296.

17. Dennett, Daniel., “The Self As The Centre of Narrative Gravity”, *Times Literary Supplement*, September, 1998.
18. Vedanta, an Indian school of thought propounds such a theory of self.
19. Dawkins, Richard., “The Extended Phenotypes”, Freeman, 1982, Oxford.
20. Ibid.
21. Eko basi sarvabhutantaratma ekam rupam bahudha yah karoti”, *Kathha Upanisad*
22. Damasio, Antonio., “Descartes Error”, Harper Collins Publishers, 2000, New York.
23. Perceiving should be generalized into any kind of activity and an object here encompasses responses to any external stimulus.
24. darsana sparsanabhyamekarthagrahanat’, *Nyaya Sutra*, 1/199
25. James, William., “The Principles of Psychology”, 1890, Dover Publications, England.
26. Especially for those cognitive scientists like Dennett and others who conceives of self in terms of a narrative.
27. Memes are units of culture. Blackmore, Susan., “The Meme Machine”, Oxford University Press, 1999, Oxford.
28. Blackmore, Susan., “The Meme Machine”, Oxford University Press, 1999, Oxford.
29. Gazzaniga, Michael., “The Mind’s Past”, University of California Press, 1998, alifornia, p.25