

# THE BUDDHIST TURN IN CONTEMPORARY PHILOSOPHY OF MIND

**Napoleon M. Mabaquiao, Jr.**  
*De La Salle University, Manila*

*Contemporary philosophy of mind is generally characterized by its project to naturalize the mind. Utilizing the findings of the different sciences involved in cognitive science, especially those of artificial intelligence and neuroscience, it continues to explore ways to explain the workings of the mind in purely scientific terms. But despite the rigor and sophistication of its methods, certain questions critical to its success have remained unanswered, such as how consciousness emerges from the brain's physical processes and how the phenomenal properties of our conscious experiences arise from the physical properties of our bodily experiences. This has led some scholars to seek alternative perspectives. One such perspective that is widely explored today is Buddhist thought. The centrality of the mind in Buddhist thought and its perceived compatibility with the findings of modern science make it an attractive alternative framework to carry out the naturalization project. In this paper, I aim to examine the plausibility of this strategy. In particular, I shall evaluate whether Buddhist thought provides the needed insights to overcome the challenges facing the said project.*

*Keywords: mind, Buddhism, cognitive science, computationalism, hard problem of consciousness*

“All phenomena are preceded by the mind.  
When the mind is comprehended,  
all phenomena are comprehended.”  
Buddha<sup>1</sup>

## INTRODUCTION

Contemporary philosophy of mind is primarily geared towards the naturalization of the mind or, what comes to the same, the assimilation of the mind into the scientific world-view. Motivated by current advances in the areas of artificial intelligence and neuroscience, most current philosophical investigations concerning the mind focus on exploring ways by which the mind can be explained in purely scientific terms. This project to naturalize the mind, also known as the *naturalization*

*project*, has a two-fold objective. On the one hand, it intends to give rigor to how we think about minds and their processes. On the other hand, it aims to complete the so-called scientific principle of physical closure, according to which all phenomena occurring in this world can be sufficiently accounted for using only physical explanations.

However, while this project has surely made progress in explaining the functionality of the mind, generally referring to the system of inputs and outputs of mental processes, it is still plagued with difficulties when accounting for the mind's consciousness and its related properties. The so-called *hard problem* of consciousness, referring to the problem of showing how consciousness arises from the physical processes of the brain's neurons, is far from being resolved (See Chalmers 1995). The epistemic gap between physical facts and phenomenal facts, referring to the problem of how knowledge of phenomenal facts can be inferred from knowledge of physical facts, has not been satisfactorily closed (See Mabaquiao 2015). The usual reaction of proponents of the naturalization project to these difficulties is to simply ignore them. They allege that either these difficulties are not really essential to the understanding of the mind (the general position of materialism), or they are pseudo-problems for allegedly being grounded in mistaken assumptions or in a false theory about the nature of the mind (as argued, for instance, by the Churchlands in their theory of eliminative materialism) (Mabaquiao 2015, 29-31). On the other hand, there are also those who argue that the solutions to such difficulties lie beyond the natural limits either of our cognitive abilities (See McGinn 1997) or of our language as a representational system (See Mabaquiao 2009). As such, while they support the materialistic framework of this project (for maintaining that what lies beyond our cognitive limits is still physical in nature), they do not believe that the naturalization project will succeed.

Still, there are those who go beyond the usual confines of Western philosophy and modern science to handle these issues. They explore alternative perspectives, most especially those from the East. If not the Eastern philosophies themselves, these perspectives are either rooted in or inspired by such philosophies. A perspective that is widely recognized and whose appeal to philosophers and cognitive scientists alike continues to grow is Buddhism or Buddhist thought (or philosophy) (See Arnold 2012 and Spackman 2012). There are two main reasons, among others, for this: first is the centrality of the mind in the overall Buddhist philosophy (as evidenced by its focus on training the mind as the way to achieve its spiritual goal of enlightenment), and second is the perceived compatibility of the central tenets of Buddhist philosophy with the findings of modern science (especially in the area of quantum physics). In this paper, I aim to critically examine the alternative strategy to overcome the obstacles of the naturalization project using insights coming from Buddhist philosophy. In particular, I intend to support the view that argues for the non-feasibility of this strategy. As a framework for this position, I shall advance an account of the Buddhist philosophy of mind, which is a dualism of the sort that deviates from the usual conceptions of it in traditional, Western philosophy.

I shall divide my discussion into three parts. In the first part, I will introduce the project of naturalizing the mind. I shall look into its motivations and the various ways by which it has been proposed to be carried out by philosophers of different

metaphysical persuasions. I will also discuss in this part some of the major questions that have been raised concerning the feasibility of this project. In the second part, I will examine the strategy of using Buddhist thought to overcome the major obstacles of the said project. I will first show why it has seemed natural for cognitive science to consider Buddhist thought as an alternative approach to handle its difficulties. In the third part, I will critically assess the various viewpoints on the use of Buddhist thought as a framework for cognitive science. After pointing out their gaps and weaknesses, I shall argue for an alternative view.

## NATURALIZING THE MIND

In addition to our physical activities, we go about our everyday lives engaging in mental activities such as thinking, making decisions, wishing, imagining, and experiencing pains and pleasures. We usually attribute our physical activities directly to the workings of our bodies, though we also believe that some of these activities, the ones we usually call “voluntary actions,” may have been brought about, directly or indirectly, by some mental processes. On the other hand, as regards our mental activities, we usually attribute them to the workings of our minds. However, as we continue to know a great deal about our bodies and the natural environment with which they interact, we remain to know very little about our minds. The irony is that while it is with our mental states that we are most immediately and directly acquainted with, our mind is still largely shrouded in mystery (Dennett 1991, 21-22). However, our minds or conscious states undoubtedly form an important part of what defines who we are as humans and as individual persons. For it is essentially in reference to our mental attributes such as our rationality, freedom, and consciousness that we distinguish ourselves from the rest of nature; and to our individual memories, beliefs, and attitudes that we distinguish ourselves from one another.

Previously it was thought that science could afford to ignore the workings of the mind or the nature of consciousness in its investigations of nature. The pervading view then seemed to be that the mind has no critical place or role in a systematic and comprehensive explanation of the natural world. As it were, it was presupposed that whatever the nature of the mind is would not really make a difference in the scientific account of the world. As such, it was thought that the mind was a subject matter better left to the speculations of philosophers and theologians. However, somehow the idea that the scientific account of nature cannot be truly complete unless it also deals with the nature of the mind or consciousness persists. The preeminent mathematician and physicist Roger Penrose (1997, 8) put this more clearly when he remarked: “A scientific world-view which does not profoundly come to terms with the problem of conscious minds can have no serious pretensions of completeness. Consciousness is part of our universe, so any physical theory which makes no proper place for it falls fundamentally short of providing a genuine description of the world” (See also Chalmers 1997, xi).

This sentiment articulated by Penrose has slowly been asserting itself, as evidenced by the growing number of scientists and mathematicians who have joined

the materialist-oriented philosophers in the attempt to come up with a scientific account of the workings of the human mind. This project to explain the mind in purely scientific terms or, more simply, to naturalize the mind has two main goals. The first, which is a precondition for the second, is to provide rigor in the way we theorize about the mind. This is a reaction to the speculative ways by which the mind had been investigated and explained in the past. The second is to carry out the completion of the scientific account of the natural world. As Penrose earlier noted, if the mind is part of the natural world, then a complete account of the science of this world should include the mind. The mind, in this regard, is seen as the last piece of the puzzle to complete the scientific picture of nature.

However, initial efforts to carry out the naturalization project have been met with doubts and skepticism. The mind did not seem at the time to lend itself to the scientific methods of observation and quantification. However, with the entry of computer technology into the scene, things have taken a revolutionary turn. With its enormous capacity for simulating a wide variety of phenomena, thereby proving to be a very powerful tool for understanding the mind, the dreamt-of science of the mind has generated high hopes. The use of computers in understanding the human mind culminated in the development of a theory called the *computational theory of mind* (henceforth, simply *computationalism*), according to which the human mind is a sophisticated kind of computer and that the computer of appropriate sophistication has itself a mind. More specifically put, the human mind is seen as a sophisticated software implemented by the hardware of the brain; and that the sufficiently sophisticated computer software is itself a mind. The seriousness in which the computationalist project is being pursued has given birth to the development of a new science, a projected interdisciplinary science of the mind, involving the disciplines of philosophy, psychology, linguistics, artificial intelligence, neuroscience, and anthropology, which has come to be known as *cognitive science* (Freidenberg and Silverman 2006, 6).

With computationalism as the reference point, we can strategically divide the development of this projected science into three stages, namely: the *pre-computational*, *computational*, and *post-computational stages* (See Mabaquiao 2009, 481-487; 2012, 193-222). The theories at the pre-computational stage are directly reacting to *Cartesian dualism*, which divides reality into the mutually exclusive physical and mental realms. According to this dualism, reality consists of matter and minds: matter being spatially extended but unconscious, while minds being conscious but non-spatially extended. The critical consequence of this dualism is that there can only be a science of matter but never of the mind. That being the case, one important preliminary task in establishing a science of the mind is to demonstrate the falsity of Cartesian dualism as a theory of reality. Two theories serve as paradigm representatives of this stage, which reject the Cartesian independently-existing mental realm, namely, *behaviorism* and (*mind-brain*) *identity theory*. For behaviorism, expressions allegedly referring to mental states are reducible to expressions merely referring to behavioral dispositions. On the level of reality, behaviorism claims that there really are no mental states for what we have; in place of these states are mere behavioral dispositions. On the other hand, identity theory claims that expressions

allegedly referring to mental states are merely non-scientific equivalents of scientific expressions referring to brain states (or the brain's neural states). On the level of reality, this theory claims that there really are no mental states for what we have; in place of these states are merely neuronal states of the brain.

At the computational stage, computationalism becomes the dominant view of the mind. In the discipline of philosophy, computationalism is the logical offshoot of *functionalism*, that theory of mind which defines mental states in terms of the causal relations among input, internal, and output states. In the discipline of artificial intelligence, computationalism has come to be known as *strong AI*. Accordingly, a distinction has been made (mainly due to John Searle 1980) between the views of strong AI and weak AI. Strong AI boldly claims that the human mind is a sophisticated computer (more specifically, it is a computer software run by the hardware of the brain), and the sufficiently sophisticated computer itself has a mind (or, more specifically, the sufficiently sophisticated computer software is itself a mind). In contrast, *weak AI* merely claims that the computer serves as a very powerful tool for understanding how the human mind works. It has no metaphysical claims about the mind. Computationalism presently has two dominant models: the *classical* or *symbolic model*, which defines mental computations as manipulations of symbols according to certain rules, and the *connectionist model*, which defines mental computations as interactions among the units comprising neural networks. However, in either model, the computationalist view that thinking, as Zenon Pylyshyn (1989, 51) has put it, is “a species of computing” is maintained.

Finally, the theories at the post-computational stage are those critically reacting to computationalism. They consider the account of computationalism as faulty due to its leaving out one or more of the following fundamental features of the mind: consciousness, intentionality (referring to the directedness of mental states towards some objects or states of affairs in the world), and qualia (referring to the subjective or the “what-it-is-like” qualities of our conscious experiences). Consequently, these theories see the actualization of the projected science of the mind in the possibility of a science that accommodates these features of the mind. Foremost of the post-computational theories of mind are the *biological naturalism* of Searle, which regards mental states as higher-level biological states that are caused by but not reducible to the causal properties of the brain, and the *quantum view of consciousness* of Penrose (1994), which regards mental states as quantum states in the microtubules of the neurons. Before advancing their respective theories, both philosophers have presented arguments against computationalism. Searle presented his Chinese room argument, according to which human minds can never be computers because while human minds are aware of what their mental states refer to in their world (their intentionalities), computers are never aware of what the symbols that they manipulate refer to in the world. Penrose (1994), for his part, presented an argument based on *Gödel's incompleteness theorem*, according to which human minds can never be computers because while humans can recognize the truth of a proposition outside of the rules of the system in which the proposition belongs, computers cannot (for the actions of computers are inescapably rule-governed).

Theories in both computational and post-computational stages are attempts to come up with a science of the mind using different models. These theories, needless to say, have encountered challenges or criticisms. Let us, in this section, briefly examine some of these challenges. The arguments used to challenge these theories of mind come in various forms. Some arguments point out certain inconsistencies in their key ideas. Behaviorism, for instance, is shown to be inconsistent or guilty of circular reasoning for eventually making references to mental states when pressed to explain why different persons respond with different behavioral dispositions to the same input stimuli. It claims that mental states are nothing but behavioral dispositions, but in explaining why we have such and such behavioral dispositions and not otherwise, it refers to how our mental states affect our behaviors. Another example is the (type) identity theory which is charged with *neuronal chauvinism*—referring to the undue preference for the human brain. The identity theory claims that mental states are nothing but the neural states of the human brain, which implies that other organisms, such as animals, are not endowed with the same kind of brains that humans have will never have mental states. But this conflicts with the common-sense view that animals too have mental states though their brains may be different, structurally and/or materially, from ours.

Some arguments point out that the theories intended to explain the mystery of the mind just generate further mysteries. For instance, Cartesian dualism is often criticized for making the mind-body relation as something utterly mysterious. For how can two qualitatively different substances, the physical substance which is governed by the deterministic laws of nature, and mental substance, which is not, causally interact with one another? However, most alternative accounts of the mind-body relation, which aim to sidestep the Cartesian mystery, have created further mysteries, which seem more serious than what they intend to explain away or resolve. For instance, Leibniz's parallelism tries to explain away the Cartesian mystery by saying that there really is no mind-body causal interaction at all and that the seeming causal interaction between them is just an illusion brought about by the pre-established harmony set by God (Look 2020). It may have sidestepped the Cartesian mystery, but it has created more mysteries. Among others, what is the point of God creating a totally deterministic world where morality and responsibility have no place? Another alternative account to Cartesian dualism is the so-called *epiphenomenalism*, which claims that while mental states are indirectly caused by the physical states of the body, mental states, on the other hand, do not have any causal powers. However, having no role to play in our survival, it appears mysterious why we have mental states in the first place.

Lastly, some arguments point out that something fundamental about the mind is left out in the explanations of certain theories. For instance, Searle (1980) argues that functionalist and computationalist accounts of the mind leave out the inherent property of consciousness to be intentional or to be directed towards objects or facts in the world. On the other hand, Thomas Nagel (1974) and Frank Jackson (1991), in their knowledge arguments, argue that the said accounts, along with other materialist theories of the mind, leave out the qualia, referring to the subjective (or that "what-it-is-like") qualities of our subjective experiences. The central point of their arguments

is that if the nature of mentality is exhausted by some physical mechanism, then just by knowing this mechanism, we would know everything about mentality. However, this is simply not the case. For instance, saying that pain is C-fiber stimulation tells us nothing about how pain feels or what it is like to be in pain. This failure of purely physical explanations to account for the subjective qualities of conscious experiences has been called by Joseph Levine (1983) the *explanatory gap*.

## BUDDHISM AND COGNITIVE SCIENCE

As the problems besetting the naturalization project have remained unresolved or have not been satisfactorily answered within the areas of traditional (Western) philosophy and modern science, alternative perspectives have been explored. One that has attracted a great number of scholars is Buddhism. And this is for at least two reasons: the centrality of the mind in overall Buddhist thought and the *perceived* compatibility of the central tenets of Buddhism with the findings in modern science, most especially in the area of quantum physics.

The centrality of the mind in Buddhist thought can immediately be demonstrated through the *four noble truths*, as Allan Wallace (2002, 16) does:

The first task is to recognize the nature and full range of suffering to which humans are vulnerable. The first noble truth formulates that as the problem to be addressed. The second noble truth presents the hypothesis that the essential causes of suffering are to be found within the mind, specifically in terms of cognitive, emotional, and attentive imbalances. The third noble truth hypothesizes that these afflictive tendencies can be irreversibly dispelled from the mind. And the fourth noble truth presents detailed procedures for collecting data by observing mental processes and experimenting with techniques for transforming the mind and eliminating its afflictive elements.

Accordingly, craving (referring to the kind of desire that leads a person to be attached to the objects of his/her own desires), which arises from ignorance, is the cause of suffering; and thus, to eliminate suffering is to eliminate craving and ignorance. Now since craving and ignorance are mental phenomena, the way to eliminate them would be to train the mind using techniques described in the *eightfold path*. Buddhist scholar Peter Santina (1984, 64) tells us that among these techniques, the three that directly deal with mental development are *right effort*, *right mindfulness*, and *right concentration*. Briefly, right effort “means cultivating a positive attitude towards our undertakings” (Santina 1984, 65). More specifically, it refers to the effort to free the mind of unwholesome thoughts and fill it up with wholesome ones. Right mindfulness is “awareness or attention, avoiding a distracted and clouded state of mind” (Santina 1984, 68). Right concentration, also called meditation or tranquility, is “the practice of focusing the mind single-pointedly on a single object” (Santina 1984, 68).

The centrality of the mind in Buddhist thought can also be seen in its *no-self doctrine*, which rejects the idea of the self as a permanent and independently-existing substance or entity, which is popularly known in Western philosophy as the *Cartesian self*. In place of this Cartesian self, Buddhism puts forward a notion of the self as a mere convenient collection of certain physical and mental factors—the so-called *five aggregates* that account for the possibility of personal experience. The motivation behind the no-self doctrine is clear: eliminating the belief in a permanent, independently-existing, and substantial self eliminates our attachment to our own personal desires, which in turn eliminates our attachment to existence in general and eventually leads to the recognition of the impermanence of things. In this consideration, the no-self doctrine is sometimes regarded as the core of Buddhist thought. Santina (1984, 127) explains:

Once we identify, imagine, or conceive of ourselves as an entity, we immediately create a schism, a separation between ourselves and the people and things around us. Once we have this conception of the self, we respond to persons and things around us either with desire or aversion. In this sense, the self is the real villain of the piece. Seeing that the self is the source and cause of all suffering, and seeing that the rejection of the self is the cause of the end of suffering, rather than trying to defend, protect, and preserve the self, why should we not do our best to reject and eliminate this idea of the self?

With regard to the compatibility of Buddhist views with the findings of modern science, one classic demonstration of this can be found in Fritjof Capra's book *The Tao of Physics* (1975). In this book, Capra shows how Eastern mysticism, represented by Taoism, Hinduism, and Buddhism, has anticipated the findings of modern science, such as those in quantum physics and the theory of relativity. One such finding is the idea in quantum physics that reality is an interconnected whole. Capra (1975, 140-141) particularly notes that this finding has been anticipated in Buddhist thought:

Quantum theory has abolished the notion of fundamentally separated objects, has introduced the concept of the participator to replace that of the observer, and may even find it necessary to include the human consciousness in its description of the world. It has come to see the universe as an interconnected web of physical and mental relations whose parts are only defined through their connections to the whole. To summarize the world view emerging from atomic physics, the words of a Tantric Buddhist, Lama Anagarika Govinda, seem to be perfectly apropos: "The Buddhist does not believe in an independent or separately existing external world, into whose dynamic forces he could insert himself. The external world and his inner world are for him only two sides of the same fabric, in which the threads of all forces and of all events, of all forms of consciousness and of their objects, are woven into an inseparable net of endless, mutually conditioned relations."

Santina's (2002, 8) following remarks shed further light on this point:

This growing interest in Buddhism and these many areas of affinity between the teachings of the Buddha and the tendencies of modern science, philosophy, and psychology have reached their apex at this very time in the suggestions now proposed by quantum physics, the latest developments in experimental, theoretical physics. Here too, we find that not only is the method of science—observation, experiment, and analysis—anticipated by the Buddha but that some of the very specific conclusions about the nature of man and the Universe that are indicated by the latest developments in quantum physics were also indicated by the Buddha. For instance, the importance of the mind. A noted physicist not long ago has remarked that the Universe is really something like a great thought. And it is said in the Dhammapada that the mind precedes all things, that the mind is the maker of all mental states. Similarly, the relativity of matter and energy is mentioned. There is no radical division between mind and matter. All these indications are now gradually being revealed by the latest developments in science.

Now, since the mind is central to Buddhist thought and Buddhist thought in general and given its *perceived* compatibility with modern science, it is natural to suppose that Buddhist thought must have something significant to contribute to the project of cognitive science to naturalize the mind. As Richard Payne observes in his essay “Buddhism and Cognitive Science: Contributions to an Enlarged Discourse” (2002, 9): “Given Buddhism’s long history of placing primary emphasis on consciousness in the transformation of human existence, the future development of Buddhist thought would seem to necessarily entail a continuing dialogue with cognitive science.”

However, how exactly can Buddhist thought contribute significantly to the development of cognitive science? It is not enough to show that there are parallels between the Buddhist concept of the mind and the scientific concept of the mind that is pursued in cognitive science, like, for instance, both concepts reject the Cartesian self. The various materialist theories of mind all reject the Cartesian self, but they all seem to face the same problems. What is needed is to show that Buddhist thought can help cognitive science overcome its difficulties. In this regard, in the following section, we shall critically examine the main competing viewpoints on whether and how Buddhist thought can contribute to the goals of cognitive science.<sup>2</sup>

## THE CONTENDING VIEWS

One viewpoint has it that Buddhist thought has something positive to contribute to cognitive science by way of methodology. Accordingly, what is lacking in cognitive science is a refined instrument by which to study the mind scientifically. As Wallace

(2002, 16) writes: "...the modern West has developed a sophisticated science of behavioral and neural correlates of consciousness, but no science of consciousness itself, for it has failed to develop sophisticated, rigorous means of exploring the phenomena of consciousness first-hand. And this is the first step towards an empirical science of any class of natural phenomena." According to Wallace, Buddhism supplies this needed refined instrument in the form of the mind as refined by a certain type of training—the so-called *meditation*. Wallace explains: "The only instrument humanity has ever had for directly observing the mind is the mind itself, so that must be the instrument to be refined. The untrained attention is habitually prone to alternating bouts of agitation and dullness, so if the mind is to be used as a reliable tool for exploring and experimenting with consciousness, these dysfunctional traits need to be replaced with attentional stability and vividness." Wallace, however, thinks that the kind of science of consciousness that will emerge here will not be the one aimed at by cognitive science, which is a physicalist science of the mind. Meaning to say, if cognitive science were to adopt the Buddhist way of studying the mind, then it would have to abandon its physicalist stance. As Wallace (2002, 30) writes in the conclusion of his paper: "It took the scientific community fifty years to recognize that the mind couldn't meaningfully be reduced to a set of behavioral dispositions. Hopefully, it will not take that long before neuroscientists open their minds to the possibility that the mind may not be meaningfully reduced to neural mechanisms either."

Another viewpoint states that Buddhist thought has nothing positive to contribute to cognitive science because the goals of Buddhism and cognitive science are in stark contrast. This viewpoint is well articulated by Kamran Karimi et al. in their essay "A Buddhist Perspective on Artificial Intelligence" (2001). According to Karimi *et al.*, while the goal of Buddhism is to develop the mind to achieve enlightenment, the goal of cognitive science is to advance a computational model of the mind. What makes these goals at odds with one another is the fact that the computability of the mind is an essential part of what should be transcended to achieve enlightenment. In this connection, Karimi et al. (2001, 1) write: "Buddhist beliefs are in stark contrast to the current approaches in AI because in Buddhism the ultimate goal is to stop the calculating mind and thus achieve enlightenment. An enlightened person does not show signs we traditionally attribute to intelligence." They further state that "even though researchers may come very close to emulating a human mind, the result will probably never gain enlightenment. This is because enlightenment cannot be achieved with an intelligent mind that calculates and follows rules" (2001, 4). This means that the human mind that AI intends to simulate, namely the calculating mind, is that "ignorant" state of the human mind that Buddhist enlightenment intends to transcend.

This viewpoint is also endorsed by Asaf Federman in his essay "What Buddhism Taught Cognitive Science about Self, Mind, and Brain" (2011). According to Federman, as a result of the opposing goals of Buddhism and cognitive science, the alleged affinities between these two systems concerning the nature of the mind are actually naïve or superficial. For instance, Federman argues that while Buddhism and cognitive science are alike in rejecting the Cartesian self, the motivations for this rejection are different: it is practical for Buddhism while it is theoretical for modern science. Speaking of the systems of Buddhism and cognitive science, Federman (2011, 47) explains:

Both systems are united in rejecting the idea of an eternal, immaterial soul that is separated from the body, and yet controls it from within. But this is more or less where the similarities end. Buddhism rejects a kind of self (atman) that is eternal, blissful, and identical with the creative force of the universe (Brahman). It identifies the attachment to such a self as a source of misery and thus provides logical considerations (philosophy) and practical exercises (meditation, morality) as antidotes. The prime motivation is therefore ethical and is part of a larger soteriological system. Cognitive science, on the other hand, lacks both the ethical motivation and the practical application. Its rejection of the Cartesian self is a result of scientific theorizing that emerges from experimental work....

The third and last viewpoint claims that Buddhist thought can critically dialogue with cognitive science by way of providing alternative models of consciousness. Richard Payne articulated this view in his essay “Buddhism and Cognitive Science: Contributions to an Enlarged Discourse” (2002, 5):

The project of juxtaposing Buddhist psychology and cognitive science should not be seen as an uncritical acceptance of the authority of science, but rather as a dialogue in which the phenomenologically-based teachings of Buddhism concerning the workings of the mind can interact critically with cognitive science. Nor is it an attempt toward some grand resolution demonstrating the perfect harmony of Buddhism and cognitive science. And, finally, it is not an appeal to the authority of cognitive science as legitimating Buddhism—an all too common project in the discourse on science and religion. Rather, the goal is twofold: to find a critical perspective from which to evaluate the contemporary relevance of traditional Buddhist teachings, and to provide Buddhist models of consciousness and its transformation critically different from those already under consideration in cognitive science.

Now, whether the contribution that will result in this dialogue is positive or negative depends on the kind of alternative model of consciousness being advanced. John Spackman, in his essay “Contemporary Philosophy of Mind and Buddhist Thought” (2012), identifies the following Buddhist-based/inspired models of consciousness: (a) Wallace’s non-materialist view, (b) Flanagan’s materialist view, (c) Thompson’s phenomenological approach, and (d) Spackman’s own neutral non-dualism. Spackman, in his paper, aims to show how the first three models fail to overcome the same difficulties faced by the standard models in cognitive science; and, consequently, how his own proposed model, the fourth model, is able to overcome these challenges. In what follows, let us examine Spackman’s analyses in this regard.

Wallace’s non-materialist view, called *consciousness-based view* by Spackman, is built on a Buddhist idea that reality ultimately consists of a primordial consciousness, “which is wholly undifferentiated, without qualities or

conceptualizations” (Spackman 2012, 743) such that “at this level, subject and object are entirely transcended” (ibid.) thereby forming a unity. Because it claims that reality is fundamentally mental, this view is classified as a form of idealism, but it varies from traditional forms of idealism precisely because it rejects the distinction between the subject and object of consciousness (ibid., 742).

Flanagan’s materialist view is built on the idea that it is possible to naturalize Buddhism in that it is possible to take “from Buddhist accounts of the mind only what is compatible with physicalism” (ibid., 745). This view has been recently defended by Mark Siderits in his thought-provoking essay “Buddhism and Techno-physicalism: Is the Eightfold Path a Program?” (2001). Siderits argues that computationalism (which he refers to in his paper as “techno-physicalism”) provides a good framework for explaining some main Buddhist tenets, such as impermanence and nonself. According to Siderits, impermanence and nonself are very much coherent with the materialism of computationalism, which presupposes no underlying permanent reality. On the other hand, the computationalist idea that the human mind, like computer software, is just a collection of programs is able to explain the Buddhist notion of nonself where the self is taken as a mere collection of various kinds of elements. However, more importantly, the computationalist model is able to explain the problematic Buddhist issue of how to explain the possibility of rebirth, given that there is no enduring self. For who is reborn if there is no enduring self? Siderits very cleverly points out that this question can be answered by the principle of multiple realizability of computer softwares. Accordingly, if minds are computer softwares, then they can be transferred from one kind of hardware to another as long as these hardwares have the necessary sophistication to run the softwares. Siderits (2001, 311) explains:

And once again, the computer metaphor helps us make sense of the Buddhist claim, helps us see how rebirth might be possible without a transmigrating self. For we can see how one computer might, in going out of existence, set in motion causal processes resulting in some new computers having states importantly related to those of the original. Rebirth might be the organic equivalent of using a Zip drive just before the final crash of the old computer, then installing selected files on a new machine.

Siderits’s approach uses cognitive science as a framework to resolve the issues of Buddhism. Be that as it may, if Siderits proves successful in doing so, the idea that cognitive science and Buddhist thought have fundamental similarities will gain more credibility. Siderits, however, states that one fundamental problem with the computationalist framework is its materialism, which contradicts Buddhist concepts pertaining to morality and spirituality, such as karma, rebirth, and enlightenment. Indeed, as shown above, computationalism can explain the possibility of rebirth (of non-enduring selves), but the Buddhist model of rebirth is inextricably tied up with karma, which escapes the materialist framework. In short, though the computationalist framework may be able to explain the metaphysical side of Buddhism, it fails to account for its ethical and/or spiritual side. Siderits’s take on the matter is that though

karma, along with its related concepts, has “played an important role in many Buddhist cultures, it is not crucial to the central project of Buddhism” (Siderits 2001, 312). Thus, like Flanagan, Siderits thinks that Buddhism is essentially a materialist philosophy of mind; for while it may contain non-materialist elements, these elements are negligible.

The phenomenological approach to cognitive science, also called the *enactive approach* or *embodied-enactive approach*, is inspired by the Buddhist principle of *dependent co-arising*, which regards the mind and the world as dependent on one another, and by the phenomenology of Maurice Merleau-Ponty (Spackman 2012, 747). This approach was articulated in the book *Embodied Mind: Cognitive Science and Human Experience* by Varela, Thomson, and Rosch (1991). This approach generally defines cognition as emerging from the activity of embodied agents. This particularly means, in the words of Varela et al. (1991, 9), that “cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs” (also quoted in Spackman 2012, 747).

Spackman argues that these Buddhist-based/inspired models of consciousness have not been able to overcome the challenges faced by cognitive science. As regards Wallace’s consciousness-based model, Spackman shows that while this view may be able to sidestep the problem of the explanatory gap between the objectively physical and the subjectively mental, it, however, runs into a problem with the (scientific) *principle of causal closure*—roughly the principle that causation is sufficiently explainable in physical terms—for excluding the physical in the causal picture.

As a materialist view, Flanagan’s model is shown by Spackman to be susceptible to the standard anti-materialist arguments such as the so-called conceivability arguments—consisting, among others, of the *zombie argument* and *inverted-qualia argument*. In the case of Siderits, his conclusion is not really necessary. He falsely assumes that the computationalist model is the best model to explain the concepts and issues of Buddhism. Dualism can very well do the job so long as the substantiality of the mental is not postulated. More importantly, however, the idea that the ethical or spiritual tenets of Buddhism are not central to its project is clearly mistaken, as emphasized by the viewpoint of Karimi et al. and Federman, as we have earlier discussed. For when metaphysical questions compete with ethical ones in terms of priority or value, Buddhism insists on choosing the ethical ones. It is not that metaphysical questions are unimportant, but that these questions should aid us in dealing with the practical questions of ethics, which for Buddhism mainly concern how humans can overcome their sufferings. This point is articulated very clearly in the *parable of the poisoned arrow*, which goes:

The Buddha always told his disciples not to waste their time and energy in metaphysical speculation. Whenever he was asked a metaphysical question, he remained silent. Instead, he directed his disciples toward practical efforts. Questioned one day about the problem of the infinity of the world, the Buddha said, “Whether the world is finite or infinite, limited or unlimited, the problem of your liberation remains

the same.” Another time he said, “Suppose a man is struck by a poisoned arrow and the doctor wishes to take out the arrow immediately. Suppose the man does not want the arrow removed until he knows who shot it, his age, his parents, and why he shot it. What would happen? If he were to wait until all these questions have been answered, the man might die first.” Life is so short. It must not be spent in endless metaphysical speculation that does not bring us any closer to the truth.<sup>3</sup>

Furthermore, the fact that the Buddhist notion of the no-self can be explained by the computationalist concept of the mind as a computer software does not necessarily mean that Buddhism shares the physicalism of computationalism. As Arnold (2012, 4) notes:

Pursuing this thought, Mark Siderits asks (in the subtitle of a recent article): “Is the Eightfold Path a Program?” (2001). That is, can characteristically Buddhist accounts of the person be harmonized particularly with those cognitive-scientific projects that, informed by the availability of the computer model, take thought to be somehow “computational”? Among other things, this amounts to the question of whether the basic Buddhist commitment to selflessness might be compatible with *physicalism*.... Siderits proposes that the basic Buddhist project *is* finally reconcilable with cognitive-scientific physicalism.... There remains, however, a significant obstacle to the view that Buddhist thinkers elaborated a position that is uniquely compatible with scientific understanding: while cognitive-scientific accounts of the mind are generally physicalist in character, *Buddhist philosophers are emphatically not physicalists*.

With regard to the enactive approach, Spackman argues that it has not really been successful in resolving the problem of the explanatory gap, for whether the body is a “lived body” or not, the question of how consciousness arises from the physical processes of the body remains. The consequence of this analysis to our purposes is that these alternative Buddhist models of consciousness will not help in the project of cognitive science to naturalize the mind. In short, from the perspective of Spackman, they have nothing positive to contribute to the feasibility of this project.

Now, what about Spackman’s own alternative model, which he refers to as *neutral non-dualism*? This model is built on the Buddhist idea (of Nāgārjuna) that all phenomena are empty of *svabhāva*, which Spackman (2012, 748) translates as intrinsic nature, “that is, a nature possessed independently of other entities, causes, parts, and so on.” This implies for Spackman (2012, 748) that “No positive statements at all can be asserted about reality at the ultimate level.” Consequently, this view is “neutral” in that “the ultimate relation between mind and matter cannot be construed in terms of materialism, dualism, or any consciousness-based view” (*ibid.*, 749), while it is “non-dualistic” in that “properly, we cannot assert that reality is one, only that it is not two”

(*ibid.*, 748)—by “two,” Spackman here presumably refers to the categories of the physical and the mental. Now, whether or not this model truly overcomes the difficulties faced by the standard theories in cognitive science—the explanatory gap and others—the problem with this model is that it endorses a kind of *mysterianism* about the nature of consciousness, according to which the nature of consciousness will forever remain a mystery to us since we can never really know what constitutes it. Spackman (2012, 749) writes:

Neutral non-dualism, as I have described it, is in some ways reminiscent of the “New Mysterianism” proposed by Colin McGinn, and yet it is importantly different (McGinn 2000). Both views hold that we cannot grasp the ultimate nature of the relation between mind and matter. But for McGinn, the reason for this inability is the limitation of our cognitive powers. There is, in fact, an answer to the question of whether mental events and physical events are identical or distinct; it is just that this fact lies beyond our cognitive abilities. According to neutral non-dualism, the relation between the mental and the physical is not just unknowable by us but unintelligible.

The consequence of this mysterianism, for our own purposes, is that Spackman’s own proposed model has likewise nothing significant to contribute to the feasibility of the project to naturalize the mind. What it shows us is akin to Wallace’s proposal for cognitive science to adopt the introspective method of Buddhist meditation: that if we are to resolve the difficulties faced by cognitive science, we will have to abandon its project to naturalize the mind.

Spackman’s arguments against the first three models of consciousness are plausible, but I find his arguments in support of the model of consciousness that he advances in need of refinement. The label used by Spackman, namely “neutral non-dualism,” is confusing. If the descriptions “neutral” and “non-dualistic” are both applied to things on the fundamental level, their combination is contradictory. If things are ontologically neutral or mysterious in an unintelligible sense [not in the sense of McGinn’s (1997) brand of mysterianism which results from our cognitive limitations], then any assertion of what they are or what they are not ontologically will not make sense. So saying that things are non-dualistic on this level is as senseless as saying that they are ontologically dualistic, monistic, or pluralistic. In this regard, the combination of these descriptions only makes sense if they refer to different levels, say “neutral” to the ontological status of things on the fundamental level and “non-dualistic” to the ontological status of things on the non-fundamental level.

The non-fundamental level may refer to the higher level of existence resulting from the combinations or interactions of things on the fundamental level, such as when the basic elements of nature combine to form substances (like when hydrogen and oxygen combine in a certain way to produce water). Given this, is Spackman, therefore, claiming that Buddhism takes the nature of things on the non-fundamental levels to be non-dualistic? This is hardly the case, for Buddhism talks of physical and

mental things, processes, and activities on these non-fundamental levels. One good example would be the five aggregates that combine to form the notion of the self, which is a combination of physical and mental states and activities. The same combination can also be seen in the four noble truths and the eight-fold path.

Spackman is correct in describing the ontological nature of things on the fundamental level as neutral in consideration of the fact that they are devoid of intrinsic nature. However, he is mistaken in describing these things as non-dualistic, either in the fundamental or non-fundamental levels. On the fundamental level, describing things as non-dualistic contradicts the description that these things are ontologically neutral. On the non-fundamental levels, describing things as non-dualistic conflicts with the way Buddhism talks about things in the world.

The Buddhist philosophy of mind, in this regard (on the aspect of ontological nature), is better described as “neutral property dualism.” It is neutral in the same sense that Spackman thinks of it, but it is property dualism in the sense that on the non-fundamental levels, things have the properties of being physical or mental. While it is true that on the fundamental level, things cannot be categorized as either physical or mental, these things can have mental and physical properties on a higher level. Such properties may emerge from the interactions of things on the fundamental level.

That things are devoid of intrinsic nature does not necessarily lead to a denial of dualism in general. It certainly does lead to a rejection of the traditional kinds of dualism usually discussed in Western philosophy, namely substance dualism, and property dualism. Substance dualism, standardly represented by Cartesian dualism, claims that on the fundamental level, things are divided into material objects (matter) and minds. On the other hand, the property dualism discussed in traditional, Western philosophy of mind is of the materialist orientation. Its proponents claim that things in the world are fundamentally physical, but they have both mental and physical properties. Two prominent representative theories of property dualism of this kind are Donald Davidson’s (1991) *anomalous monism* and John Searle’s (1999) *biological naturalism*. Davidson, on the one hand, claims that the relation between mental and physical properties of physical objects is anomalous as this relation is not governed by strict deterministic laws. Searle, on the other hand, claims that consciousness is a higher-level biological property as it is a property that emerged from the interaction of the biological properties of the brain in the course of evolution.

If we separate the materialist orientation from property dualism, we will have a general theory of property dualism that can admit of kinds. In light of the considerations above, we can distinguish between two kinds: *materialist property dualism* (the one assumed in discussions in the traditional, Western philosophy of mind) and what we have referred to as *neutral property dualism*. While materialist property dualism is ontologically committed to physicalism on the fundamental level, neutral property dualism has no ontological commitment on the same level. They, however, are both committed to the view that on the non-fundamental levels, things have physical and mental properties. If our description of the Buddhist philosophy of mind in the ontological sense proves to be correct, then we have expanded the traditional, Western classifications of possible ontological positions about the nature of things and minds.

It may be asked why it is important to maintain dualism on the non-fundamental level in our account of the Buddhist philosophy of mind? First, as earlier noted, Buddhism makes ample references to the dualism of the physical and the mental states in many of its doctrines. Second, which follows from the first, it will consistently explain most, if not all, of the main tenets of Buddhism. Earlier, we noted that a materialist reading of Buddhism would not be able to explain the ethical and spiritual facets of Buddhism, as contained in its doctrines of karma, rebirth, and enlightenment. Furthermore, without the dualism, the Buddhist philosophy of mind will not be able to deal with the various arguments leveled against physicalism, such as the knowledge argument. On the other hand, an idealist reading of this philosophy will not be able to explain its adherence to the reality of causation operating both in the mental and physical realms, as presupposed in the doctrine of karma.

Finally, with its characterization as neutral property dualism, we can explain why the Buddhist philosophy of mind can accommodate some features of the computationalist theory of mind or strong AI. Siderits have shown that the Buddhist philosophy of mind can accommodate the view that minds are multiply realizable in the sense that they can function in various physical bases. In the language of strong AI, this means that minds are computer softwares that can be run by various hardwares of sufficient sophistication. In the language of Buddhism, minds are aggregates of impermanent physical and mental things and processes that can transport from one physical body to another following the principle of karma. What was wrong with Siderits's analysis, however, was that he applied the computationalist theory wholesale, with its materialist commitment, to interpret Buddhist philosophy of mind; and the result was that he was willing, as Flanagan did as well, to take out from Buddhism doctrines, such as karma, that do not fit in the materialist framework of the theory. With our characterization of Buddhism, we do not need to resort to this. We can maintain that the Buddhist philosophy of mind can accommodate some of the important views of contemporary philosophies of mind without sacrificing its spiritual or ethical aspects.

## CONCLUSION

Our investigations have shown that the insights provided by Buddhism on the nature of the mind do not help cognitive science in making its project to naturalize the mind more feasible or in overcoming the challenges that it faces in its pursuit of the said project. The Buddhist-based method for a rigorous study of the mind conflicts with the materialist commitment of cognitive science. Furthermore, the proposed Buddhist-inspired models of consciousness have their own difficulties to deal with. What is needed is an account of the Buddhist philosophy of mind that will be consistent with its central tenets. In this regard, I advance the view of neutral property dualism as a way of characterizing the ontological commitment of the Buddhist philosophy of mind. As a form of property dualism, Buddhism shares some of the features of contemporary philosophies of mind which subscribe to property dualism. Such features include the multiple realizability of minds or mental states as advanced by

strong AI or computationalism. However, as a neutral form of property dualism, Buddhism deviates from the materialist directions of these philosophies of mind. Therefore, the Buddhist turn in cognitive science will not help cognitive science overcome the challenges of its naturalization project—as this project is materialist in orientation. What Buddhism can offer to cognitive science, however, is a framework by which cognitive science can evaluate its goals and motivations. For instance, Buddhism can provide the spiritual and ethical guide to cognitive science in the pursuit of its projects, especially those involved in the construction of highly sophisticated AI or intelligent machines. It may be explored, for instance, in what ways AI can contribute to the minimization of suffering in the world.

## NOTES

1. Quoted in Wallace 1999, 176.

2. The term “mind” can generally be understood as referring to the faculty responsible for conscious states and processes, which include our cognitive and phenomenal abilities. An entity having this faculty is generally said to have a mind. The nature of the mind—pertaining to whether the mind is physical or not, identical to the brain or not, a substance or a class of phenomena or properties, causally determined by bodily and other physical events or not, a computational system or not, and others—is what is philosophically contentious. The various theories of the mind, including those of Buddhism and cognitive science, offer different accounts of the nature of the mind.

3. There are many versions of this parable. The quoted version is by the famous Zen monk from Vietnam, Thich Nhat Hanh, as cited in Zen 101.

## REFERENCES

- Arnold, Dan. 2012. *Brains, Buddhas, and believing: The problem of intentionality in classical Buddhist and cognitive-scientific philosophy of mind*. New York: Columbia Press.
- Capra, Fritjof. 1975. *The tao of physics*. Colorado: Shambhala Publications, Inc.
- Chalmers, David. 1997. *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Chalmers, David. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2(3): 200-219.
- Davidson, Donald. 1991. Mental events. In *The nature of mind*. Edited by David M. Rosenthal Oxford: Oxford University Press, 247-256.
- Dennett, Daniel C. 199. *Consciousness explained*. Boston: Little, Brown, and Co.
- Federman, Asaf. 2011. What Buddhism taught cognitive science about self, mind and brain. *Enrahonar. Quaderns de Filosofia* 47: 39-62.
- Flanagan, Owen. 2011. *The Bodhisattva's brain: Buddhism naturalized*. Cambridge: MIT Press.
- Freidenberg, Jay and Gordon Silverman. 2006. *Cognitive science: An introduction to*

- the study of mind*. California: Sage Publications, Inc.
- Garfield, Jay and Graham Priest. 2002. Nagarjuna and the limits of thought. In *Empty words: Buddhist philosophy and cross-cultural interpretation*. Edited by Jay Garfield. Oxford: Oxford University Press, 86–105.
- Hanh, Thich Nhat. 2018. Parable of the poisoned arrow. In *Zen 101* (n.d.). <http://buddhazen101.tumblr.com/post/25078328254/parable-of-the-poisoned-arrow>. Accessed 10 February 2018.
- Jackson, Frank, 1991. What Mary didn't know. In *The nature of mind*. Edited by David M. Rosenthal. Oxford: Oxford University Press, 392-394.
- Karimi, Kamran, et al. 2001. A Buddhist perspective on artificial intelligence.” In *The 38th Western Canadian Philosophical Association Conference (WCPA '2001)*, Regina, Saskatchewan, Canada. <http://myweb.ncku.edu.tw/~stli/www/teach/ke/buddhist.pdf> Accessed: February 10, 2018.
- Levine, Joseph. 1983. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64 (4): 354-361.
- Look, Brandon C. 2020. *Gottfried Wilhelm Leibniz*. *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2020/entries/leibniz/>. Accessed November 5, 2021.
- Mabaquiao, Napoleon. 2009. Transcendence and the elusive science of the mind. *Budhi: A Journal of Ideas and Culture* 13 (1 & 2 & 3): 481-487. (Special Edition, edited by Philippe Capelle-Dumont).
- Mabaquiao, Napoleon. 2015. The phenomenal concept strategy and a master argument. *Kemanusiaan* 22(1): 53-74.
- Mabaquiao, Napoleon. 2012. *Mind, science and computation*. Quezon City: Vibal Foundation, Inc.
- McGinn, Colin. 1997. Can we solve the mind-body problem? In *Philosophy of mind: A guide and anthology*. Edited by John Heil. Oxford: Oxford University Press, 781-797.
- Nagel, Thomas. 1974. What is it like to be a bat? *Philosophical Review* 83(4): 435-50.
- Payne, Richard K. 2002. Buddhism and cognitive science: Contributions to an enlarged Discourse. (Symposium Proceedings) *Pacific World: Journal of the Institute of Buddhist Studies* 3(4): 1-14.
- Penrose, Roger. 1994. *Shadows of the mind: A search for the missing science of consciousness*. Oxford: Oxford University Press.
- Pylyshyn, Zenon W. 1989. Computing in cognitive science. In *Foundations of cognitive science*. Edited by Michael Posner. Cambridge: The MIT Press, 49-91
- Santina, Peter. 1984. *Fundamentals of Buddhism*. Taipei, Taiwan: The Corporate Body of the Buddha Educational Foundation.
- Spackman, John. 2012. Contemporary philosophy of mind and Buddhist thought. *Philosophy Compass* 7(10): 741–751.
- Searle, John. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences* 3(3): 417-457.

- Searle, John. 1999. *Mind, language and society: Doing philosophy in the real world*. London: Weidenfeld and Nicolson.
- Siderits, Mark. 2001. Buddhism and techno-physicalism: Is the eightfold path a program? *Philosophy East and West* 51(3): 307-314.
- Thompson, Evan. 2007. *Mind in life: Biology, phenomenology, and the sciences of the mind*. Cambridge: Harvard University Press.
- Varela, Francisco, et al. 1991. *The embodied mind: Cognitive science and human experience*. Cambridge: The MIT Press.
- Wallace, Alan B. 1999. The Buddhist tradition of Samatha: Methods for refining and examining consciousness. *Journal of Consciousness Studies* 6 (2-3): 175-187.
- Wallace, Alan B. 2002. A science of consciousness: Buddhism (1), the modern West (0). *The Pacific World: Journal of the Institute of Buddhist Studies*, 3(4): 15-32.