AN ENCOUNTER BETWEEN ARISTOTLE AND CONTEMPORARY PHILOSOPHY OF MIND

THE CASE OF REDUCTIVE PHYSICALISM AS ESPOUSED BY JAEGWON KIM

by

EUGENE EZENWA OGUAMANAM CSSP

A thesis submitted in fulfilment of the requirements for Doctor of Philosophy

DPhil

In the Department of Philosophy at the

UNIVERSITY OF PRETORIA

FACULTY OF HUMANITIES

SUPERVISOR: PROF. EMMA RUTTKAMP-BLOEM

JUNE 2020

Table of Contents

Acknowledgement

I would like to thank my supervisor Prof. Emma Ruttkamp-Bloem for her assistance, guidance, encouragement, and unfailing patience throughout this project. I will always remain grateful. She showed me through her approach to academic works that academic excellence entails humility. Thank you very much, Prof.

I am very grateful to His Lordship, Bishop Graham Rose of Dundee diocese, for financial support. University of Pretoria (UP) Postgraduate Bursary (2017-2018) is highly appreciated too.

My gratitude also goes to the Department of Philosophy, University of Pretoria, and the library staff of University of Pretoria, St. Augustine College Johannesburg and Ekurhuleni library Benoni, and the students and staff of St. John Vianney Seminary, Pretoria.

To my mother, Felicia Oguamanam, my brothers (Chima and Damian) and their families, thank you very much for your support and encouragement. To my aunts Alice and Grace, and my uncles Aloysius and Ambrose, I sincerely thank you for being there for your sister; my mother. And may God grant eternal rest to my father, Mr. Livinus Oguamanam and all the souls of the faithful departed.

To my Spiritan family, thank you all for your support during my term as the superior of South Africa Group. Your cooperation was a great support during my studies.

I would like to thank the parish of St. Mother Teresa of Calcutta, Crystal Park for their care and love, and the Archdiocese of Johannesburg for giving the Spiritans the opportunity to serve the parish. I thank also the Carmelite Sisters in Benoni.

And to my friends, Dr. Fr. G. Walmsley, Dr. Fr. Peter Sodje, Fr. Mike Okoro, Fr. Alfred Igwebuike, Dr.Fr. Chikadi Anyanele, Ms. Chinwe Abasilim, Ms. Miriam Moepya, Ms. Bonke Nkosi, Ms. Gladys Letshedi, etc. for all the words of encouragement. I say to you all, thank you.

May God bless you all.

Abstract

I argue in this thesis that Aristotle's hylomorphic metaphysics, supported by his theory of causality and his theory of the soul (*De Anima),* holds the key to solving the problem of mental causation in contemporary philosophy of mind. A core aspect of the contemporary mind-body problem is the problem of mental causation (how does the mind interact with the body to cause actions in humans). Without mental causation, in the realist sense of the word, it is difficult to see how humans are held responsible for their actions. There have been different approaches to solving the mind-body problem, but each has met with its own set of problems, except, I argue, Aristotle's hylomorphism. Jaegwon Kim argues that Davidson's anomalous monism *cum* supervenience renders mental causation epiphenomenal, and that a mental state is causally efficacious only when reduced to the physical properties. I argue that it is the phenomenal consciousness that accounts for our actions, and while neither Davidson's nor Kim's accounts of action can adequately deal with phenomenal consciousness, Aristotle's metaphysics can. I argue that the ancient and neo-Aristotelian notion of self-knowledge is akin to our contemporary notion of phenomenal consciousness and that Aristotle saves the notion of autonomous mental causation through his theory of hylomorphism that holds every substance is a composite of matter (body) and form (soul). My thesis is thus a novel invitation to rethink Aristotle's psychology and philosophy of mind in the context of contemporary philosophy of mind.

Key Words

Mind-Body Problem, Reductive and Non-Reductive Physicalism, Functionalism, Anomalous Monism, Supervenience, Causal Closure Argument, Exclusion Argument, Hylomorphism, Phenomenal consciousness.

Introduction

The so-called 'mind-body problem' is a centuries-old problem in metaphysics. It raises fundamental questions concerning the nature, dignity and worth of a human person and has one of its first Western philosophical formulations in Aristotle.

Humans, just like any other biological animal, have a body that is spatio-temporal, in other words, physical. On the other hand, what differentiates humans from most other animals is the quality of our capacity to think, reason and plan, as well as the ability to reflect on our actions and plans in order to improve on them. Eric Matthews (2005:1) writes:

> We are certainly animals, primates like apes and monkeys but we tend to distinguish ourselves from other primates by our capacity to think things through, to reflect on our existence and to change our actions in the light of that reflection, to plan and organize our lives, to control our emotions and desires – in short to be rational.

Thus, human beings are beings with both physical and mental properties. One of the biggest issues concerning mental properties relates to the mind-body question. The mind, with its mental properties such as self-awareness and consciousness, seems to be located in the brain, which is a physical organ in the body. The mind, in this sense, is thus embodied. On the other hand, mental properties of mind, such as subjective experience and intentionality, seem to suggest that the mind is a non-spatial, immaterial substance. This characteristic of the mind, together with its embodiedness, creates a tension, which underlies all mind-body debates (McGinn 1982:16-17).

The mind-body problem thus relates to the relationship between the immaterial mind and the physical body. But, how does what is mental interact with physical substance? Some issues to consider in answering this question include for instance: How can we characterize a causal connection between mental events (e.g. my desire for a drink) and physical events (e.g. my getting up from my seat and walking to the fridge)? Furthermore, it is a known fact that damage to the brain can also affect the rational capacity and even the emotions of a human

person. But how can this be? The mind-body problem, in summary, is the problem of mental causation (the problem of interaction) and the problem of the nature of the mind (its metaphysics). In this thesis, the focus will be more on the former than the latter problem – i.e. how can something strictly mental have an effect on a material or physical substance? By the strictly mental phenomena, I mean consciousness – particularly phenomenal consciousness, which is the subjective feeling one has in being in a particular mental state. For instance, the sweetness that one experiences in tasting the sweetness of an orange. It is the sweetness of an orange that will cause me to love oranges. The other aspect of consciousness is access consciousness which refers to the causal relation between mental states in producing particular behaviour (see Gennaro 2017, 4).

The mind-body problem debate is about the nature of a human person; it is a debate about what it means to be a rational and conscious human person. Thus, any theory of mind should be able to account for rationality and consciousness in a human person.

In contemporary times, there are mainly two umbrella approaches to the mind-body problem, namely dualism and physicalism. Currently, the more influential of the two seems to be physicalism (Kind, 2018). Dualism holds that the mind is immaterial while the body is physical and the mind is independent of the body. Physicalism comes in two broad forms: reductive physicalism and non-reductive physicalism. Reductive physicalism is the view that every mental state or property is identical or reducible to a physical state or property. It holds that mental states do not exist over or above physical processes. Non-reductive physicalism holds that mental states are physical states but denies that mental states can be reduced to behavior or brain states. Instead, it acknowledges some kind of causal relation between the two kinds of states.

Jaegwon Kim, a great proponent of reductive physicalism, argues that if mental states are to have causal power over physical states, then mental states must be reduced to or identified with physical states. Hence, non-reductivism is not possible for Kim. Kim (2001a), in his book, *Mind in a Physical World: An Essay on the Mind- Body Problem and Mental Causation*, writes that the dominant view or theory of the mind-body problem for the past twenty-five years has

been non-reductive physicalism (the view that every substance is physical although mental states are irreducible to physical states). In other words, mental states are autonomous on this view.

Kim (ibid.) explains that functionalism and Donald Davidson's anomalous monism are two contemporary theories of mind that have argued for the autonomy of mental. The main argument of functionalism is that mental kinds or states are functional states of their physical or biological realizers. Mental kinds and properties are functional kinds at a higher level of abstraction than physiochemical or biological kinds (ibid.). Functionalism is committed to multiple realizability (i.e., the argument that a mental state is realizable by different physical states) (Putnam, 2002). Davidson's anomalous monism holds on the one hand, that all mental events are physical events (monism) and on the other hand, that mental events are anomalous because there are no laws connecting mental events and physical events (no psychophysical laws) (Davidson, 2001). For Davidson, therefore, mental events are both autonomous and irreducible to physical events.

Davidson (2001) develops anomalous monism in an attempt to allow for mental causation – i.e. in support of the notion that mental events have causal power. Davidson (ibid.) takes mental events to be physical events implying monism, which means that, in effect, all events are physical events. Hence, mental events are in a causal relation with physical events. This causal relation does however not hold between mental event *types* and physical event *types* as they are in an anomalous relation in the sense that the usual deterministic or strict laws of physics do not connect them. Rather it holds between mental event *tokens* and physical event *tokens* only. For instance, a particular pain (*M*) at t_1 is identical with a particular C. fibers (*P*) firing at t_1. Thus, "the distinction between a *type* and its *tokens* is an ontological one between a general sort of thing and its particular concrete instances" (Wetzel, 2018).

Kim (2001a) criticizes Davidson's anomalous monism because he argues that it does not in fact adequately explain how mental and physical states are related. Kim holds that every mind-body theory should be able to explain how the mind is related to the body. In response to Kim's criticism, we may consider Davidson's supervenience theory. This is the idea that the

mind supervenes or is dependent on the physical, which implies that if two events are the same in all physical respects, they must be the same in all mental respects and there cannot be a change in the physical without a change in the mental. Kim argues that the supervenience theory is not a mind-body theory because it does not actually account for *how* the mind supervenes on the body. He also accuses functionalism of violating his principle of causal closure (physical events can only have physical causes).

In this thesis, I will argue that considering Aristotle's philosophy of mind in relation to Kim's attack on non-reductive physicalism offers a very interesting perspective on the current impasse in contemporary philosophy of mind between reductionist and non-reductionist approaches to the mind-body problem. My overall claim is that a philosophy of mind based on a reductive physicalist framework, understates the reality of the mental to an unacceptable extent. My research aim is to argue for a non-reductive physicalist view of the mental in terms of Aristotle's formulation of the mind-body problem[1] (based on his hylomorphic analysis of individual substances) to critique Jaegwon Kim's reductionist position. I will argue that Aristotle's hylomorphic theory is what non-reductive physicalism needs to offer an adequate solution to the mind-body problem, while upholding mental causation. This does not imply that my solution will not be based on any neuro-scientific evidence, as it is still a physicalist position after all, albeit a non-reductive one.

In his *De Anima (On the Soul)*, Aristotle (1986) applies his metaphysical theory of matter and form (hylomorphism), to explain the nature of life in living things. Matter is the physical body of a living thing, while the form of a living thing is its soul, and the notion of 'form' relates to whatever determines the way a living thing functions. The theory of the soul as the form of a living thing is central to Aristotle's psychology in *De Anima*. For Aristotle every good account of the soul must account for two things namely 1. How living bodies generate physical processes or spontaneous movements; and, 2. How living bodies could have cognition and mental awareness of their surroundings, which entails knowing how physical and mental processes could form a unit in living bodies. Aristotle believed that his predecessors (and we must add,

[1] Though, Aristotle did not have a notion of mind in the contemporary sense.

most of his successors) were unable to account for the unity of physical and mental processes. The work in *De Anima* does not have a modern concept of the soul and it does not discuss the soul's immortality directly, rather a general conception of the soul and its many activities such as perception, imagination, desire and intellect, are discussed – all of which are of great interest to modern philosophers.

I believe Aristotle, as an ancient philosopher, still has this interesting contribution to make to contemporary physicalist-neuroscience-based philosophy of mind because contemporary non-reductive philosophy of mind, which is deeply functionalistic in nature (emphasizing only a causal relation between mental events and physical events in the brain), in arguing for the autonomy of mind and mental causation. I am motivated to use Aristotle to salvage non-reductive physicalism from Kim's criticism rather than exclusively using contemporary physicalist-neuroscience-based philosophy of mind, because neuroscience solutions to mind-body problem are reductive in nature. Reductive physicalism, as I will argue and show in my thesis, does not offer adequate and sufficient solutions to the mind-body problem because it denies an independent mental causation.

Thus, by bringing Kim and Aristotle – contemporary and ancient; and reductionist and non-reductionist – into debate with each other, I will show that Aristotle's hylomorphic theory of individual substances offers a novel way to salvage non-reductive physicalism from Kim's contemporary criticism.

My main purpose with this research is to argue that an Aristotelian critique of Jaegwon Kim's philosophy of mind provides a novel response to the contemporary debate on the mind-body problem. My overarching aim is an invitation to rethink Aristotle's theory of mind as I argue it is still relevant in this modern time and offers a viable challenge to contemporary reductionist philosophy of mind. Using Aristotle's hylomorphic theory, I will show that reductive, exclusively neuro-science-based solutions to mind-body problem cannot adequately show or explain what it means to be a rational and conscious human being. I will show that Aristotle's biological functionalism (the idea that something is alive by virtue of the function of the form (soul) within its biological constitution) that is based on his hylomorphic metaphysics, has a

serious potential to change contemporary thinking about the mind-body problem, and I will show this in this work in particular in relation to his psychology.

My work in this thesis is to invite contemporary philosophers to look back beyond Descartes to find a possible solution to the mind-body problem. The uniqueness of my project lies in the fact that I am using Aristotle's psychology to challenge the psychology of a well renowned modern day philosopher, Jaegwon Kim and that I suggest a perspective on the mind-body problem that is not exclusively based on neuro-scientific data, although engaging with it in arguing against reductivism, and still maintaining a *physicalist* perspective, albeit a non-reductive one.

To achieve my objective in this work, I will commence in chapter one, by defining the key terms such as mental reality, mental causation, the immaterial nature of mind, and phenomenal consciousness whose causal efficacy is the epicenter of this thesis. I will state that the kind of mental causation under discussion is not the Cartesian dualist understanding of mental causation, but rather Kim's concept of mental causation as I want to save non-reductive physicalism from Kim's criticism that it renders mental phenomena epiphenomenal (causally inert).

In chapter two, I will explain what physicalism is. Furthermore, will discuss the functionalist notions of mental causation as both Kim and Davidson are functionalists. However, while Kim is a realist functionalist, Davidson is a role functionalist. I will show that physicalism, which comes in two broad forms namely reductive physicalism and non-reductive physicalism, and functionalism, both have truncated ideas of mental causation, and hence cannot account for phenomenal consciousness or qualia required for an adequate mental causation. I will show that their inadequate solution to mental causation is the cause of their failure to offer a theory of action that is human oriented, rather than mental states oriented.

I will examine in chapter three, the debate between Kim and Davidson on mental causation, which I call the modern discussion on mental causation. Davidson argues for autonomous mental causation based on his theory of causality and anomalous monism *cum* the supervenience principle. But Kim using his principles of mental causation argues that Davidson's

theory of mental causation renders mental phenomenal causally inert as the causal power of the mental phenomenal reside in the causal power of physical phenomena. However, as Kim is interested in saving mental causation, he proposes a functional model of reduction that holds mental phenomena are causally efficacious only when they are functionally reduced to physical phenomena. I will argue that Davidson does not save mental causation as his theory is physically based just like Kim's theory. In other words, he does not succeed in arguing for an autonomous mental causation. And neither Kim nor Davidson can account for phenomenal consciousness with their respective theories, even if Davidson makes a compelling argument in relation to the four principles of anomalous monism. Hence, we have to look elsewhere for a solution to mind-body problem.

I will argue in chapter four that it is Aristotle's hylomorphism and related notion of causality that can offer an argument for mental causation, and also, I will argue that his theory can account for phenomenal consciousness being causally efficacious. By borrowing a bit from medieval philosophers and neo-Aristotelians such as Thomas Aquinas, I will argue that he accounts for phenomenal consciousness as self-knowledge, which I will argue is the modern notion of self-consciousness. Hence, Aristotle is able to give a theory of action that is human oriented, instead of mentally states based, as in the case of Davidson.

I will argue in chapter five that Aristotle's theory of hylomorphism and theory of causality seem to offer an adequate response to Kim's criticism of non-reductivism that is enshrined in his principles of causal closure and causal exclusion. I will also show that Aristotle offers a better response to Kim than Davidson does.

I will conclude the thesis by showing that my thesis is a novel contribution to the mind-body problem. And if my thesis is acceptable, the contemporary mind-body problem will have seemingly been solved.

Chapter 1: Reality of the Mental and Mental Causation

1. Introduction

In this chapter, I will set out the playing ground for this thesis by explaining concepts such as the reality of the mental, mental causation, and the immateriality of the mental. I take consciousness to refer to all mental states, and these mental states fall under either 'access consciousness' or 'phenomenal consciousness' or to use David Chalmers's coinage, the 'easy problem' or 'hard problem' of consciousness. The 'easy problem' refers to functionally definable phenomena such as response to environmental stimuli, while the hard problem is the problem of subjective experience or phenomenal consciousness - "[w]hy should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does" (Chalmers 1995: 3). In this thesis, when I argue that a mental state is real and causally efficacious in the physical world, I mean to refer to mental states of the phenomenal consciousness; the qualitative aspect of consciousness which is normally described with Thomas Nagel (1974) expression "what is it like' (to see the redness of a red flower). I argue that phenomenal consciousness, though immaterial – in the sense of not being physical or physiological – is real and causally efficacious in the physical world. Furthermore, I will explain in this chapter that the notion of mental causation under discussion is Jaegwon Kim's notion of mental causation, and not the Cartesian dualist notion. It is Kim's notion that is under discussion as my aim is to save mental causation from Kim's claim that mental properties are inefficacious in the physical world unless they are functionally reduced to physical properties.

2. The Reality of the Mental/Mental Realism

Mental Realism is the theory that mental properties and events are causally active in the physical world. They are not inert; they cause behaviours and actions as mental properties, not as reducible physical properties (Miller, 2019). In this sense of the word, in any production of behaviour or action, "mental properties are real properties of objects and events; they are not merely useful aids in making predictions or fictitious manners of speech" (Kim 1993, 344). For Kim, to be a mental realist is to believe that "your mental properties must be *causal properties* – properties in virtue of which an event enters into causal relations it would not have otherwise have entered into" (Kim 1993, 279). Thus, in the causal structure of the world, mental properties are found doing causal work *qua* mental, not as reducible physical properties. There are mentalistic explanations of our actions and behaviour because mental properties are causally active in their production. For instance, I opened the door because I desired that it should be opened. In this case, desire is the mental event that caused me to open the door. Because mental properties have causal powers, they exist (Alexander dictum).

The reality of the mental implies that the mental does not need to be reduced to physical properties for them to cause mental events and physical events. Thus, the reality of the mental opposes reductive physicalism's view that holds that mental properties are only causally efficacious in so far as they are reducible to physical properties and that on their own, *qua* mental properties, they are epiphenomenal (causally inactive). Epiphenomenalism is contrary to our common-sense expectation because we believe that there is mental causation; "[e]piphenomenalism is against our experience of mental events for example, the way we experience pain seems to be a good reason to avoid pain. It feels bad. However, if epiphenomenalism is true, then we don't actually try to avoid pain because of how it feels. We would merely avoid pain (when we do) because of our non-mental brain activity" (Gray 2010: 1).

Though mental realism holds that mental states, properties, and events (together referred to here as mental phenomena) are real and causally active, supporters of the view do

not necessarily claim that mental phenomena are separable from the physical as professed by Cartesian dualism. Ricardo Echavarria (2010: 9) writes:

> Mental realism asserts the existence of mental reality, mental properties, mental events, and mental substances. It may come as a surprise that mental realism asserts the existence of mental properties without having Cartesian dualism as a consequence. However, this surprise is quelled when we consider that a mental substance just is a substance with mental properties, in our case, selves or minds …

Tamas Demeter (2009: 65) writes about the reality of the mental:

> Mental entities exist. What they are or what their nature is, how they work, etc. is not a matter of being realist about them, not a matter of ontological commitment, but of an empirical hypothesis or speculative metaphysics, depending on one's cognitive tests, which is connected to realism only in a contingent way. Thus, the thesis can be expressed more informatively as mental entities exist independently of what we think about them.

The claim defended in this thesis is that mental phenomena do not need physical phenomena to cause events in the physical world, and thus they do not necessarily in any way share the same properties with physical phenomena. They are not of the same nature as the physical. Since physical phenomena are known to be extended in time and space and quantifiable and measurable (Robinson, 2017: 6), mental properties are non-extended and immaterial or nonphysical (this description is Cartesian, although I am not professing substance dualism here). The further claim of the thesis is that although mental substances are immaterial, they have causal power in the physical world. Hence, they are real. Thus the notion of mental causation in hand here is of a kind of causation that affirms that mental states are real.

3. Mental Causation

Mental causation is the question of the possibility of an immaterial mind, interacting with the body, or of mental properties affecting physical properties. How could mental, *qua* mental, cause the physical body to act? To ask how mind and body are related is to ask how they could possibly affect one another (Robb et al., 2018: 1-2). Another instance of mental causation is when the pain I feel when I break my ankle causes me to enter my car and drive to see a doctor for treatment. More so, the mental images we have in our memory help us to identify things around us. Or, my desire to be a doctor causes me to register for medical studies at a university. We are moral agents because mental properties, such as reason[2], can help influence our action and behaviour. And we are held responsible for our actions because our mental activities such as deliberations and decisions do cause our behaviour (ibid.).

Kim (2001a: 57) notes too that without the causal efficacy of mental states, human agency is an illusion. We are responsible for our actions because our desires, beliefs, intentions, decisions, and hope cause us to move our physical bodies to perform actions in the physical world.

> Agency of the sort required for freewill and moral responsibility appears to require mental causation. If your behaviour is not caused by your mind's activities-its deliberations, decisions, and the like – what sense would it make to hold you responsible for what your body does? You would appear to be scarcely more than a passive observer of your body's activities. (ibid.: 2)

Perception, memory and reasoning are possible due to mental causation. Moreover, without them, there will be no human knowledge. In reasoning, mental states cause one another. Acquisition of new beliefs or knowledge may cause me to change my previously held beliefs and knowledge, which, without mental causation, will not be possible (Kim, 2001a: 57).

[2] I take reason to be the cause, explanation or justification for an action or event.

> In the philosophy of psychology, it is widely believed that psychological explanation hinges on the possibility of mental causation. If your mind and its states, such as your beliefs and desires, were causally isolated from your bodily behaviour, then what goes on in your mind could not explain what you do.
> (Robb, et.al. 2018: 2)

The term 'mental causation' differentiates involuntary actions from voluntary actions. Voluntary actions are the actions we perform intentionally and freely and we can be held morally responsible for them (Yoo, 2018: 1). Mental causation is fundamentally central to our understanding of voluntary actions as mental states such as intentions or desires are the direct causes of our actions; but mental causation is not applicable to involuntary actions. Involuntary actions such as blinking of eyelids and yawning are not subject to mental causation, and hence we are not held responsible for them (ibid.). Thus, we are agents of our action because our mental activities cause our bodily and other mental activities.

There are other ways in which the mental and the physical can interact with one another or exert causal influence on one another (Maslin, 2007: 175):

(1) The thought of going on holiday made him smile (mental to physical).

(2) The desire to go on holiday made him start thinking of how to raise money for the holiday (mental to mental).

(3) She struck her toes on a piece of stone and felt pain (physical to mental).

I will note here that there is a distinction between the dualist notion of mental causation and Jaegwon Kim's notion of mental causation. I will be working extensively with Kim's notion of mental causation, and have focused on explaining that interpretation of the concept in this section, as I intend to use Aristotle theory of hylomorphism and four causes to 'save' mental causation from Kim's criticism that mental properties *qua* mental are causally

inefficacious in the physical world unless they are reduced to physical properties. Let me now – briefly – explain the dualist notion of mental causation.

3.1. The Dualist Notion of Mental Causation

Kim (2001a:57) in his book, *Mind in a physical world*: *An Essay on the Mind-Body Problem and Mental Causation,* traces the history of mental causation. Kim (ibid.) claims that Descartes is responsible for both of the problems of mental causation and the mind-body problem, because he propounded the ontology of two radically separate types of substances, namely material bodies that are spatial and temporal and conscious minds that are immaterial and non-spatial. With the mind being different from the body; mind without any physical characteristics, it becomes a problem to understand how the mind is supposed to interact with the body or cause the body to action. Descartes' contemporaries such as Pierre Gassendi and Princess Elizabeth of Bohemia (see e.g. Skirry, 2008: 135) raised objections to the possibility of interaction between the mind and the body. Gassendi holds that the mind cannot interact with the body because the mind is immaterial while the body is material:

> How can there be effort directed against anything, or motions set up in it, unless there is mutual contact between what moves and what is moved? And how can there be contact without a body when, as is transparently clear by the natural light, naught apart from body, can touch or yet to be touched? (AT V11 341: CSM 11 237)

In her letter to Descartes, Princess Elizabeth noted that contact is required between the mind and the body for causal interaction to occur. Thus, she asked Descartes for an explanation of his position that the mind and the body interact:

> I beseech you to tell me how the mind of man (being only a thinking substance) can determine the spirits of the body in order to make voluntary actions. For it seems that every determination of movement is made by the impulsion of the

> thing moved by the manner in which it is pushed by what moves it ....(AT 111 661)

Gassendi and Princess Elizabeth's reading of Descartes' theory of dualism implies that a human being is not one substance but two completely separate and independent substances that can only causally interact by contact and motion. For Descartes, however, the mind and the body as independent substances are not joined together as two things that causally interact by contact and motion only, but somehow they are one, a whole forming a human being – this is the aspect of the mind-body problem related to the embodiment of the mind. Descartes uses the sailor in a ship metaphor to illustrate the interaction and union between the mind and body. He argues that the mind and the body are not united in the way a sailor is present in her ship. Descartes writes:

> Nature also teaches me, by these sensations of pain, hunger, thirst and so on, that I am not merely present in my body as a ailor is present in a ship, but I am very closely joined and, as it were, intermingled with it, so that I and the body form one thing *(unum quid*). If this were not so, I, who am nothing but a thinking thing would not feel pain when the body was hurt, but would perceive the damage by the pure intellect (*pure intellectu*), just as a sailor perceives by sight if anything in his ship is broken. Similarly, when the body needed food or drink, I should have an explicit understanding of the fact, instead of having confused sensations or hunger and thirst. For these sensations of hunger, thirst, pain and so on are nothing but confused modes of thinking, which rise from the union, and, as it were, intermingling of the mind with the body. (AT V11 81: CSM 11 56)

So, mind and body are distinct from one another even though they are composite. And Descartes insists on a mind-body composite because its purpose is to sustain and preserve

one's life as a human being. And this explains why we have sensory perceptions as well as passions such as love, hatred, desire, joy and sadness:

> Regarding this, it must be observed that they are all ordained by nature to relate to the body, and to belong to the soul only in so far as it is joined with the body. Hence, their natural function is to move the soul to consent and contribute to the actions, which may serve to preserve the body. (AT X1 429-30: CSM 1 376)

The mind-body composite is best understood through Descartes' theory of passion. In that theory, Descartes states that 'the soul is united to all parts of the body conjointly', which is just a simpler way of stating that the whole soul is in the whole body. Though the whole soul is in the whole body, according to Descartes, the soul performs its functions more particularly in the pineal gland located at the centre of the brain (AT X1 351-3: CSM 1 339-40). The mind is the human soul for Descartes and as such, it is the principle of rational, human life. The mind animates a human body to make it a living human body in order for the entire body to be truly alive. Descartes does not claim that the mind is entirely in the pineal gland to the exclusion of any other body part, but only claims that this is where the mind performs its proper functions most particularly. For Descartes:

> … sensory perceptions and passions are received from the body through the pineal gland, as are the actions of the will or volitions transmitted to the body through it. Therefore, the whole soul is in the whole body and the whole in any one of its parts by virtue of its functions as a soul or animating principle, but it has its principal seat in the pineal gland in so far as the mind exercises its primary powers of perception and volition through it. (ibid.: 149)

Thus, it is one small part of the brain (*the pineal gland*) that directly mediates causation from body to mind and mind to body. The mind is not immediately affected (*immediate affici*)

by all parts of the body, but only by the pineal gland (Voss, 1993: 131). Thus, "Descartes proposed that we could locate the workings of mental causation in the pineal gland, which he believed to be the gateway between the mind and the body" (Yoo, 2017: 8). Dale Jacquette (2009) notes that the reason Descartes opted for the pineal gland as the centre of causal interaction could be the fact that the pineal gland, unlike other structures in the brain, is not divided into hemispheres, hence serves as a single point of contact between mind and brain. The question remains however how mental causation can occur in the pineal gland that is spatially located in the brain while the mind lacks spatial dimension. Thus, Descartes does not solve the problem of mental causation. Mel Thompson (2012:20)) writes:

> How an un-extended mind could cause even the slightest movement in the physical body remained a theoretical problem, but Descartes tucked the problem away in the most inaccessible place possible. In doing so, he dodged the fundamental issue that his dualism had raised.

Descartes' theory of mental causation (interactionism) has been criticised based on two factors, namely spatial location (as mentioned) and energy conservation. Firstly, to reiterate in terms of spatial location, it is the nature of causation that the cause and its effect must share the same spatial location or stand in contiguous space to each other: a ball does not move unless something pushes against it. Thus, based on spatial location, there is no way mind and body can interact or cause each other because thought or feelings have no spatial location like the body (Gennaro, 2017: 9).

The other problematic factor of Descartes' notion of mental causation is the principle of conservation of energy. Causation occurs when energy is transferred from one thing to another, for instance, when a moving car hits a tree and the tree falls down, the car's momentum has been transferred to the tree. It is a well-known fundamental physical law that the energy in the universe is always conserved; there is no loss or gain of energy. The total quantity of energy (kinetic, chemical, electrical, thermal, etc.) in the universe is always fixed. This is called the conservation of energy principle, and it entails that the universe is physically causally closed.

There is no energy outside the universe that can cause something to happen in the physical universe and no energy in the physical universe can leave the physical universe and cause something outside it. Since the mind is not extended, non-spatially located as Cartesian dualism holds, it is outside the physical universe and thus it cannot cause or affect the body that is extended and spatially located in the physical universe. For the mind to causally effect the body, will be a violation of the conservation of energy principle. Hence, there is no mental causation (Gennaro, ibid.).

Some philosophers, like Daisie Radner (1993: 122), go so far as to argue that Descartes' theory of causation is actually against the very notion of mental causation. For Descartes, there must be a likeness between an effect and its cause. Descartes writes in the *Third Meditation*, "Now it is manifest by the natural light that there must be at least as much (reality) in the efficient and total cause as in the effect of that cause" (AT VII, CSM II, 28). Descartes went on to explain his view: "For where, I ask, could the effect get its reality from, if not from the cause? And how could the cause give it to the effect unless it possessed it?" (AT VII, 40; CSM 11, 28). For Radner (1993: 123), Descartes' causal principle implies that a cause must possess the same sort of modification that it brings about in its effect. In other words, a cause produces what it has in an effect. For instance, the heat from fire will cause water to boil. So causal interaction between mind and body would be ruled out, since they are quite opposite to each other and do not possess the same modifications.

In the contemporary period, the dualist notion of mental causation no longer attracts much attention. In contemporary philosophy of mind, it seems most often to have been replaced by Kim's notion of mental causation, embedded in some version of property dualism (The idea that there is one physical substance that possesses two kinds of property, namely mental properties and physical properties.

3.2. Kim's Notion of Mental Causation

In his 1993 book, Kim (ibid) treats the topic of mental causation, which he considers a central question in the philosophy of mind, in detail. There are three principles that play a very important role in Kim's argument for mental causation namely

(1) The criterion of reality, which states that to be real is to have causal power.
(2) The causal exclusion principle, which states that for each event there must be one complete and independent causal explanation.
(3) The causal closure principle, which states that the physical universe is causally closed, that is, every physical event has a complete causal explanation in terms of another physical event.

For Kim, only physical events and properties have causal power. Mental events have causal power only when they are reduced to, or identified with the physical. Specifically, for Kim, the mental supervenes on the physical. Similarly, the causal power on the higher level must supervene on the causal power at the lower level. The supervening causal power is dependent on that on the lower level. Kim (1993) holds that all the events in the world can be causally explained in terms of microphysics and by the causal principle there can only be one causal explanation of each event. All causal powers need to be reduced to those of microphysics. He states that there is no irreducible causal power, which then becomes his fourth principle. As a fourth principle, there is the claim that the causal power of any object, event or property is reducible to the lower level causal power unless of course there is no lower level.

Kim holds that mental properties *qua* mental properties do not have causal power in the physical world. In the physically closed universe, it is only physical properties that are causally efficacious in the physically closed universe. For Kim, mental properties can be causally efficacious only if they are reducible to physical properties and mental properties are reducible to physical properties through functional reduction (and they are dependent on physical properties, through supervenience). Thus, only through functional reduction, can intentional and cognitive properties become causally efficacious.

Kim (2002: 643) writes:

> But are mental properties physically reducible? I have argued that if they are to be causally efficacious, whether with respect

> to physical properties or other mental properties, they must be reducible to physical properties. But to 'solve' the problem of mental causation, another step must be taken: we need to show that mental properties are in fact physically reducible. The position I favour here is similar to the position recently defended by several philosophers: (i) cognitive/ intentional mental properties, including belief, desire, perception, and the like are physically reducible (via functional reduction); however, (ii) sensory qualities of conscious experience ('qualia') are not so reducible.

Kim argues against epiphenomenalism of the mental by the functionalization of mental properties so that mental properties, functionally reduced to physical properties, can enjoy the causal efficacy of physical properties in a physically closed universe.

It is in the context of Kim's notion of mental causation that I will argue, against Kim, via an interpretation of Aristotle's philosophy of mind, that mental properties *qua* mental are causally efficacious. They do not need to be reduced to physical properties for them to cause mental and physical events. The whole argument will be expanded in later chapters. In this thesis, I will thus treat mental causation as the causal efficacy of immaterial consciousness. The reason for this is because consciousness embodies all other mental states as I will show in the next section.

4. Mental Causation as the Causal Efficacy of Immaterial Consciousness

I argue that mental states have causal power in the physical world. They do cause mental events and physical events. Mental states are sensations, cognitions, emotions, perceptions, and conative states, as well as subjective experiences (the feelings what it is like) we experience as human beings. They can be characterised as follows:

> Sensations – pains, aches, tickles, itches, throbs, tingles

> Cognitions – believing, knowing, understanding, conceiving, thinking, reasoning
>
> Emotions – fear, jealousy, envy, anger, grief, indignation, enjoyment
>
> Perceptions – seeing, hearing, tasting, smelling, touching,
>
> Quasi-perceptual states – dreaming, imagining, seeing in the mind's eyes, hallucinating, seeing after images
>
> Conative states – acting, trying, wanting, intending, wishing
>
> Qualia – experiential states. (Maslin, 2007: 6-7)

I take all these mental states to be immaterial, but not in the sense of Cartesian dualism that separates mental states completely from physical states. Though they are immaterial, they operate through the physical bodies. They are thus embodied as in the Cartesian case, but not separate substances, they are composite not separable. Through mental states human beings interact with their environment and live accordingly in the world. More so, mental states are embodied in consciousness. In other words, consciousness entails them. Kim (2007 : 407) claims that:

> Concerns about the efficacy of consciousness can arise either as part of a broad concern about the efficacy of mentality in general, or as a more specific worry focusing on conscious mental states, or the conscious aspects of mental states. It might be that although mental states, including those that are conscious are causally efficacious, there is a further question of whether the fact that these states are conscious is causally relevant. Thus, consider a conscious thought or belief. Assume it is causally efficacious in generating further thoughts and beliefs and in the production of appropriate bodily behaviours. Even so,

> its being a conscious thought might be causally irrelevant; it might be what makes a causal difference is only its content. The conceptual distinction that have led to the separation of the two issues {mentality in general and consciousness} are relatively new, and the philosophical writings on the issue of causal efficacy of consciousness up to the nineteenth century appear to have addressed the issue in terms of mentality in general, although it is quite clear that the consciousness was the focus of attention. (2007: 406)

Thus, consciousness would include feelings, thought, memories, feelings, sensations, desires, beliefs and intentions.

The notion of consciousness that embodies all other mental states is immaterial and ontologically independent of physical states. Colin McGinn (1999: 12-13) discusses the immateriality of consciousness thus:

> Consciousness itself could not be introduced simply on the basis of what we observe about the brain and its physical effects. If our data, arrived at by perception of the brain, do not include anything that brings in conscious states, then the theoretical properties we need to explain these data will not include conscious states either. Inference to the best explanation of purely physical data will never take us outside the realm of the physical, forcing us to introduce the concepts of consciousness.

I argue that this notion of immaterial consciousness depicts immaterial consciousness as a real phenomenon in the physical world because it causes mental and physical events. Let us look more closely now at the notion of subjective or phenomenal experience to grasp more reasons for the claim that consciousness is immaterial.

4.1. Qualia and Subjectivity of Mental Phenomena

Thomas Nagel (2002) in his article, "*What Is It Like to be Bat"*, notes that consciousness may occur in various forms and degrees in different kinds of organism. However, for an organism to have basic conscious experience there must be something 'what it is like' for the organism to be itself and not another organism. He (ibid.: 219) calls this the subjective character of experience. Another term for this subjective experience is *quale*. Qualia[3] are the ways things seem or look or taste or feel or smell or sound to us. Thus, qualia are the phenomenal qualities of experience (Dennett, 2002: 226). They are private to any particular organism. For Nagel, there is no reductive theory of mind capable of explaining this kind of experience, because it is experience that is subjective in character. It is beyond physicalism to explain, because physicalism deals with objective data or external objects.

The qualitative character of consciousness is equated with 'raw feels' such as the redness someone experiences when she looks at ripe tomatoes, or a particular taste she experiences when she eats a ripe orange. It is the scent of roses she smells, or the pain or lust she feels, or sounds of music she hears. Thus, as mentioned already, the colour, taste, smells, sounds, pain and lust as she experiences or perceives them are the phenomenal features of conscious states called 'qualia' (Van Gulick, 2017: 9). Thus, qualia entail our experienced sensory states, thoughts and desires.

Although the terms 'phenomenal properties' and 'qualia' are used interchangeably, the two should not be co-mingled together (ibid.: 10). Phenomenal structure entails all the structures available within the domain of experience while qualia are the properties of a phenomenal or experienced object. The phenomenal structure of experience is intentional and entails sensory properties (colour, shape, feels, and tones), qualities, representations of time, space, cause, body, self and the lived reality of the world. In other words, "the phenomenal structure of consciousness encompasses both sensory qualia and the spatial, temporal and conceptual organization of our experience of the world and ourselves as agents in it" (ibid.).

[3] Clarence Irving Lewis coined the term 'qualia' (Crane, 2000: 169-194).

Subjectivity (subjective experience), though "sometimes equated with the qualitative or the phenomenal aspects of consciousness … is a distinct aspect of consciousness, nevertheless, related to the qualitative or phenomenal aspects of consciousness" (ibid.: 11). However, it is different from these aspects of consciousness, because it is about the limit on the knowability or understandability of facts of conscious experience. Subjectivity is the view that the "facts about conscious experience can be at best incompletely understood from an outside third person point of view, such as those associated with objective physical science" (ibid.). Subjective experience is not subject to scientific observation and examination as it is not extended in space.

Nagel (2002: 220) offers a thought experiment to explain the subjective character of experience (which in contemporary philosophy mind seems to be equated with (an aspect of) consciousness, see the next section) by considering what it is we can know about being a bat, an organism with traits of mental states like human beings, but having a form of perception called echolocation that assists it in its functions and movements. And this form of perception is not like any of the senses humans have. This unique feature of bats makes it difficult for humans to imagine what it is like to be a bat or to experience the bat the way it experiences itself. Nagel (ibid.) thus holds that as experience is a subjective matter, it will be difficult for humans to understand what it is like to be a bat. He argues that a physicalist may explain the physical features of a bat, the chemical components and cellular functions of its brain; she can even imagine what it would be like for a human being to act and behave like a bat. However, the physicalist will fail to account for how the bat experiences its experience; what/how the bat feels as itself.

Brisbane Scott (2004: 3) argues for the immateriality and subjectivity of mental phenomena along the same lines as Thomas Nagel. She argues that mental properties such as thought, emotion, pains, and sensations do not have physical properties such as mass and shape (ibid.). Moreover mental properties are private and subjective. They occur in the private domain of the individual, who is conscious of them. The bitterness or sweetness of food is a private experience of the individual. Mental properties are thus distinguishable from physical

qualities that lie in the public domain due to their mass, weight and extension in space. Scott (ibid.: 4) holds that the existence of secondary qualities or properties like colours, taste, smell, sound etc. which are distinct from primary qualities like mass, weight, shape and size proves that mental properties are separate from physical properties and she then concludes that this favours (some version of) dualism (ibid.).

Nagel (2002: 223) believes that every conscious experience of any organism is only comprehensible and understandable from the organism's point of view. There is no way we can comprehend the subjective character of experience of another being from our own subjective viewpoint. Thus, for Nagel (ibid.: 225), for a proper understanding of a physical theory of mind, we should ask how does the subjective aspect of reality (consciousness in general, or subjective experiences in particular) relate to the objective aspect (for example, the brain)? His question implies that for us to understand the physicalist hypothesis that a mental event is a physical event (ibid.), there is a need to understand how mental terms and physical terms are the same. (Is the meaning of the term 'consciousness' the same thing as the meaning of the term 'brain'?) At present, we do not have a mechanism according to which we can have this kind of understanding. Hence, for Nagel (ibid.), we cannot yet claim that a physical term and a mental term refer to the same thing.

The reality of qualia or subjective experiences supports a notion of mental causation. It is the pain that I feel subjectively that will cause me to avoid situations that bring about pain. It is the sweetness of an orange that will cause me to love oranges because if an orange is bitter I may not love to eat it because of the bitterness. More so, the way I perceive and react to the outside world is determined by my subjective experience of the world.

Let us now discuss this phenomenon 'consciousness' in detail as it occupies a very important place in the philosophy of mind and in my thesis.

4.2. Consciousness

According to David Chalmers (1996: 1), the nature of consciousness is a very difficult phenomenon to grasp:

> Conscious experience is at once the most familiar thing in the world and the most mysterious. There is nothing we know about more directly than consciousness, but it is far from clear how to reconcile it with everything else we know. Why does it exist? What does it do? How could it possibly arise from lumpy gray matter? We know consciousness far more intimately than we know the rest of the world, but we understand the rest of the world far better than we understand consciousness...

"The abstract noun 'consciousness' is derived from the Latin *cons* (with) and *scire* (to know)" (Gennaro, 2017: 2). Generally, in contemporary philosophy of mind, a conscious mental state is characterised according to Thomas Nagel's notion of consciousness, ""what it is like". When I am in a conscious mental state, there is "something it is like" for me to be in that state from the subjective or first person point of view" (ibid.). For instance, when I hear Bob Marley's music, there is "something it feels like" to hear it from my perspective.

Ned Block (1995: 227) identifies two forms of consciousness namely phenomenal consciousness and access consciousness. Phenomenal consciousness is similar to Thomas Nagel's notion of subjective experience as noted above. Access consciousness is having access to what one thinks, believes, or desires. These mental states can be verbally reported or accounted for by the individual, reflected or reasoned about, even controlled by her (see e.g. Sturm, 2012.: 3). Access consciousness has no qualitative aspect like phenomenal consciousness. Thus, access consciousness plays a functional role as it is concerned with what a particular mental state does in relation to other mental states to produce a particular behaviour (Gennaro, 2017: 4).[4] In addition, consciousness is not the same as self-consciousness, which is a kind of awareness or consciousness of one's own mental states. Self-consciousness is either a bodily awareness or "the ability to reason and reflect on one's own mental states, such as one's

[4] Louise Antony (2007:148) argues for the reality of mental phenomena based on access consciousness as she appeals to reasoning and deliberation, intentional inexistence, opacity and predictive power. For her, these four points provide the strongest prima facie case for psychological realism.

beliefs and desire" (ibid.). However, I will argue that phenomenal consciousness is self-consciousness that is causally efficacious in the physical world in chapter four.

Many philosophers of mind are battling today with questions such as how a conscious mental state is related to the body. What makes a mental state to be a conscious mental state? Can brain activity explain consciousness? Arguably, the key issue in the philosophy of mind today is the problem of consciousness (ibid.: 1). The two competing theories in the philosophy of mind that attempt to address the issue of consciousness are dualism and materialism (physicalism) (ibid.: 6). The dualists, as noted earlier on, hold that the conscious mind or a conscious mental state is non-physical. The physicalists hold that the mind or conscious mental states are the products of neural activity. One reason for this stand could be the close connection between brain damage and various states of consciousness (ibid.: 10). Moreover, one is declared dead once a brain death occurs (ibid.). These two theoretical approaches, dualism and materialism (physicalism) are the foundations of other theses that are engaged in solving the problem of consciousness.

However, none of the theories or mechanisms that have been offered to solve the problem of consciousness are without their critics. The materialists have the problem of explaining how and why some brain states are always accompanied by conscious experience. Levine (1983) holds that there is an 'explanatory gap' between materialist accounts of the mind and consciousness, i.e. materialism has difficulty in explaining consciousness. How could a material body such as a brain that is spatio-temporal give rise to consciousness that is immaterial and non-spatio temporal? Thus, there is a huge gap in our attempt to explain the relationship between phenomenal properties and brain properties. Consequently, there is doubt about the possibility of a scientific treatment of consciousness. Dualists, in their turn, have the problem of explaining how the non-physical body, the mind, or conscious mental state can causally interact with the physical body, or brain properties (Gennaro, ibid.: 1; Sturm, 2012.: 3). One contemporary philosopher who has written extensively on consciousness is David Chalmers. I will consider some of his views on consciousness in the next section.

4.2.1. David Chalmers on Consciousness

David Chalmers is known for his numerous and important works on consciousness. It was he who fashioned or formulated the 'easy' and the 'hard' problems of consciousness. In his article (2002: 247) "*Consciousness and Its Place in Nature",* he notes that the easy problem of consciousness is the problem of the phenomena of mental states "to discriminate stimuli, to report information, to monitor internal states, or to control behaviour" (ibid.). A physical system such as a computer could be 'conscious' in any of those ways. The phenomena of easy problems are susceptible to the standard methods of cognitive science, and they could be explained in terms of computation and neurobiology (ibid.). Their task is to explain how some causal role is played in the cognitive system, ultimately in the production of behaviour. Thus, they can be functionally defined, and relate to Block's notion of access consciousness.

The hard problem of consciousness is not the problem of explaining cognitive functions or, in other words, to explain causal roles of the cognitive system, but rather to explain why these causal roles such as discrimination, integration, access, verbal reporting and control are *accompanied by experience*. The hard problem of consciousness includes explaining states of perceptual experience, bodily sensation, mental imagery, emotional experience, and occurrent thought. There is something it is like to see a vivid green, to feel a sharp pain, to visualize Paris, to feel a deep regret, and to think that one is late. Each of these states has a phenomenal character, with the phenomenal properties (or qualia) characterizing what it is like to be in the state (ibid.). Thus, the hard problem relates to Block's notion of phenomenal consciousness.

Chalmers argues that the claim that experience is closely associated with the brain is never doubted, but how and why does the brain give rise to experience? Explaining why physical processes in the brain are accompanied by states of experience, is the central problem of consciousness (ibid.: 248). A reductive explanation of consciousness will address it based on physical principles. A physicalist will treat consciousness as a neuro-physical or neurobiological process. For the non-physicalist, consciousness is non-physical, even if it is closely associated with physical processes (ibid.). A non-reductive physicalist will treat consciousness as a mental phenomenon as part of the solution to the problem of understanding consciousness.

Frank B. Dilley (2004: 1) argues that materialism, as a reductive view, cannot account for consciousness because **'**consciousness**'** or 'self' is not physical but rather is a non-material substance. He (ibid.: 1) holds that we need a 'substance view of self or consciousness' (Cartesian dualism) to account for the reality of consciousness. He holds that abandoning Cartesian dualism as the materialistic philosophers of mind are advocating, would mean abandoning the values that define and characterise our human nature such as free will, moral responsibility, apportioning of credits and blames, and rationality. To abandon consciousness, self, or soul implies embracing materialism that promotes determinism, which cancels out our sense of moral responsibility and free will.

The values Dilley has mentioned are (linked to moral responsibility and free will) meaningful and real based on mental causation (that the mental properties or events are capable of causing other mental events and physical events as *qua* mental). However, mental causation is a big problem for Cartesian dualism, as noted before, because of the interaction problem between two independent substances, one material and the other immaterial. On the Cartesian dualist theory, moral responsibility and free will cannot be accounted for. For reductive theorists such as Kim, mental properties or events are causally efficacious only if they are reduced to physical properties. Thus, mental properties and events cause events as physical properties. I argue otherwise that mental properties and events (consciousness) cause events *qua* mental and they are causally efficacious *qua* their immateriality.

Chalmers (2002: 248), in turn, holds that the concept of consciousness seems to resist any materialist solution as shown by means of three arguments namely the explanatory argument, the conceivability argument, and the knowledge argument (ibid.). These arguments show that there is an epistemic gap between phenomenal and physical facts. In other words, physical facts do not entail phenomenal facts. And it is argued that if there is an epistemic gap between phenomenal and physical facts, then there is an ontological gap (ibid.).

The explanatory argument is about the difference between the easy problem and the hard problem of consciousness (ibid.) – and as such we have already touched on it. The easy problems deal with the explanation of behavioural and cognitive functions, which do not

concern the hard problem (ibid.). As stated before, the ""easy problems" of consciousness are, for example, the ability to discriminate and categorize stimuli, the ability of the cognitive system to access its own internal states, and the difference between wakefulness and sleep generally have more to do with the functions of consciousness, but Chalmers urges that solving them does not touch the hard problem of consciousness" (Gennaro, 2017: 11). These explanations are however not adequate to explain the hard problem of consciousness (the problem of consciousness). "The hard problem of consciousness … basically refers to the difficulty of explaining just how physical processes in the brain give rise to subjective conscious experience. ... How can we explain why there is something it is like to entertain a mental image, or to experience an emotion?" (Chalmers, 1995: 201).

The explanatory argument runs like this:

1. "Physical accounts explain at most structure and function.
2. Explaining structure and function does not suffice to explain consciousness.
3. (Therefore) no physical account can explain consciousness" (Chalmers, 2002: 248).

A physical account can thus solve the easy problems (which involve explaining functions), but something more is needed to solve the hard problem. According to Chalmers, most philosophers only try to solve the easy problems such as Block's "access consciousness" while ignoring the hard problem (phenomenal consciousness) (Gennaro, 2017: 11).

Another argument that shows that materialism cannot account for consciousness is the conceivability argument. Briefly the idea is that we could conceive a zombie, a cognitive system that is physically identical to us (conscious beings), but which has no conscious states. That we can conceive of zombies shows that it is metaphysically possible that there exist zombies. If it is metaphysically possible that zombies exist, then consciousness is non-physical (Chalmers, 2002: 249). This argument is like this:

1. "It is conceivable that there be zombies

2. If it is conceivable that there be zombies, it is metaphysically possible that there be zombies.

3. If it is metaphysically possible that there be zombies, then consciousness is non-physical.

4. (Therefore) consciousness is nonphysical" (ibid.).

Brie Gertler (2008: 305) also uses the conceivable argument to argue for the reality of the mental *via* a thought experiment she calls 'a conceivability test' that shows the presence of pain in the absence of a body (disembodied pain). Because pain can be conceived in the absence of a body, it shows that pain is not identical to the body. In her thought experiment, she asks an individual to engage in the act of imagining or conceiving. Let the individual pinch herself to experience pain. In this act of pinching, these five processes should be followed by the individual being tested (ibid.: 306):

1. I should believe strongly that I could conceive of experiencing pain without my physical features or in a disembodied body. 2. If I can conceive of experiencing a particular scenario then that scenario is possible. 3. It is possible that this very pain occurs in a disembodied being. 4. If this pain is identical to some physical state, then it could not possibly occur in a disembodied being. 5. This very pain is not identical to any physical state. She (ibid.: 307)) concludes that the identity thesis (reductive materialism) is false based on these five steps, since these steps imply pain can occur in a disembodied being. Pain is thus not identifiable to a physical state.

In its turn, the knowledge argument is intended to show that consciousness is not deducible from the physical. Someone could know all the physical facts about an event or phenomenon and reason properly about them and still be unable to know all conscious – subjective – facts. Frank Jackson (1982)'s famous thought experiment on this issue supposes that Mary is a neuroscientist who knows everything about the physical processes relevant to colour vision. However, Mary, all her life, has lived in a black and white room and has never experienced a red colour. In other words, she does not know what it feels like to see red. But,

when Mary is taken outside her black and white room one day, she experiences red for the first time, and this implies that she will know a new fact which previously she did not know of. She will come to know what it is like to see red. Thus, physical facts do not exhaust all facts about consciousness. And so, materialism is false in terms of its explanation of consciousness.

1. "Mary knows all the physical facts
2. Mary does not know all the facts.
3. (Therefore) physical facts do not exhaust all the facts" (Chalmers, 2002: 250).

Chalmers (ibid.: 250) notes that physicalists have many ways of resisting the above-mentioned epistemic arguments that hold that that there is a difference between phenomenal facts and physical facts, in other words, arguments claiming that consciousness is not a neuro-physical or biochemical event. Some reductive physicalists hold that physical facts and phenomenal facts are about the same reality; hence, there is no epistemic gap. For them, conscious states are brain states. There is then no hard problem of consciousness on their terms. Some reductive physicalists hold that there may be an epistemic gap between the physical and the phenomenal domains, but there is no ontological gap (ibid.). For them, conceivability does not imply possibility. Thus, the zombies and the like might be conceivable, but they are not metaphysically possible (ibid.).

Some of the reductive physicalists acknowledge a deep epistemic gap between phenomenal facts and physical facts, but believe that the apparent gap may be due to our own limitations and it is closable in the future. Nagel (1974) has played with the idea that one day we would understand how consciousness could be physical, just as pre-Socratic philosophers could have understood today how matter could be energy. McGinn (1989), on the other hand, argues that the gap may be unclosable by humans due to the serious limitations in our cognitive skills, while Churchland (1997) seems to be with Nagel by holding that future science

will overcome the apparent epistemic gap, however, acknowledging that at present we do not have a complete physics (Chalmers, 2002: 258-259).

However, Chalmers' commitment to the hard problem of consciousness, the 'what it is like' experience, makes him discard any reductive materialist's response to the three epistemic arguments. Chalmers holds that physical facts and phenomenal facts are fundamentally different. Thus, according to Chalmers (ibid.: 259), no complete physics can close the gap between the physical facts and phenomenal facts in so far as physics deals with structure and dynamics and can only solve the easy problems. But some reductive physicalists, like Jaegwon Kim, disagree with him as we shall see in chapter three. For Chalmers however, those novel theories such as relativity and quantum mechanics may introduce new structures and dynamics over old structures but the gap between physical facts and phenomenal facts will remain (ibid.: 259).

Chalmers looks elsewhere for the solution to close the gap and argues that the solution must be one that takes consciousness seriously. There are two possibilities. First, consciousness or phenomenal properties are a fundamental part of the world just like space or time. Second, consciousness may not be a fundamental part of the world, but it is caused by some more fundamental property 'X' that is not itself caused by physical entities. Here 'X' is seen as the proto-phenomenal property, which is the fundamental property. The question is how to integrate this proto-phenomenal property with the principle that the microphysical features of the world are causally closed.

Non-reductivism may be of help here (ibid.: 261). And there are three options for the non-reductivist. The first option is to accept a version of dualism, a view that holds that phenomenal facts are distinct from physical facts, and microphysics is not causally closed since phenomenal facts and physical facts affect each other. This position is also known as 'interactionism', which is the same as Cartesian dualism. The main objection to this theory is that it is incompatible with physics as microphysics is in fact causally closed so there is no way a mental state can cause a physical state.

According to Chalmers (ibid.: 264) another option for the non-reductivist is to accept that phenomenal properties are ontologically different from physical properties, and to hold that phenomenal properties are epiphenomenal, because they cannot causally affect the physical properties. This view is called epiphenomenalism. According to this view, physical states can cause phenomenal states but not vice versa. Here, psychophysical laws run in one direction only, from physical states to phenomenal states. Phenomenal states, such as beliefs, have no effect on our actions according to this view. But epiphenomenalism is often rejected because it is counterintuitive. For instance, is may be true that a sensation of pain plays no causal role in my hands moving away from a flame, while nevertheless, there is a natural inkling that a sensation of pain does cause me do something (ibid.: 264).

The last option for the non-reductivist to approach to the hard problem of consciousness is the view that Chalmers advances as '-type F monism' (ibid.: 265). Monists hold that there is only one kind of reality as opposed to dualists who profess two kinds of reality. On Chalmers' view, consciousness and physical properties are ontologically intertwined. The intrinsic properties of the fundamental physical world are inherently phenomenal properties, and they underlie physical reality itself. This view has its roots in the work of Bertrand Russell (1927), who, in *The Analysis of Matter*, has pointed out that physics defines physical entities and properties by their relations to one another and to us. Russell held, in addition, that there should be some underlying intrinsic properties that ground the disposition and relations of the entities among themselves. But physics cannot account for the intrinsic nature of these entities; it is silent about them. Philip Goff (2017) reflecting on Russell's work:

> Throughout the whole hierarchy of the physical science [neuroscience to physics] we learn only about causal relationship. And yet there must be more to the nature of physical entity, such as cerebellum, than its causal relationship. There must be some intrinsic nature of the cerebellum, some way it is in and of itself independently of what it does. About the intrinsic nature physical science remains silent.

Thus, we have a metaphysical problem to answer, namely what are the intrinsic properties of fundamental physical reality? (Chalmers, 2002: 265). Another metaphysical problem is how to integrate phenomenal properties with the physical world. Phenomenal properties as intrinsic properties seem foreign to the dynamic character of physical theory; nevertheless, they are the only intrinsic properties we know (ibid.).

Russell proposed a solution for both problems; either the intrinsic properties of the physical world are the phenomenal properties themselves or they are not phenomenal properties but constitute phenomenal properties. Thus, consciousness (as phenomenal consciousness) and physical reality are deeply intertwined (ibid.). This view will integrate phenomenal and physical properties together in the natural world. Phenomenal properties and physical properties are intertwined and are in a causal relation with each other. In this case, nature has entities with intrinsic (proto) phenomenal qualities, which are in causal relations within a space-time manifold (Ibid). Physics emerges from the relations between these entities, whereas consciousness or phenomenal properties emerge from their intrinsic nature. Thus, phenomenal properties are the intrinsic properties of fundamental physical systems (ibid.).

This view, as it acknowledges the reality of physics, is compatible with the principle of causal closure of microphysics and physical laws. More so, it advances the causal role for consciousness in the physical world, as proto-phenomenal properties are the fundamental foundation for all physical causation (ibid.).This view also has some common characteristics with both materialism and dualism. It is akin to materialism since physical terms not only refer to physical properties but also to the proto- phenomenal properties. It has something in common with dualism because it holds that the proto-phenomenal properties are ontologically fundamental, and there is separation between structural properties that define physical theory and intrinsic proto-phenomenal properties that define consciousness.

Russell's view also has some aspects in common with neutral monism because the underlying proto-phenomenal properties constitute the physical domain through their relations and the phenomenal domain through their collective intrinsic nature. Furthermore, in relation to phenomenal form, it can stand as a kind of idealism as the proto-phenomenal properties

constitute physical properties, but these properties need not appear as mental properties in the mind of an observer. This view can also be likened to pan-psychism as the proto-phenomenal properties are everywhere at the fundamental level of reality. Hence, the view is somehow called *pan-protopsychism* for the fact that the proto-phenomenal properties comprise all the physical reality. On these terms, consciousness is found in all things (ibid.).

For Chalmers (ibid.: 267), similar to Russell, his type F-monism, though speculative in nature, creates an integrated view of nature whereby physical properties and phenomenal properties are intertwined. However, Chalmers admits that no one has delved into this area of study deep enough and holds that this view will provide grounds for further research in order to show the complete integration of consciousness and physical properties in the natural world. I will show in chapter four that this view has something in common with Aristotle's hylomorphism in terms of which a every substance is made of *hyle* (matter) and represents physical properties on the one hand, and of *morphe (*form) which represents conscious or mental properties; so that just as physical properties and phenomenal properties are intertwined for Russell and Chalmers, so matter and form are intertwined in every substance according to Aristotle's hylomorphic theory.

The immaterial nature of consciousness is also displayed in the unity of consciousness because it is only a thing that is immaterial, not defined by a particular sensible object, can unify all our experiences into one, just as the immaterial mind can think about anything as it is immaterial, not limited to any particular physical object. Let us look at the unity of consciousness.

4.2.2. The Unity of Consciousness

Unity is another aspect of consciousness that concerns the organization of our experiences. We do have many experiences at the same time such as such as sights, sounds, smells, and other sensations as well as memories, thoughts and emotions, and all these experiences are sometimes unified into one experience. Gennaro (2017:29) writes:

> ... When one looks at how the brain processes information, one only sees discrete regions of the cortex processing separate aspects of perceptual objects. Even different aspects of the same object, such as its colour and shape, are processed in different parts of the brain. Given that there is no 'Cartesian theatre' in the brain where all this information comes together, the problem arises as to just how the resulting conscious experience is unified. What mechanisms allow us to experience the world in such a unified way? What happens when this unity breaks down, as in various pathological cases? The problem of integrating the information processed by different regions of the brain is known as the binding problem.

It is clear that the binding problem and the hard problem of consciousness are intertwined in the sense that the solution to the binding problem arguably implies the identification of the elusive neural correlate of consciousness (ibid.: 29).

Chalmers (2010: 497-548) has given different kinds of explanations for the unity of consciousness, namely the unity of elements of consciousness, subject unity and subsumptive unity. According to the unity of elements of consciousness, all the different aspects of an object form a unity. For instance, someone looks at an object and sees that the object is a ball; it is white, circular in shape and hard. And she sees that this object, ball, is on the top of a table, etc. She understands that these different elements make up her subjective experience of the ball.

According to subject unity, two elements of consciousness that are experienced both belong to the self at the same time (Chalmers, 2003: 5). Subject unity is divided into two types 'phenomenal unity' and 'access unity'. In terms of 'phenomenal unity', two conscious states are experienced at once. In a phenomenal conscious state; there is something it is like to be in that state. Being in that state is something private and subjective to me. It is a sort of subjective experience. For example, there is something it is like for me to listen to Bob Marley's music and to feel love (ibid.).There is something it is like to feel headache. There is a phenomenal unity

when two phenomenal conscious states are jointly experienced as representing a single object. For example: when I am conscious of a blue car, I experience the presence of blue and I experience the presence of a car, but I also experience a blue car. In this instance, there is a unity of what it is like to experience the *blueness* and *carness* at the same time; two states are unified by being experienced as parts of a single object (ibid.).

While in terms of access unity, two conscious states can be accessed at once for verbal report, reasoning and deliberate control of behaviour. For example: I look at a blue car, I can report the presence of the car ("there is a blue car"), I can think about the car (that it is a good car and it will be nice to test drive it), and its presence may direct my behaviour (I take my driving license in order to drive it to the town). My seeing the car affords me a sort of access information about the blue car (ibid.: 4). I have access conscious of the blue car. Access unity is when two conscious states are unified so that their contents are jointly accessible and their junction is available for verbal report, reason and behaviour direction. For example, I am conscious of the blue car. I can report the presence of blue as well as the presence of car (ibid.). I can also report the presence of the blue car, and the presence of the blue car can influence my reason and behaviour. Hence, my seeing of blue and my seeing of a car are access unified, not just individually access conscious. I have access to the contents of both states at once (ibid.).

In terms of subsumptive unity, a unified consciousness involves particular experiences being subsumed in a more complex experience or a single state of consciousness (Brooks, et. al. 2017), particular experiences are "aspects of a single encompassing state of consciousness" (Bayne and Chalmers, 2003). The particular experiences subsumed into a single state of consciousness yields to what is called a conjoint phenomenology: a phenomenology that encompasses the phenomenology of the individual states (Brooks, et.al. 2017) "there is something it is like for the subject to be in [two conscious] states simultaneously" (Bayne, 2010: 32). For instance, my single experience of a car may involve different experiential parts such as *carness*, *blueness*, motion and sound of the car. Subsumptive unity is that these different experiential parts; *carness, blueness*, motion and sound are subsumed into a single phenomenological experience.

The unity of consciousness or the consciousness of many things seems to create a problem for any neuroscience of consciousness because of the fact that the whole contents of consciousness should be unified on the one hand but at the same time, on the other hand, the distinctiveness of each aspect of its contents should be maintained. Can neuroscience account for this? Chalmers (1991) has reviewed this possibility but we shall discuss his critique in chapter two after we have discussed physicalism.

4.3. Personal Reflection on Consciousness

Based on Chalmers's type 'F' monism that holds that consciousness is a fundamental part of reality and it is intertwined with the physical reality in every substance, every object is fundamentally conscious. But I would say it is of different degrees. The consciousness of plants cannot be of the same degree with that of human beings who are capable of thinking and having qualia. As mentioned, for Block consciousness has two forms namely access consciousness and phenomenal consciousness. I don't think we should see these two forms as two separate forms of consciousness. It is one consciousness which can, at certain time, be of access form or phenomenal form. As access consciousness, consciousness is related to mental states such as thinking, desire and beliefs, and as phenomenal consciousness, it is about the qualia and subjective experience. Therefore, every mental state entails consciousness. I would say then that the mind-body problem could be regarded as a 'consciousness-body' problem.

Consciousness is not divided into two parts but in acting in human beings, it can appear as access consciousness or phenomenal consciousness, and unpacked in terms of the easy or hard problem of consciousness, depending on the context. This notion of consciousness can be likened to the concept of 'form' in Aristotle's theory of hylomorphism, the theory that holds that every substance is made of form and matter. I will offer arguments for this claim in chapter four. In other words, I will argue that every substance is made of consciousness and matter. Consciousness as form defines the function of the organism. Thus, plants and animals have different types of consciousness. However, plant's consciousness and animal's consciousness are also found in human beings as human beings are capable of performing most of plants and animals' functions such as nutrition and perception. This relates also to Aristotle's theory of the

soul – the form of human beings is the soul, the principle of life in human beings. And according to Aristotle, there are three kinds of soul namely the vegetative soul that is responsible for nutrition (found in plants, animals and human beings), the sensitive soul that is responsible for feeling and perception (found in animals and human beings) and the rational soul that is responsible for thinking (found only in human beings).

More so, I will argue in chapter four that consciousness in the sense of both access and phenomenal consciousness can be unpacked via Aristotle's notion of the intellect (rational soul) which distinguishes between the passive intellect and the active intellect. Be reminded that Block in his turn distinguishes between two aspects of consciousness in human beings access consciousness (the easy problem) and phenomenal consciousness (the hard problem). And both the access consciousness and phenomenal consciousness operate together in helping human beings to perceive, react and navigate the world. Through the access consciousness, sensory information is made available to the rest of the mind for a human person to have an experience of what it is like to perceive something. Before one experiences 'what it is like' to hear Bob Marley's music, one has to be able to actually hear the music. I will discuss Aristotle's hylomorphism in detail later in chapters four and five.

Furthermore consciousness is immaterial as noted earlier. If it were material, the phenomenal or hard problem of consciousness which is identified as the qualia problem, would be solved by physicalism that holds that everything is physical or reducible to physics. As we shall see later in this thesis, accounting for phenomenal consciousness or the subjective quality of experience (qualia) is the biggest problem of materialism or physicalism. Neuroscience can today cater for access consciousness and solve the easy problem of consciousness but the hard problem of phenomenal consciousness is not subject to neuro-scientific findings. Mel Thompson (2012: 48) notes that neuroscience can help us to understand how various parts of the brain align with physical and mental operation. Neuroscience, today, helps us to observe the activity of the living brain, which was impossible before. Hence, some people hope for the possibility in the future that neuroscience will be able to answer the traditional questions about the mind and its relationship to the body, or about the nature of the human person (ibid.: 67).

Through neuroscience we know that certain parts of the brain are associated with particular mental and conscious functions – sight, hearing, language, emotions and so on and that, if the part of the brain is disabled (e.g. through a stroke), other parts may be able to take over some of its functions (Thompson, 2012: 78). Though there could be a correlation between the mind and the neural activity in the brain, neuroscience cannot solve the mind-body problem as its investigation into the brain only informs us about neurons, neural pathways or neural activities but not about conscious acts such as hearing, seeing, touching, remembering etc. Neuroscience cannot answer the question 'what it is like' to be us as it is a physicalist theory. It cannot explain how the mind and body interact. It cannot even describe the nature of the correlation between the mind and the brain. We thus need some other solution to the hard problem.

5. Conclusion

Mental causation is concerned with the causal efficacy of the mental, especially the phenomenal form of consciousness. I have promised to offer arguments for claiming that phenomenal consciousness is causally active in the physical world, because it is real. Because of the causal efficacy of the mental, which I have argued to be interpreted as (both access and phenomenal) consciousness, human beings are agents responsible for their actions and behaviour. Phenomenal consciousness does not need to be reduced to physical properties for it to be real or cause events. I will show in what follows that the causal efficacy of phenomenal consciousness is not undermined by Kim's causal exclusion and causal closure principles that underlie his notion of mental causation. Let us now first turn to a discussion of the functionalist notions of mental causation.

Chapter 2: Physicalist and Functionalist Notions of Mental Causation

1. Introduction

In the previous chapter, I mapped out the playing ground for my thesis by defining some of the key concepts such as mental realism, mental causation and the immateriality of the mental. I stated that mental reality means causal efficacy of the mental *qua* mental, not in Kim's understanding of it, which is that mental causal efficacy is only obtained when it is functionally reduced to physical properties. And I stated too when I refer to mental causation I mean the causal efficacy of the immaterial consciousness that embodies all other mental properties and events. Though consciousness is of two kinds namely access consciousness and phenomenal consciousness, I purported that I mean in the context of this thesis primarily phenomenal consciousness and its causal efficaciousness. The mental in this sense is real for it has power to cause human action, and affect mental and non-mental events. Human action in these terms is voluntary action because of the causal efficacy of the phenomenal consciousness, which is beyond the explanatory power of physicalism.

Given the focus of this thesis on Kim's notion of mental causation, which is reductive physicalistic in nature, I remind the reader that my aim is to argue that Aristotle's philosophy of mind can save non-reductive physicalism from Kim's epiphenomenal accusations. In this chapter, I will show that physicalism, which comes in two forms namely reductive physicalism and non-reductive physicalism, and functionalism, both have truncated ideas of mental causation, and hence cannot account for phenomenal consciousness or qualia.

I will begin by considering a general definition of physicalism. Thereafter I will I discuss the historical development of physicalism and types of physicalism. Then it will be a discussion on the reductive physicalism and mental causation that will cover identity theory. Having discussed identity theory, I will look at property dualism, strengths and weaknesses of non-reductivism and functionalism, and criticism and reflection on functionalism.

1.1. Physicalism Defined

The terms 'physicalism' and 'materialism' are interchangeable in the sense that both are monistic and both claims everything that exists is physical. However, the term 'materialism' is very old, while the term 'physicalism' was introduced in philosophy only in the 1930's by Otto Neurath and Rudolf Carnap (1959/1932) (Stoljar, 2015:1).

Physicalism is a theory that states that everything is or may be reduced to the physical or that everything supervenes on the physical. Thus, the universe is physical in nature. Physicalism denies what dualism affirms, which is that humans are made up out of two substances, namely the mind and the body. For physicalists, everything, including thought, desire, consciousness, etc. is purely material or physical (Maslin, 2007: 64).

1.2. The Term 'Physical'

Andrew Melnyk (2007:2) in explaining the term 'physical' writes that a physical thing is a spatio-temporal object, unlike the Cartesian dualist notion of the mental. Physical objects possess essentially the properties of shape, size, and solidity and are capable of motion. According to Melnyk (ibid.), the term 'physical' can be explained in two ways namely narrow sense and broad sense. In the narrow sense of the term, mind and mental properties are clearly not physical as they are not spatio-temporal phenomena, and they are not objects of any theory in fundamental physics. However, in the broad sense of the term physicalism, human minds, mental properties and processes can be physical. They are physical because they can meet two conditions that make an individual item physical (1) they can be accounted for or explained by using the vocabulary of fundamental physics; and (2) they are physically realized items of functional kinds. They are reducible to the physical. For instance, pain is not mental qua mental but it is the C.fibers (neurons) firing. Hence, pain is described as physical because it is what the C.fibers firing does. A functional item, according to Melnyk, is physically realized, when its realizer conforms to the theories of fundamental physics and it is subject to physical laws and other physical conditions (ibid.).

Having explained the term 'physical', let us now look at the historical development of physicalism.

2. The Historical Development of Physicalism

Stephen Priest (1991), in his book, *Theories of Mind,* discusses the history of physicalism. Physicalism, or rather materialism as it was referred to, began with Democritus in the fourth century BC. Democritus held that every substance is made of atoms, the smallest and imperceptible particles of elements (ibid.: 99). Epicurus, another materialist, built his humanistic ethical theory on the atomism of Democritus. He held that the human soul is mortal because it is composed of atoms that are separable (ibid.). For Epicurus, qualities such as sweetness, bitterness, and colour exist by convention, and only the atoms exist in reality. During the 17th century, Thomas Hobbes, an English philosopher, was attracted by the atomic theory of Democritus. He believed that just like atoms, everything that exists has a physical dimension of size, height, depth and weight. In other words, everything is measurable. In addition, he claimed that all our thoughts and sensations are caused by matter and they are physical in nature (ibid.). Hobbes also taught that even God and the soul are physical and are composed of invisible particles of matter.

Julien Offray de La Mettrie, an enlightenment thinker, presented a purely mechanistic view of human thought and action in his book *Man a Machine* (*L'Homme machine)* (1748). He described the human person as a highly complicated physical object with moving parts in contrast to Descartes' notion of a human person as essentially an immaterial soul. Thought and sensation for De La Mettrie are nothing but complicated motions of matter. It was Baron d' Hollbach (1770), who in his book, *System of Nature*, gave complete expression to the materialism of the French Enlightenment when he argued that the universe is one enormous deterministic system of physical objects (Priest, 1991: 101). According to Hollbach, the world should be viewed from the scientific point of view and there was a clarion call from many angles to apply this scientific approach to the study of mind.

In the twentieth century, with science having made big strides and progress in atomic theory, evolution, neuroscience, and computer technology, the term 'physicalism' came to be used more often and physicalism of various types became a prominent doctrine in the philosophy of mind. J.J.C. Smart (2000:86) writes:

> It seems to me science is increasingly giving us a viewpoint whereby organisms are able to be seen as physical-chemical mechanisms … that is, for a full description of what is going on in a man, you would have to mention not only physical processes in his tissues, glands, nervous system and so forth, but also his states of consciousness: his visual, auditory and tactual sensations, his aches and pains.

Having traced aspects of the history of physicalism as a theory in the philosophy of mind, let us look at the various types of physicalism in relation to mental causation in the philosophy of mind.

3. Types of Physicalism

Physicalism comes in two main versions namely reductive physicalism and non-reductive physicalism. Reductive physicalism holds that every mental state or property is identical to or is reducible to a physical state or property. The mental is nothing other than the physical. According to this view, the mental *qua* mental cannot cause any physical body to act. Non-reductive physicalism on the other hand, holds that a mental state is a physical state but denies the reduction of mental states to physical or brain states. According to this view, there are mental phenomena or properties that supervene on or emerge from the physical. Non-reductive physicalism thus acknowledges at least some kind of causal relation between mental properties and physical properties.

Reductive physicalism is also known as the identity thesis or mind-brain identity theory or central state materialism, as it is a theory that states that the mind is identical to the brain in all respects. Even human consciousness is reducible to or explainable in terms of natural processes in reductive physicalist accounts of the mind. Jaegwon Kim, himself a reductive physicalist, writes that: “Mental states and processes are to be construed as states and processes occurring in certain complex physical states such as biological organisms, not as states of some ghostly immaterial beings” (Kim, 1995: 579).

Now, let consider the reductive physicalist response to mental causation, and then we will look at the non-reductive physicalist response to mental causation. It will be good to remind the reader that mental causation in this context is the ability of mental properties and events to cause other mental and physical events or properties *qua* mental. I will show that neither reductive nor non-reductive physicalism offers an adequate account of mental causation. Their inability to account for the mental *qua* mental, will then lead us to looking at functionalism, as a view that professes also to account for mental causation, to see if it does so adequately.

4. Reductive physicalism and Mental Causation

4.1. Behaviourism

Gilbert Ryle (1949:11) refers to Descartes' dualistic theory of mind as the 'official theory'. It is referred to as the 'official theory' because intellectually influential persons ranging from philosophers, psychologists, lawyers, scientists, and doctors, to religious clerics accepted it as the standard view of human nature. However, Ryle (ibid.) maintains that the principles of the Cartesian dualistic theory are unsound, self-defeating and against our knowledge about the mind.

As we have seen in section 1.3.1, Descartes holds that every human being possesses a mind and a body, and that while the body is subject to mechanical laws; the mind is subject to non-mechanical laws. Thus, according to Ryle, "Descartes divides a person's life into two, one is external and the other is internal. Ryle (ibid.: 12) therefore calls the official theory a metaphor, because the body and mind cannot influence each other as they belong to the same category; even if Descartes claims the mind is un-extended and the body is extended. The official theory gives us no direct way of knowing the mind or inner life of others. Ryle's counter claim is that following Descartes's theory, we can never know anything about other's mental lives, while, in fact, we do know a lot about the minds of others through their behaviour.

Ryle (ibid.: 16), furthermore, calls Descartes theory a myth. Ryle accuses Descartes of creating a 'ghost in the machine' (ibid.); mind being the ghost that is in a machine that is the

body. Moreover, for Ryle, Descartes' theory is a false theory because it is built on a mistake that he calls a category mistake (ibid.). According to Ryle, category "mistakes are those made by people who are perfectly competent to apply concepts, at least in the situations with which they are familiar, but are still liable in their abstract thinking to allocate those concepts to logical types to which they do not belong" (ibid.: 17). For Descartes, the mind is not in the same category as the physical body, while actually it is, according to Ryle. For Descartes, the mind and the body are two different sorts of things. Mental processes and bodily processes are and have different sorts of causes and effect. The implication of this is that there cannot be causal interaction between them. However, Ryle argues that Descartes actually did put the mind and body in the same category by assuming that the mind is governed by rigid non-mechanical laws as the body is governed by rigid mechanical laws (Nath, 2015:3).

Contrary to Descartes, Ryle holds that minds or mental phenomena do not refer to some ghostly mental events or states but, rather, that they are at issue when we observe the concept-specific behaviour of others. Ryle (1949: 15), thus offers a behaviouristic counter to Descartes. For him, every mental phenomenon is a *disposition* to behave or act in a certain way. Mental concepts are dispositional concepts because they often signify our abilities or tendencies to act in certain ways. For Descartes, the first person characteristics such as consciousness and self-knowledge are real human characteristics even if they are different from the bodily processes that are open to the public. However, Ryle argues that self-knowledge cannot be known through introspection or consciousness. We have knowledge of others and ourselves through observing the daily behaviour or conduct of others and ourselves. Ryle thus does not see the mind or consciousness as separate from our daily activities, rather he holds a behaviourist view of minds. Mental talk is a description of how people behave or are disposed to behave. For instance, we know that one is intelligent through her utterances or behaviour, or one is known to be kind when one performs actions that are deemed kind.

Ryle's criticism of Descartes's theory led him to develop his own theory of mind. Ryle's theory is referred to as logical behaviourism (Ryle, 1949: 15). Logical behaviourism deals with the meaning of mental terms (Fodor 1981). A mental term is about the behaviour or disposition

of an organism to behave in a certain way. That John is thirsty would mean that if water were available, then John would drink some water**.** Logical behaviourism has been labeled as a physicalist theory of mind because of the "if-then statement", called a behavioural hypothetical that expresses a behavioural disposition (Fodor, 1981: 2). Fodor argues further that "By definition a behavioural hypothetical includes no mental terms. Then if-clause of the hypothetical speaks only of stimuli and the then-clause speaks only of behavioural responses. Since stimuli and responses are physical events, logical behaviourism is a species of materialism**"** (ibid.).

Ryle does not deny the existence of the mind, but argues against Descartes' bifurcation of mind and body. However, Nath (2015: 4), just like Hilary Putnam (2012: 45) in 'Brains and Behaviour' and Ned Block (2012: 97) in 'Troubles with Functionalism', holds that Ryle's concept of mind has many defects such as its inability to elucidate higher processes of mind like creative thinking, artistic imagination and integral vision of things – and we may add phenomenal consciousness, as these are mental concepts that occur independent of behavior. The argument is that these mental processes occur internally in the mind without any externalization of them in public but Ryle's theory is externally oriented as it focuses on behavioural displays. Therefore, the claim is that Ryle does not do justice to the realm of mental life. Nath (2015: 4) attributes this to the fact that Ryle was influenced by behaviourism's prejudices against a dualistic philosophy of mind.

The fact that a mental state can be explained through physical action does not on its own mean the same as the identification or reduction of that mental state with or to the physical processes through which it is expressed. This means that Ryle fails to account for mental states as his theory focuses on external behaviour while mental states like thinking and artistic imagination are internal occurrences in the mind. And to equate the mental state with the physical disposition creates difficulty in accounting for the influence of one mental state on another or causing another. Logical behaviourism is more or less silent about mental causation. Jerry Fodor (1981: 3) writes:

> Event causation actually seems to be quite common in the realm of the mental. Mental causes typically give rise to behavioural effects by virtue of their interaction with other mental causes. For example, having a headache causes a disposition to take aspirin only if one also has the desire to get rid of the headache, the belief that aspirin exists, the belief that taking aspirin reduces headaches and so on. Since mental states interact in generating behaviour, it will be necessary to find a construal of psychological explanations that posits mental processes: causal sequences of mental events. It is this construal that logical behaviourism fails to provide.

It is an obvious fact that we plan an action before that particular action is carried out. This planning is done only in the mind. Through thinking and logical analysis, we compare and contrast different plans with one another before adopting a particular one for execution. In other words, planning comes first before action (see, e.g. Thompson 2012:12). Moreover, and related, Ryle's theory cannot account for the phenomenology of our own mental states, the feel which is neither dependent on nor available to third-person conditions. It cannot account for qualia as it ignores completely an account of what goes in the mind.

However, there are positive aspects to Ryle's view; for instance, his argument becomes useful in the verification of the meaning of words we sometimes employ in our daily communication. We can only understand the meaning of the sentence that I was kind by showing the action performed by me (ibid.: 22).

U. T. Place (2000: 78), in arguing against behaviourism, holds that consciousness for the behaviourist is a special kind of behaviour. However, according to Place (ibid.: 78) while some mental events such as wanting, knowing, remembering, and believing can be understood behaviourally as dispositions, nevertheless, there are mental concepts and events that unavoidably require an inner process story such as consciousness, experience, sensation and

mental imagery. For Place, nevertheless, the reality of an inner process account does not entail dualism because it is correlated with physiological processes.

J. J. C. Smart (2000: 85), in turn, holds that behaviourism has no room for sensations as, according to behaviourism, a person is just a physical being and any mental event is an expression of behaviour or a disposition to behave in a certain way. For instance, when I say I am in pain it means that I am crying or making a sophisticated wince. However, for Smart, behaviourism is an inadequate theory for accounting for mental events or processes (ibid.) as he argues that to be in pain means to be in distress, which means to be in a certain agitated condition. Thus, there is something going on in the brain when we have sensations (ibid.).

Armstrong, like Hilary Putnam, argues that Ryle's particular version of behaviourism is problematic because there are mental processes that are without outward behaviour. People may be angry or thinking without doing or saying anything (ibid.). For Armstrong, 'thought' itself is a mental process that is not waiting to be activated or fulfilled, "When I think, but my thoughts do not issue in any action, it seems as obvious as anything is obvious that there is something actually going on in me that constitutes my thought, it is not simply that I would speak or act if some conditions are unfulfilled" (ibid.: 139). Thus for him, behaviourism is an unsatisfactory account of the nature of the mind (ibid.).

However, on Armstrong's view, the behaviourist is not completely wrong. She is wrong in identifying mental occurrences with behaviour (ibid.); however, she is right in asserting that the nature of mind is logically tied to behaviour. For Armstrong, mind or thought is not behaviour itself or disposition to behave, but the inner causes of behaviour. The "thought is not speech under suitable circumstances, rather, it is something within the person that in suitable circumstances, brings about a speech" (ibid.: 140). Armstrong holds that the verdict of modern science is that mental states are nothing but the causes of behaviour and therefore, we can identify them with the physical states of the central nervous system. The mind is the actual physical state of the person who has dispositions, states that have actual causal power to bring about behaviour in suitable circumstances. A mental state is "a state of the person apt for

bringing about a certain sort of behaviour and a state of the person apt for being brought about by a certain sort stimulus" (Armstrong, 1968: 82). There is a causal relation between mental states and physical states in affecting or causing behaviour. We can explain action by appealing to its mental causes. Thus, Armstrong presents a materialist causal theory of the mind, which has also some functionalist features, as will become clear below.

David Lewis (1983:100), in his turn, was upfront in his identification of mental states or events with the physical, and also in noting the causality of the mental states.

> My argument is this: The definitive characteristic of any (sort of) experience as such is its causal role, its syndrome of most typical causes and effects. But we materialists believe that these causal roles which belong by analytic necessity to experiences belong in fact to certain physical states. Since these physical states possess the definitive character of experiences, they must be experiences.

Lewis seems also to be laying the foundation of a functionalist theory of mind. This becomes explicit in his (1970) paper "*How to Define Theoretical terms*":

> Folk psychology contains words such as 'sensation', 'perceive', 'belief', 'desire' 'emotion', etc. which we recognise as psychological. Words for colours, smells, sounds, tastes and so on also occur. One can regard common sense platitudes counting both these sorts of words as constituting a theory and we can take them as theoretical terms of common sense psychology and thus as denoting whatever entities or sorts of entities uniquely realise the theory. Then if certain neural states do so too (as we believe) then mental states must be these neural states.

Thus, the causal theories of Armstrong and Lewis are closely related to functionalist theory of mind as we shall see later in this chapter.

Another reductive physicalist theory of mind, closely related to central state materialism, is the identity theory.

4.2. The Identity Theory

In the twentieth century the identity theory of mind occupies an important place in the philosophy of mind. Philosophers like U. T. Place (1956), J. J. C. Smart (1959), H. Feigl (1958) and others developed this theory. The central theme of this theory is that mental events and physical events are identical and it came as a reaction to behaviourism. The early identity theory is a type identity theory. Criticisms of it lead to focus on a token version.

Token-token materialism (physicalism) identifies a particular mental state token or individual thought to a particular brain state token or a particular physical-chemical neurophysiological object or event. For example, my feeling of pain at this particular time is reducible to or identical with the chemical reaction currently occurring in my brain or nervous system in a particular neural path, the current C. fibers firing (Dale Jacquette, 2009: 57). Thus an instance of a mental state (pain) is also an instance of a physical state (C. fibers firing). Type-type identity theory (type physicalism), on the other hand, holds that pain in general is theoretically reducible to the C. fibers firing, a neurophysiological event (ibid.). Type identity is a theory about mental universals, or properties, while token identity is a theory about mental particulars. Type identity theory does not entail token identity:

> Because if mental kinds themselves are physical kinds, then a particular instance of a mental kind will also be a particular instance of a physical kind. But in no way the former is identical to the latter because a concrete particular that belongs to both a mental kind and a physical kind is a contingent fact (Nath, 2004: 2).

Type-type identity is more restrictive than token-token identity, because it allows only human beings to be capable of feeling pains because they are the only organisms possessing C. fibers. However, token-token identity does allow organisms other than human beings to feel pain, because pain is not limited to the reality of C. fibers. It allows a particular token of a mental state to be realized in many different tokens of physical states (ibid.). "Token physicalism does not rule out the logical possibility of machines and disembodied spirits having mental properties. Type physicalism dismisses this possibility because neither machines nor disembodied spirits have neurons**"** (Fodor, 1981: 4).

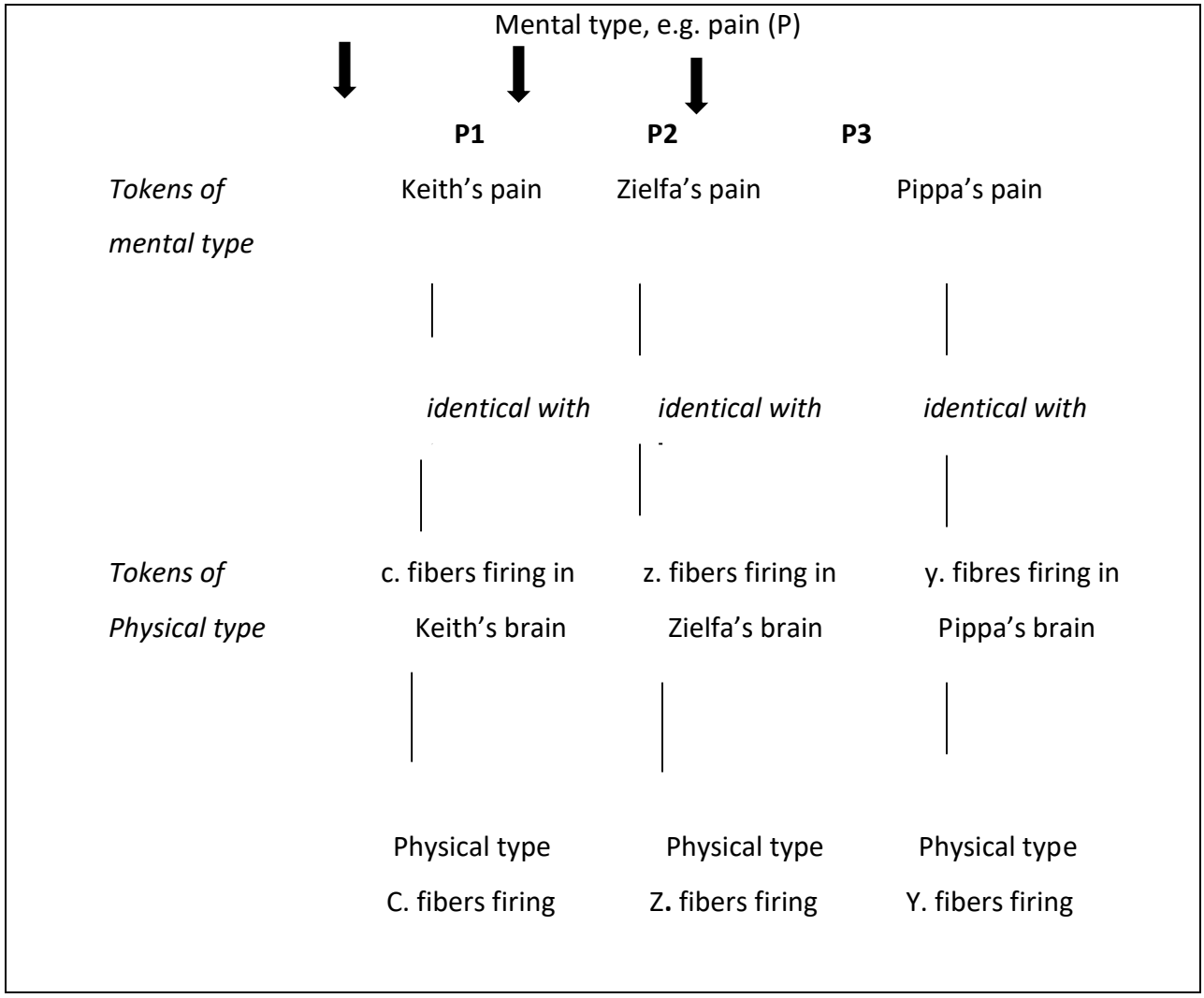
The illustrations below (Maslin, 2007: 69-72) will help us to distinguish token-token identity from type-type identity

Type-type identity (ibid.: 69):

Type of phenomena	Water	lightning	Pain
is identical with			
Type of phenomena	H_2O	pattern of electric charges	C. fibers firing

In the above diagram, type water is identical with H_20, while lightning is identical with the pattern of electric charges, and pain, a mental state, is identical with C. fibers firing, a physical state.

Token-token Identity (ibid.: 72):



According to token-token identity, Keith's pain is a token of the mental type 'pain' that is identical with Keith's brain's C. fibers firing, which is a token of the physical type "C. fibers firing".

Having defined the identity theory, a version of physicalism (type identity as a version of reductive physicalism and token identity as a version of non-reductive physicalism), let us look now look in more detail at the arguments of the proponents of the identity theory for arguing their position that the mind is identical with the physical, the brain or brain processes.

U. T. Place (2000), in the article, "*Is Consciousness a Brain Process*?" argues for the truth of the identity theory but he does not argue simply that consciousness (a state of being aware of an external object or oneself) is identical with a brain process. For him, the thesis that

consciousness is identical with a brain process is a valid scientific hypothesis. For Place, the statement "Consciousness is a process in the brain" is neither necessarily true nor necessarily false, but rather it is a reasonable scientific hypothesis in the same way as the statement "Lightning is a motion of electric discharges" is a sound scientific hypothesis (ibid.: 78). For Place, the argument that consciousness is not a brain process is based on a lack of understanding of the difference between the "is" of definition and the "is" of composition (ibid.:79). For him, the statements containing the "is" of definition are true by definition while those containing the "is" of composition are true by observable verification (ibid.). Any statement containing "is" to indicate a definition entails a relationship between the subject and the predicate. For example, a square is an equilateral triangle or the colour is red. If something is a square, it must be an equilateral angle and if something is red, it must have a colour (ibid.).

However, this relationship is not obtained in statements containing the "is" of composition. For example, in the statement, "His table is an old packaging case", there is no relationship between "His table" and an "old packaging case". It just happens that both expressions characterize one thing namely a table. For Place, those who argue against the statement "Consciousness is a brain process" are using the "is" of definition instead of the "is" of composition (ibid.). This is the case because the meaning of the statement about consciousness and the meaning of the statement about brain processes are different and yet they do refer to one and the same thing. In other words, the "is" of composition allows that one can be in pain without any knowledge of what is going on inside one's brain or be in possession of a table without possessing an old packaging case (ibid.).

For Place, in other words, consciousness and brain processes are two different sets of observations of one event, possibly made at different points in time, just like the cloud and water particles in the sky are two sets of observations of a single event made at different times. More so, in verification of consciousness and brain processes, different methods are applied. Brain processes are observed empirically while consciousness or mental events are observed introspectionally. The operations used to determine the nature of consciousness are radically different from the ones used to determine the nature of one's brain processes (ibid.: 81). The

point Place is making is that even though the terms have different meanings and associations, they can nevertheless refer to one and the same thing. In other words, the arguments that their meanings are different is not an argument for dualism.

In critique, the question can be asked has it been proven scientifically or is there any scientific theory that supports a layperson's observation that consciousness is a brain process? Is there a causal connection between brain activity and consciousness, as there is a causal connection between motion of the electric discharge and an average person's observation of lightening? It is good to remember here that Ned Block (2007) distinguishes between access consciousness and phenomenal consciousness. According to him, phenomenal consciousness is of a qualitative nature; what it is like to have an experience. It is about qualia. Access consciousness relates to the causal relationship among mental states in producing behaviour. It plays a functional role. It does not possess any phenomenal characteristics. There might be a scientific theory that concurs with the theory about access consciousness, as it is objective in nature in its functional role. However, in relation to phenomenal consciousness, there is no scientific theory yet about it. Presently, it defies science as it is subjective in its operation, not open to external observation. I think that Place does not take these two different aspects of consciousness into consideration in his argument. One does not know whether he is referring to the meaning of access consciousness or phenomenal consciousness. What form of consciousness can refer to brain processes?

More so, Place has likened consciousness and brain processes to lightning and motion of electric discharges. But he is comparing examples that are not of the same nature. In terms of consciousness and brain processes, one expression refers to the mental and immaterial, and the other refers to the physical and material. On the other, in the case of lightning and electric discharges, both expressions refer to the physical and material. Place may be here comparing two pairs of concepts that are different in nature.

However one has to acknowledge that Place claims that the assertion made by some people that consciousness is not a brain process is due to a logical mistake, which he calls the phenomenological fallacy (ibid.: 82). The phenomenological fallacy is the mistaken notion that

someone's description of her experience of phenomena is the description of her experience or awareness of the properties of phenomena as she experiences them on an internal cinema screen inside her brain known as a phenomenal field. Thus, when the phenomenological fallacy is committed, when she reports a 'green after image', she will be viewed to be reporting the greenness of an internal object inside of her brain, instead of the greenness of the external object in her environment that causes the internal greenness. The issue is that this internal greenness is beyond physical observation as it is completely independent of the physical object outside (ibid.).

For Place, the phenomenological fallacy occurs because we think that our ability to describe objects in our environment depends on our awareness of their phenomenal properties and not on their external properties. We describe our experience instead of the external object. That is the problem of perception. Through the ages from the Lockean (taken over from Galileo) distinction between primary and secondary qualities to Thomas Nagel's (1987) exposition of possible replies to the question "How can we know anything?" and beyond to scepticism about the external world, we have been confronted by this problem. For Place, this mistake can only be corrected when we understand that it is the perceptible properties of an object outside us that are responsible for the phenomenal experience we have inside us, which we normally describe. "It is only after we have learned to describe things in our environment that we can learn to describe our consciousness of them" (ibid.: 83). Thus, Place argues that if brain processes such as sensation and perception give rise to consciousness, then consciousness is indeed a brain process.

> In other words, when we describe the after image as green, we are not saying that there is something, this after image, which is green, we are saying that we are having the sort of experience which we normally have when, and which we have learned to describe as, looking at a green patch of light (ibid.).

Thus, the green after image, which we are normally conscious of, is caused and produced by the greenness of the external green object in our environment. Place holds that once we rid

ourselves of the phenomenological fallacy, we realise that every conscious experience or introspective observation is explainable in terms of brain processes (the activities of the sense organs and the brain). Thus, mental description and physical description are just one description of the same object.

In turn, J. J. C. Smart (2000), in his paper, "*Sensations and Brain Process"*, addresses some of the objections raised against U. T. Place's paper (2000), "*Is Consciousness a Brain Process"*, with the aim of confirming the type-type identity theory. For Smart, there are no acceptable philosophical arguments in favour of dualism (Smart, 2000: 85).

Smart laments the fact that physical science can explain all the universal phenomena except consciousness and sensations. However, he holds that these phenomena should be subjects of scientific inquiry. Sensations and consciousness do not correlate with the brain. To say that consciousness and sensations are correlates of the brain implies that they are something over and above brain processes. Hence, they would be then what Herbert Feigl called "nomological danglers" that is, it would imply that sensations and consciousness operate outside the laws of physics that govern other realities (ibid.: 86). Rather, for Smart, the report of having an 'after image' or an 'ache' is a report of a process that happens to be a brain process (ibid.: 87). Thus, if a report of sensation "is a report of something, that something is a brain process. Sensations are nothing over and above brain processes" (ibid.). For Smart, the "is" of the statement "Sensation is a brain process" is the "is" in the sense of strict identity. It is similar to the identity between the number "7" and the smallest prime number greater than "5". It is not a kind of identity of a mere spatial-temporal continuity such as the statement that Janet is the same woman who as a little girl did her primary education fifteen years ago (ibid.: 88).

His responses to the criticisms against Place's paper form part of what is known today as the type-type identity theory of mind. As a criticism against materialism (physicalism), it is argued that an illiterate person is able to talk about her after image, pain, taste or smell without any knowledge about neurophysiology, and thus the mental is not identical with the physical (ibid.). In response, Smart offers a different example by referring to Venus appearing

both in the morning and in the evening time. Hence, the term 'Morning Star' is identical with the term 'Evening Star'. However, one can know about the Morning Star but be ignorant of the Evening Star, because she is asleep at the time when the Evening Star normally appears. She knows and talks about the Morning Star without knowing that the Evening Star is also the Morning Star. That she can talk about the Morning Star without any knowledge of the Evening Star does not prove that the expression the 'Morning Star' is not the same as the expression 'Evening Star' (ibid.). Hence, in the same way, 'mental' and 'physical' are two different ways of talking about one thing.

Another objection to type-type identity theory is that an after image or pain is not spatio-temporal while a brain process is. Hence, after images or pains cannot be identical with brain processes. Smart calls this kind of reasoning *ignoratio elenchi* (ibid.: 91). Smart does not argue "that an after image is a brain process, but rather that the *experience* of having an after image [or pain] is a brain process" (ibid.). A reported experience is a brain process. There is no such thing as an after image or a sense datum as such, however, there is such a thing as the experience of having an after image (ibid.).This particular experience is described in relation to material objects because there is no such thing as a phenomenal language. For instance, we see a yellow-orange patch on the wall paper, but not a yellow-orange image *per se.* Trees and walls can be green in colour, but not the experience of seeing or imagining a tree or wall paper (ibid.).There is no greenness without an object.

For Smart then, U. T. Place's position that the identity theory is a scientific hypothesis is partly right and partly wrong. Place would have been completely right if the thesis is a statement about identity between a brain process and a kidney process, which is a completely empirical investigation. However, if the thesis is between a brain process or kidney process and mental process, it becomes a non-empirical topic and cannot be solved by an experimental test. However, according to Smart, the principle of parsimony, Ockham's razor and evolutionary theory (material objects existed first before the mind) will always favour materialism (physicalism), because dualism will leave us with entities that are not subject to physical laws.

Smart therefore holds that sensation is a brain process, and the talk about sensations is talk about brain processes, and the logic of a sensation statement is the logic of a brain process statement. This will imply that there is no inherent difference between mental and brain processes. I, on the other hand, will argue that sensations must have properties that make them mental; and these properties subsequently make their logic different from the logic of brain processes.

> The problem is that, no matter how complete the physical description of the person, or no matter how complete an explanation is in terms of a person's physical constituents, all mention of the mental is omitted. The mental is not 'captured' by any physical theory. For example, it is possible to describe in physiological detail light waves contacting a retina and the transmission of an electrical impulse to the brain-cortex, where there are links in a causal chain. But the result of this causal chain is something qualitatively different: a visual perception, an experience.... (Priest, 1991: 112)

To prove that statements about mental properties are statements about physical properties, the identity theorists may need a further argument. They need to convince that the experience as the result of the causal chain is physical.

> Even if it is true that the mental is the physical, it would seem to require further argument to show that the mental does not exist over and above the physical. After all, the mental must exist in some sense to be identified with the physical. Arguably, a person has irreducible mental properties – thoughts, moods and emotions – even if these are not properties of an immaterial substance.... (ibid.: 114)

4.3. Strengths and Critique of the Identity Theory

Brain state theory or identity theory or central state materialism has many advantages over dualism and behaviourism that make it an attractive theory of the mind. Maslin (2007: 73) enumerated those advantages as follows: The claim is that mental states are just brain states. Other than in the case of dualism that professes the existence of both mental and material substances and properties, here the focus is on one kind of substance or property only.

In addition, as mental states are brain states, there is no problem of causal interaction, while this is a big issue for dualists. Physical causation is among the physical events taking place in the brain and central nervous system. There is no problem of spatial location and energy conservation that was a huge problem for Cartesian dualists. It is the nature of causation that the cause and its effect must share the same spatial location or stand in close contact to each other: a ball does not move unless something pushes against it. In physical causation, both cause and effect are contiguous to each other as they share the same spatial location in the universe, unlike in the case of dualist mental causation (causation between mind and body).

Also, identity theory answers the question why an injury in the physical brain can affect the mental function quite adequately by simply stating the fact that the mind is the brain. On this theory, it becomes easy for science to study the mind. The study of the mind is now the same as the study of the brain, which consists in neurophysiology and neurochemistry. Psychology now seems to share the same methods and standards with all natural sciences. Physicalism thus gives us a physical origin for every substance. A fertilised ovum is completely an organization of molecules. There is no room to entertain the thought that along the line of development, any nonphysical substance just emerged. Hence, there is no nonphysical substance to account for. Physicalism shows that science, in the end, might explain all dualism stands for (ibid.: 22).

In terms of mental causation, it is not a problem therefore for the identity theory of the reductive type to prove mental causation in that the causal efficacy of the mental is dependent on the causal efficacy of the physical properties as they are reducible to the latter. However, mental properties do not *qua* mental properties cause things.

> There is no causal influence of the mental beyond that of the causal influence exerted by the brain. Type identity theory is the view that types of mental processes, like pain, simply are types of physical processes like C.fibers firing. It denies that mental causal efficacy is anything other than the causal efficacy of neurophysiological processes. The reductive element of the view means that mental causation is a placeholder for the real causal story which inevitably involves brain processes, microphysical particle states, or some other straightforwardly physical cause … The mental is causally efficacious only insofar as it reduces to the physical, and the physical is causally efficacious. Reductive physicalists allow that mental terms are a pragmatic convenience and don't need to be eliminated, so long as they are understood as a kind of sort-hand for the 'real' causal story ... This means that reductive physicalism is less strong than eliminative physicalism, which holds that the mental will, or at least should, be replaced entirely. (Andersen, 2009: 4)

In spite of the advantages of identity theory, it has attracted many criticisms. We are going to look at some of the criticisms, which in turn show the inadequacy of physicalism to solve the mind-body problem. If identity theories were adequate, the causal efficacy of mental phenomena would be usurped by physical properties. As we will see here, there are philosophers that have argued against the correctness or adequacy of the identity theories**.**

Saul Kripke (2002), in the article, "*Naming and Necessity",* argues against different versions of the identity theory, namely *token-token* identity and *type-type* identity. To remind the reader, token-token identity theory holds that the token or particular mental state is identical with a token or a particular physical brain state. For instance, my pain at 7 am in the morning is identical with my C. fibers firing at that hour. The type-type identity theory holds that my mental states of a certain type are identical with corresponding physical states of a

certain type. For instance, pain is the stimulation of C. fibers. It means that there cannot be a mental state of a certain type without a brain state of that type or vice versa (Kripke, 2002: 329).

In terms of token-token identity, Kripke argues that a mental state cannot be identical with a brain state, because, if mental states are identical to brain states, their identification is necessarily true (it is impossible to conceive one without the other), but since we can conceive of mental states independently of physical states, they cannot be identical. Descartes, through his, 'methodic doubt' showed the possibility of conceiving the mind independently of the body when he argued for *cogito ergo sum* (I think, therefore, I exist) (ibid.).

More so, analytical examination of the identity statement (a mental state is a physical state) shows that the identity theory is false. Kripke, in order to prove his position, applies the terms 'rigid designator' and 'non-rigid designator'. A rigid designator represents proper names such as 'Descartes' and descriptions contain a proper name such as 'Descartes' body'. A rigid designator refers to the same thing/entity in all possible worlds (ibid.). Descartes' body can only be Descartes' body (ibid.). In contrast, a non-rigid designator refers to different individuals in different possible worlds, and it is used for describing different entities. This does not hold in terms of proper names. But, consider for instance the sentence, 'the first Postmaster General is the inventor of bifocals'. Another person other than the first Postmaster General could have been the inventor of bifocals in other possible world (ibid.).

For Kripke, the identity statement between a mental state and a physical state is a necessary one because the terms 'mental state' and 'physical state' are rigid designators. They refer to the same thing in all possible worlds. It is impossible to imagine that pain could exist without being pain; or C. fibers exist without being C. fibers (ibid.: 330). This is contrary to the token-token identity theorist position that takes identity between a mental state and a physical state to be a contingent one, as in the case of the identity between the first Postmaster General and the inventor of bifocals (ibid.). As it is a contingent activity of the Postmaster General that made him into the inventor of bifocals, it seems that for them also it is a contingent property of

the brain to cause C. fibers to fire that turns this firing into a pain (ibid.). Pain is a non-physical property of the brain; but this non-physical property must be stated in physical terms.

Kripke argues, in contrast, that being a pain state is an essential property of pain; and even being a pain state of a specific type is an essential property of pain (ibid.: 331). The identification between mental state (pain) and brain state (C. fibers) is analogous to the scientific identification between heat and molecular motion or water and H_2O. Kripke argues that if both 'heat' and 'molecular motion' are rigid designators, then the identification between them is a necessary one, and so must it be for the terms 'mental state' and 'brain state'. But if their identification is a necessary one, then token-token identity theorists would be committed to the view that there could not be a C. fiber firing (stimulation), which was not a pain; nor a pain, which was not a C. fiber firing (stimulation) (ibid.). After all, the terms 'mental state' and 'brain state' are rigid designators because they refer to the same thing in all possible worlds (ibid.). Of course, one needs to buy into the argument that 'brain state' and 'mental state' are rigid designators for this critique to hold against identity theorists.

Eliminative materialists such as Richard Rorty and Paul Churchland have also criticized the identity theory of mind. Eliminative materialism is quite different from other physicalist theories of mind in that it is not reductive, in other words, it does not reduce mental notions to neurophysiological notions, but rather it claims that mental states and processes do not exist. It blames folk psychology (common sense conceptions of mental notions) for postulating falsely the non-existent properties such as phenomenal properties or states (this almost reminds of Place's phenomenological fallacy, although he did not view himself as an eliminativist). This theory holds that in the future neuro-physiological science will explain away phenomenal properties such as qualia or sensations and intentional states such as belief, desires and other mental notions, just like science has dealt with previous theories concerning witches and demons, which initially formed part of everyday belief. Churchland (2000:130) writes:

> Psychosis is a fairly common affliction among humans, and in earlier centuries its victims were standardly seen as cases of demonic possession, as instances of Satan's spirit itself, glaring

> malevolently out at us from behind the victims' eyes. That witches exist was not a matter of controversy. One would occasionally see them, in any city or hamlet, engaged in incoherent, paranoid, or even murderous behaviour. ... Modern theories of mental dysfunction led to the elimination of witches from our serious ontology. ... The concepts of folk psychology – belief, desires, fear, sensation, pain, joy and so on –await a similar fate, according to the view at issue. And when neuroscience has matured to the point where poverty of our current conceptions is apparent to everyone, and the superiority of the new framework is established, we shall then be able to set about *re*conceiving our internal states as activities within a truly adequate conceptual framework at last...

Churchland gives three reasons for his stance on eliminative materialism, which is based on his conviction that folk psychology is an out-dated and confused conception of our psychological activities. First, he argues that folk psychology has failed to explain most of things that are very familiar to us. Folk psychology has failed in its explanation of phenomena such as sleep, intelligence, memory and mental illness, while in contrast, neuroscience, in a short period of time, has explained these phenomena in some other naturalistic way (ibid.: 132).

The second reason he gives for folk psychology being out-dated is based on the history of scientific achievements. Science has made tremendous progress in explaining facts that folk theories could not explain previously. The folk theories of heavenly activities, the nature of fire, and the nature of life have given way to more improved scientific investigations and theories. He claims that while folk psychology has survived to this day, nevertheless, it has begun to feel the pressure of neuroscience, which will overthrow it one day. Thus, the phenomenon of conscious intelligence is not intrinsically a different phenomenon from other natural activities but is simply perhaps more complex a concept than natural activities (ibid.).

The third reason Churchland gives for the redundancy of folk psychology is that eliminative materialism is more plausible than identity theory and functionalism. He claims that eliminative materialism, unlike identity theory and functionalism, will not argue for parallels of folk psychological entities with physical ones in a matured neuroscience. The reductive theories must match folk psychological entities with physical ones; identity theory argues for a type-type match, and functionalism argues for a species-specific match. Churchland claims that finding a match will be more demanding in a matured science for the new theory must entail finding physical entities that match the mental ones but there is no prospect of a smooth match-up between mental and neuro-scientific terms. Eliminative materialism will not demand that from the new science. Hence, eliminative materialism will be more successful than the reductive theories of mind (ibid.: 132).

Eliminative materialism has been criticized as going against our common and general held view that we, humans, inevitably have minds and we do experience pains, sensations and other mental notions and that these mental phenomena do cause us to do things. For instance, I go to see the see the doctor because of the throbbing pain of my headache because I believe that some painkiller tablets will help in removing the pain. However, eliminativists may object that commonly held views are often mistaken or wrong. They buttress their point by arguing that there was a time people held the view that the sun rotates around the earth but today science has proven the opposite – that it is the earth that travels around the sun. Thus, dualist notions could simply be proven false with a matured neuroscience (ibid.: 134).

The opponents of eliminativism may enquire why, if mental states are not real; have they been part of our lives, and made our everyday life and talk meaningful and coherent. We, in our ordinary everyday capacity, believe that mental states have causal power in causing our behaviour and action. The assumptions that there are mental states and entities have been in existence before the wrong assumption of the movement of the sun and they still exist, after the correction of this assumption. Therefore, perhaps the eliminativists need stronger arguments to refute the reality of mental states.

Some philosophers, such as Jerry Fodor, who believes in the efficacy of folk psychology (people's common sense of understanding the mind), argue that eliminative materialism exaggerates the defects of folk psychology and undermine its success story. Folk psychology has aided people to communicate effectively among themselves in that they see themselves having the same mental states and phenomena. While the proponents of folk psychology might argue that the arrival of matured neuroscience might call for certain adjustments in the framework of folk psychology, however, its total elimination might be impossible as it can account for the subjective feeling or qualia of having sensations like pain (ibid.). Identity theory, for instance, has been criticised for its inadequacy in accounting for consciousness or qualia (the individual's conscious experience), e.g. the pain of a headache or the bitterness of lemon**.** I will return to the arguments for this position at the end of this chapter because a belief in the reality of qualia or consciousness faults all reductive physicalist theories of mind.

Let us now turn to non-reductive physicalism and mental causation.

5. Non-reductive Physicalism and Mental Causation

Non-reductive physicalism appreciates the independent causal efficacy of the mental by anchoring it on the physical yet does not reduce the mental to the physical. This is the view that mental properties are a different ontological class of properties from the class of physical properties, and that mental events cannot be reduced to physical events, although the physical is somehow prior to the mental. In this sense, this view acknowledges that the mental supervenes or emerges from the physical. In this way, it creates a kind of midway between Cartesian dualism and physicalism. I will consider its response to the issue of mental causation and argue (mostly in the next chapter, although I will touch on the argument here) that in the form of property dualism it may in some respects be similar to Aristotle's theory of hylomorphism, a theory that every living thing and non-living thing in nature are explained from the principles of matter and form.

Non-reductive physicalism is a theory of mind that holds that both mental and physical events are physical. In other words, it professes monism. However, it differentiates the mental

from the physical in terms of properties. There is only one substance, the physical, but it has both mental and physical properties, and so mental properties are different from physical properties.

Diana Mertz Hsieh (2003:1), in her article, "*Mental Causation through Constitution"***,** referring to Jaegwon Kim's philosophy of physicalism, notes that non-reductive materialism has three main theses, namely physical monism, the non-identity of the mental and the physical, and strong supervenience. Physical monism is the thesis that holds that everything real is physical and every physical effect must have a sufficient physical cause. Nevertheless, if combined with a weak property reductive physicalism, then it allows for mental properties existing but dependent either via supervenience, or via emergence upon physical properties. These mental properties are supposed to do real causal work within non-reductive physicalism, as non-reductive physicalism is not generally or necessarily an epiphenomenalism (ibid).

On the non-identity of the mental and the physical, non-reductive physicalists deny that the mental is identical to the physical. This view is based on the holistic character of mental. This is the view that a particular mental term normally begets another mental term. For instance, the desire to drink water will make me want for a cup, and wish that I will get a clean water. In this case, the mental term 'desire' begets another mental terms 'want' and 'wish'. The view is also based on the constitution force of rationality. As mental terms are logically connected to one another or begets the other, they are put under examination of the force of rationality and coherence. Human beings expect each other to be rational and coherent in conversation and they do this by trying to understand the meaning of words they use in discussion and whereby a meaning of word conflicts with other, they choose the right the word that conveys the true meaning of what they want to say for rational and coherent conversation. Finally, the view implies a multiple realization argument. This is the argument that the mental and the physical cannot be strictly identical if the same mental properties may be realized in a wide variety of physical systems as Hilary Putnam (2002) for instance argues.

The tenets of the non-identification of the mental with the physical, multiple realization, holism and rationality or charity result in functionalism. On this view, mental states are

identified with abstract relational (functional) states rather than particular intrinsic physical states – and thus the nature of a mental state is constituted by its relations to other states and to inputs and outputs (Block 1995: 324) – while the functional states are taken to supervene on physical states.

The non-identification of mental states with physical states argues for the irreducibility and autonomy of mental states. This autonomy of the mental is the basis for the non-reductive commitment to a compatibilist explanation under which every event or behaviour may be explained as the product of either physical or mental causes, each of which is sufficient to explain behaviour, but neither of which is reducible to others (Hsieh, 2003: 3). However, someone like Kim argues that non-reductive physicalism does not imply that mental properties are causally efficacious. We will see the argument in the next chapter.

The third thesis of non-reductive physicalism, as described above in Hsieh's terms, is so called strong supervenience. I shall discuss the notion of supervenience later on in this chapter in more detail, but for now let me just state that broadly, the notion holds that any mental difference or change implies an underlying physical difference or change. The principal difficulty faced by non-reductive physicalists is to explain mental causation, especially mental to physical causation (in relation to supervenience) (ibid.). We shall discuss this further in the next chapter.

5.1. Property Dualism

Property dualism has no place for the notion of immaterial soul substances, however it holds that the brain, a composite physical thing, possesses two fundamentally different sorts of features or properties; mental properties that are non-physical in nature, as well as familiar physical properties such as mass (Maslin, 2007:31). Thus, there is one substance and it is physical, but this substance has certain properties that are mental, thus non-physical, that cannot be reduced to physical properties. Property dualism in other words holds that:

> While there is only one kind of (physical) substance, the mind not only has physical or behavioural-material-functional

> properties but also non-physical behaviourally-materially-functionally ineliminable and irreducible properties. This kind of theory is usually referred to as property dualism, though it is known in related terms as a dual aspect or attribute, anomalous monism (Donald Davidson) or non-reductive materialism (Joseph Margoles), as opposed to Cartesian ontic or substance dualism. (Jacquette, 2009: 22)

Property dualism comes in various forms such as emergentism, epiphenomenalism, double aspect theory, anomalous monism and panpsychism.

5.2. Emergentism

Property dualism is sometimes described as emergentism because it allows for the evolutionary emergence of mind from a complex physical substance. The terms 'emerge and supervene' are sometimes used interchangeably to express this notion. This view holds that non-physical properties may emerge from, or supervene on, a material substance, if they achieve a certain complexity in the same manner in which life emerges evolutionarily from certain kinds of properly organized material substances (ibid.: 23).

In terms of emergence, simple elements gather to form more complex units, such that those units appear to have qualities that were not there in the simpler parts. Thus, chemicals take on qualities that are missing in atoms and cells take on qualities that are not in the constituent chemicals. The further up the chain of complexity one goes, the more qualities emerge that were not seen lower down. Some philosophers, such as John Mill (1843) and Jaegwon Kim (1999) hold that consciousness is an emergent property for it emerges when life develops to a certain level of complexity. A conscious human being can think, choose and respond creatively to situations, experience emotions and decide how to act, but the various organs that make up our bodies do not have qualities or abilities. Consciousness, for them, emerges once life reaches a certain level of complexity. Thus, consciousness is an emergent property of the brain (Thompson, 2012: 22).

5.3. Epiphenomenalism

Epiphenomenalism is a theory that states that mental phenomena or properties have no causal power, thus they are 'epiphenomena', that is 'above phenomena'. Mental properties exist outside a causally complete physical world, and thus physical properties can cause mental properties, but not the other way around. In these terms, it is implied that consciousness is a by-product of the brain and nervous system's activity that has no causal power of its own (Jacquette, 2009: 19).

Consequently, epiphenomenalism does not make room for human creative activity. It takes humans to be mechanical, and governed and ruled by physical brain activity. There are no psychological effects on our daily life (Thompson, 2012: 22) on this view. Kim argues strongly that mental properties are epiphenomenal as they are dependent on physical properties. And in a universe that is fundamentally physical, they are causally inefficacious as every event has one complete cause that is physical. I will come back to Kim on epiphenomenalism in Chapter 3.

5.4. The Double Aspect Theory

This theory, based on the philosophy of Baruch Spinoza, holds that God and nature are the same and that all reality has both mental and physical aspects to it. It holds that there is a single substance, which is both conscious and extended, i.e., that has both mental (not scientifically examinable) and physical (scientifically examinable) aspects. The mind and body are inseparable and in effect are two ways of seeing the same thing. Everything we experience as mental also has a physical aspect. Neural events and conscious experience are simply two ways of explaining the same thing (ibid.: 23).

6. Functionalism and Mental Causation

6.1. Functionalism

Let us now focus attention on and consider in more detail one of the most influential forms of non-reductive physicalism, namely functionalism. Hilary Putnam (2002) in his article, "*The Nature of Mental States"*, examines the identity theory with a particular reference to reductive

physicalism, using as example the question whether pain is a brain state. He argues that the identity theory of mind and behaviourism face many challenges in solving the mind-body problem, an important one of which is multiple realizability. This is the problem that the same mental state may be physically realized in a number of different forms, e.g. there are many different creatures who have the mental state of being in pain but their 'pain' is correlated with different physical states of their various nervous systems, and thus there cannot be any type-type identity between pain and C. fibers, as only humans and some animals (fish have very low percentages of C. fibers with sharks having 0%, for instance) have the latter. Consequently, he goes on to propose an alternative theory of mind called functionalism as a solution to the mind-body problem. Functionalism holds essentially that mental states have functional or causal roles in the brain.

Putnam (2002: 74) argues that brain states are somehow connected to mental states. They are not correlated, for if they are correlated it could be said then that the identity between light and electromagnetic radiation is a matter of correlation. The notion of correlation will still leave us with some questions to address such as what is light if it is not the same as the electromagnetic radiation, and what makes light accompany the electromagnetic radiation. For Putnam, these questions would be taken care of if we just accept that light is electromagnetic radiation. Similarly, pain being a brain state forecloses these questions being asked, "What is pain then, if it isn't the same as the brain state?" and "what makes pain accompany a brain state?"(ibid.). However, as mentioned above, the problem of multiple realizability makes it impossible to accept type-type identity, and thus Putnam has to make another plan.

Putnam then introduces his functionalist theory of mind. For him (ibid.), "pain is not a brain state in the sense of a physical-chemical state of the brain or the whole nervous system" (ibid.). For Putnam, "pain or the state of being in pain is a functional state of a whole organism" (ibid.). This functional state of the brain is modelled on a Turing Machine[5] and Probabilistic

[5] A Turing machine is a computing machine proposed by Alan Turing in 1936 in the context of research into the foundations of mathematics (Lisbeth De Moi, 2018: 1). See also Turing 1950.

Automaton. A Turing machine computes various computable tasks in a deterministic way, with certain instructions and a machine table, which specify how various states of the machine interact with one another and to the input. The machine computes deterministically, and its behaviour is produced by the input and resultant machine states, which interact with the input in a specific way producing computations and translating to other states automatically (ibid.). A probabilistic automaton is similar to a Turing machine, except that the transition between states produced by inputs are determined probabilistically rather than deterministically (ibid.). More so, a Turing Machine is a device that can "move sequentially from square to square on unlimited strip of paper squares. The machine scans the information found on each square (in a binary code (*two symbol systems*) consisting of strings 0s and Is), erases it, leaves the square blank or writes a 0 or 1. Computer programs in essence tell the machine what to do as it moves along the strip in prescribed square-by-square steps, dictating precisely and unambiguously what procedures it is to follow in erasing, writing and moving to another square on the strip" (Jacquette, 2009: 86).

Now, by analogy, Putnam's suggestion (2002) is that mental states are simply functional states (they are defined by their roles or works they do) within a system. However, this system in the case of humans is akin to a machine with a probabilistic transfer between various states. For Putnam, the functional organization or the description of a system is a complete account of all its various states and the probabilistic relation with one another and with sensory inputs and motor output (outward behaviour) (ibid.). The various states of the system are causally related in relation to input, the changing of various states from one state S_1 to another S_2 and back to S_1 or a move to S_3 and producing of output. The relationship between mental states is like the relationship between the different various states of a machine, whereby the machine computes by moving from one state to another. A mental state like hunger is identified by an organism being in that state of being hungry, which can move from that state to another mental state of wanting food, or having a desire of going out to find food, which in turn can move to the state of being satiated. This transition from one state to another supports the appropriateness of a functional approach to mental states (ibid.).

The behaviour of the system is comprehended without any knowledge of the physical and material makeup of the system. For Putnam then, mind or mental states can be understood without knowing the particular detail of the physical-chemical brain state or structure, which realises the mind's functions (ibid.). Recall that for the brain state theorist, pain is identifiable with C. fibers firing. Hence, it is only organisms with firing C. fibers that can feel pain. This means that the physical-chemical state of the brain must be the same in all organisms that feel pain. However, as mentioned above, there are organisms like an octopus that feels pain, but it does not have C. fibers.

For Putnam, in other words, brain state theory can only be viable if there is a common physical-chemical structure to all organisms that experience mental states (ibid.: 77), which is not the case. Putnam holds that identifying a functional system which is common to all organisms, is more viable than identifying the same physical-chemical structure in all organisms that feel pain. The fact that we identify organisms as being in pain, hungry, or angry through seemingly similar behavioural actions suggests that having a common functional organisation for all organisms is more realistic than them having a common physical structure. Psychological or mental states should thus be functionally defined rather than by their physical structure or what they are made of (ibid.).

Putnam also compares functionalism with behaviourism. Behaviourism is a theory of mind that states that being in pain is neither being in a brain state nor a functional state but an act of displaying a behaviour or disposition to behave (ibid.). For Putnam, behaviourism has one advantage because its identification of pain as behaviour is in conformity with our ordinary way of identifying pain through behaviour. However, Putnam claims that behaviourism defines pain with the concept of pain. In other words, it is circular in its analysing the concept of pain. For instance, pain is a disposition of X behaving as if X were in pain. In contrast, functionalism explains the concept of pain without employing the notion of pain. Pain, for the functionalist, is a state caused by receiving sensory inputs of damage to the body, which causes one to scream or wince or have the desire to remove the body from the source of sensory inputs or stopping the pain (ibid.: 78). Putnam argues further that behaviourism is inadequate to account for pain

because one can be in pain but decides to show no outward behaviour of screaming or wincing, and instead smile. Alternatively, one may not be in pain but decides to scream or wince, which is a sign for the behaviourist that one is in pain. Hence, pain is not behavioural in either case.

However, note that neither behaviourism nor identity theories can explain the nature of pain itself as it is accompanied by subjective experience, which is beyond the investigation of behaviourism and identity theory. If none can explain the nature of pain, one could wonder how and to what extent a functionalist state is different from the dispositional state of behaviourism or even the brain state of the identity theorist; the pain is either a functional state or brain state or a behavioural disposition. They are all states of the mind. However, Putnam does address this issue when he introduces functionalism as a doctrine that defines mental states such as desire, pain, thought or any other mental state not in terms of their internal constitution, but rather in terms of their function or what mental states do in the brain and central nervous system. The material or physical arrangement of the brain plays no role in the characterisation of a mental state. In terms of functionalism, a mental state is characterised by its causal relation to sensory inputs, other mental states and behaviour (Levine, 2017: 2).

For the functionalist, pain is a state that is caused by bodily injury that makes someone believe that her body is not fine and desire to get rid of that state or avoid it. Any creature that meets this condition of pain is capable of feeling pain. Thus, for the functionalist, pain can be realised by different kinds of physical state in different kinds of creatures. Pain, in other words, can be multiple realised in creatures with or without brains and central nervous systems, and even in a non-biological system such as a computer (ibid.: 2). This implies functionalism focuses on token-token relations and not type-type ones.

6.2. Types of Functionalism

There are various versions of functionalism namely psychophysical functionalism, machine functionalism, analytic functionalism, and role and realizer functionalism.

Psychophysical functionalism

Psychophysical functionalism uses a scientific approach in order to analyse what constitutes it. (Levine, ibid: 2). Physical analysis or speculation alone is not able to reveal the neurophysiological mechanism that makes pain a functional state. Psychophysical functionalism has the function of revealing the causal elements of the functional state. In human beings, the causal elements are located in the processes of the brain and central nervous system. Nonetheless, this is compatible with holding that mental functions are realizable in other physical arrangements. (Maslin, ibid: 125). Examples of psychophysical functionalists are Jerry Fodor and Zeno Pylyshyn.

Machine Functionalism

Machine functionalism is the earliest form of functionalism. It is rooted in the works of Hilary Putnam. As we noted earlier in this chapter, Putnam took the Turing Machine as a model for his explanation of mental states. For Putnam, any creature with mental states can be regarded as a Turing Machine, an idealised finite state digital computer, whose operation is specified by a set of instructions in the form of a machine table or program, with a system like this:

If the machine is in state S_I; and receives an input I_J, it will move to state K and produces an output O_n; and return to S_I for another causal relation between inputs and outputs.

On this model of a Turing Machine, mental states of a creature are identified with “machine table states”. These states are not dispositions to behave, as the behaviourists would believe, rather they are understood in terms of their relations to input, output and other mental states of the machine (Levine, ibid: 5).

Analytic Functionalism (Metaphysical functionalism)

Pain functionally defined, comprises an input in the way of tissue damage, output behaviour in the form of groaning, wincing and in relation to other mental states. As a mental state, the desire to be rid of the pain will also need its own functional analysis or specification. In analytic functionalism, the mind is seen as a black box, mediating between input and output but opaque as regards to what actually takes place within it. It does its job all right, but how it does or what

it is made of does not matter. Thus, the function of the mental is never identified with the system that instantiates it (Maslin, ibid: 125).

Analytic functionalism is about the process of understanding the mental states *a priori*. Mental states are analysed independently of the body. Inputs to mental states are either physical or non-physical in nature. It provides a neutral platform or non-personal ground for analysing and understanding our mental states (Atkins, 2014: 2). Proponents of analytic functionalism are David Lewis and Wolfgang Scharz.

Role and Realizer Functionalism

Functionalism comes in both non-reductive and reductive varieties. The non-reductive functionalists are the role functionalists while the reductive functionalists are the realizer functionalists. Role functionalism is associated with Putnam while realizer functionalism is the version promoted by David Lewis. However, both role and realizer functionalists take mental state-types to be in some sense associated with causal roles (Bennett, 2007: 323). For them, pain will be associated with the causal role of being "typically caused by tissue damage, and typically causing aversive reactions" (ibid.).

While the realizer functionalists identify the mental state-type pain with whatever occupies this causal role, the role functionalists identify it with the role itself (ibid.). Role functionalism holds "that mental-state types are second-order, relational properties. Pain is the property of having some property or other that meets a certain functional specification, or plays the pain role" (ibid.). For realizer functionalism, "pain is the first-order property that actually plays the role in some population – in humans, for example, the pain might be C.fibers stimulation" (ibid.). Role functionalists take mental predicates like 'pain' to rigidly designate second-order functional properties, while realizer functionalists take them to non-rigidly designate first-order physical properties… "(ibid.).

Role functionalists hold that pain is a higher-level relational property. Realizer functionalists argue that functionalism as theory aims at providing a concrete description of the lower-level properties that satisfy the functional characterizations. Realizer functionalism is also

known as the "functional specification theories". On this view, the pain in humans is the C. fibers stimulation that plays the causal role of pain in human beings (Levin, Janet 2018: 8). Realizer functionalism, the kind of functionalism, Kim professes, is in a better position than role functionalism to explain the causal efficacy of the mental, since mental properties are functionally reducible to physical properties. For instance, "if I stub my toe and wince, we believe that my toe stubbing causes my pain, which in turn causes my wincing" (ibid.). Some philosophers have argued (Malcom 1968; Kim 1989, 1998) that if pain is realized in me by some neural event-type, then in so far as there are purely physical law-like generalizations linking events of that type with wincing, one can give a complete causal explanation of my wincing by citing the occurrence of that neural event (and the properties by virtue of which it figures in those laws). "And thus it seems that the higher-level role properties of that event are causally irrelevant. This is known as the "causal exclusion problem", which is claimed to arise not just for functional role properties, but for dispositional properties in general" (Levin, 2018:9). Kim is credited with the proposition of the exclusion problem. It does mean that Kim is a realist functionalist. While Davidson being a non-reductivist, is an anti-realist functionalist, in other words, he professes role functionalism in his anomalous monism.

7. Davidson's Theory of Action and Anomalous Monism

Let us now consider a final specific theory in the family of non-reductive physicalism, which is core to my focus on Kim and Aristotle in later chapters, namely Davidson's anomalous monism. One of the core theses of anomalous monism is mental anomalism, according to which mental events, predicates, states or properties are not nomic, in other words, according to this view, there are neither strict psychological nor strict psychophysical laws governing the mental. Mental properties cannot participate in nomic correlations with physical properties. Thus, the autonomy of the mental lies in its being anomalous.

Moreover, according to another core thesis of Davidson's view, his monist thesis, every causally interactive mental event is token identical to a physical event (Davidson as quoted in Yalowitz 2014: 6), but, given his anomalism, we cannot know *which* physical event is at issue, although we do know that a mental event must be a physical event to be causally efficacious.

Jaegwon Kim, a strong critic of Donald Davidson, holds that mental properties are only causally efficacious when they are reduced to physical properties. From a certain point of view one can say that this criticism does not hold against Davidson, as, given his monist thesis, he does not view himself as an epiphenomenalist.

Before we consider Davidson's anomalous monism, let us first look at his theory of mental causation for it was his theory of causality that gave rise to anomalous monism. Anomalous monism was propounded to show how mental causation fits into his theory of causation. Thus anomalous monism is rooted in Davidson's action theory. In his theory of causality, reasons are taken as causal explanation.

> … reasons are a species of causal explanation … What really motivates Davidson to make this claim? Essentially, it was the felt need to explain the mysterious connection between reasons and actions that motivated him to think that reasons are causes. And given his non-reductionism, he developed the metaphysics of anomalous monism in order to make this claim intelligible …. (Hutto, 1999:380)

Davidson (2001: 225) himself writes:

> … But the explanations of mental events in which we are typically interested relate them to other mental events and conditions. We explain a man's free actions, for example, by appeal to his desires, habits, knowledge and perception. Such accounts of intentional behaviour operate in a conceptual framework removed from the direct reach of physical law by describing cause and effect, reason and action, as aspects of a portrait of a human agent. The anomalism of the mental is thus a necessary condition for viewing action as autonomous …

In the article, "*Actions, Reasons and Causes"*, Davidson (2001) sets out his causal theory. He holds here that there is a relation between reason and action. A reason explains an action by giving the agent's reason for taking that action. This kind of explanation, he calls rationalization, and rational explanation for him is a species of causal explanation.

> A reason rationalizes an action only if it leads us to something the agent saw, or thought he saw, in his action – some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory or agreeable ... Whenever someone does something for a reason, therefore, he can be characterised as (a) having some sort of pro-attitude toward actions of a certain kind, and (b) believing (or knowing, perceiving, noticing , remembering) that his action is of that kind. Under (a) are to be included desires, wantings, urges, promptings... (ibid.: 1-2).

He calls pro-attitudes such as desire and related belief the primary reason of the agent to perform an action. The primary reason for action is its cause (ibid.: 2). Therefore reasons are causes. Actions can be rationalized in all sorts of ways that have nothing to do with the actual reason. What makes something the actual reason is that it causes the action, as well as rationalizing it. Davidson holds that when one acts with a primary reason (belief and desire) he acts intentionally. In rational explanation or reasoning, desire plus belief gives rise to action. Thus, reason-explanation implies both mental to mental causation and mental to physical causation (ibid.: 3).

An action may appear to have many reasons for it but there will always be one primary reason for an action. And the primary reason answers the question, 'Why did you do it?' For instance, I desire to make John happy on his birthday and I believe buying a cake for him will make him happy. Hence, I buy the cake to wish him a happy birthday but after eating the cake, John develops a stomach upset. Wishing John a happy birthday was my primary reason for

buying the cake, which was intentional, while the aftermath, the stomach upset, was not my primary reason and it was unintentional.

> C1. *R* is a primary reason why an agent performed the action *A* under the description *d* only if R consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that *A*, under the description d, has that property (ibid.: 5).

The cause of an action is thus, as stated, the reasons one gives in doing the action. Rational explanation or rationalizing has two components: it rationalizes and justifies an action.

> Because justifying and explaining an action so often go hand in hand, we frequently indicate the primary reason for an action by making a claim, which if true, would also verify, vindicate or support the relevant belief or attitude of the agent ...The justifying role of a reason ... depends upon the explanatory role (ibid.: 8).

Davidson notes that it is justification that differentiates teleological causal explanations (rational) from non-teleological causal explanations (physical), as justification does not feature in the explanation of physical events, which lack human agency. It is only the causal aspect of the explanation that is noted in a physical event. For example, in a physical explanation, such as explaining a mango falling from the tree because of the wind, we do not justify why the mango fell down, since a wind is not a personal agent. But one may give a reason to justify one's action. For instance, an agent did not help her financially because she had no money.

> Noting that non-teleological causal explanations do not display the element of justification provided by reason ... if rationalization is ... a species of causal explanation, then justification, in the sense given by C1, is at least one differentiating property. How about the other claim: that

> justifying is a kind of explaining, so that the ordinary notion of cause need not be brought in? Here it is necessary to decide what is being included under justification. It could be taken to cover only what is called for by C1: that the agent has certain beliefs and attitudes in the light of which the action is reasonable. But something essential has certainly been left out, for a person can have a reason for an action and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason. Of course, we can include this idea too in justification; but then the notion of justification becomes as dark as the notion of reason until we can account for the force of that *because*... (ibid.: 9).

For Davidson, the reason for acting implies that it is the cause of the action in question. Moreover, the reason-explanation is a species of causal explanation that designates a reason as a rational cause.

According to Hume, any causal relation instantiates a general regularity in nature and it requires two distinct events; one being an effect of the other backed by a causal law that describes how the cause causes its effect. But, in rationalizations of actions (citing of reasons for action), there is no case of one event following the other based on the regularity of nature in terms of Hume's theory of causality. There are no precise laws that determine events in rationalizations. In other words, causal laws that are essential in physical causal explanations are missing in rationalizations (ibid.:16). For Davidson, there are elements of coherence and rationality involved in the rational explanation as entailed by his principle of charity which are not obtainable in physical causation: "In explaining action we are identifying the phenomena to be explained and the phenomena that do the explaining as directly answering to our own

norms; reason explanations make others intelligible to us only to the extent that we can recognize something like our own reasoning power at work" (Rainone, 1999: 128).

To repeat, Davidson (ibid.) explains that the difference between rationalization and ordinary causation is that rationalization has two components in that it rationalizes and justifies. And there is personal agency involved that is characterized by belief and desire (pro-attitude). On the other hand, in explaining ordinary causation, only the causal aspect of explanation is taken into consideration, there is no justificatory aspect, which characterizes rationalization.

Keith Maslin (2007: 182), explaining Davidson, reminds us that Davidson distinguishes between causation and causal explanation. Causation holds between events (singular causal statements) regardless of how they are described. Causation happens at the ontological level while causal explanation occurs at the higher level. Causal explanation is dependent on how events are described, and importantly, how they are interpreted, and it employs syllogisms of a deductive nature. For instance,

(1) If a black cloth is washed with JIK bleach, the black cloth will turn into a white cloth.
(2) My black cloth is washed with JIK bleach
(3) Therefore, my black cloth has become white.

In explanation what is deduced, based on the law, "is not the event itself, but a statement (3) which describes the event, and it is deduced from two other statements (1) and (2). That is why Davidson maintains that causal explanation obtains between statements, which describe events, and not between the events themselves, as these are non-linguistic entities" (ibid.: 183).

Davidson makes the point that causal laws must back singular causal statements in the following way:

> Hume's claim … may mean that 'A caused B' entails some particular law involving the predicates used in the descriptions 'A' and 'B', or it may mean that 'A caused B', entails that there

> exists a causal law instantiated by some true descriptions of A and B. Obviously, both versions of Hume's doctrine give a sense to the claim that singular causal explanations 'involve laws'. But the second version is far weaker, in that no particular law is entailed by a singular causal claim, and a singular causal claim can be defended, if it needs defence, without defending any law. Only the second version of Hume's doctrine can be made to fit with most causal explanation; it suits rationalizations equally as well (Davidson, 2001: 16).

Thus, a reason can be the cause of an action even if no psychophysical or psychological law exists that can link that cause with its effects, described respectively as a reason and an action. That is possible if those very events are covered under another description by a law of another kind. And for him, mental events and actions can be causally connected, although no psychophysical or psychological laws exist, because they are in theory redescribable in neurological terms and covered under those descriptions by physical, chemical or biological laws.

Davidson's argument runs like this:

1. The explanations of human actions by an agent's reason are causal explanations.
2. Every causal explanation implies the existence of a law.
3. The explanations of human actions by reasons are not covered by any psychophysical law (Nannini, 1999: 98).

Reason and action relation does not fall under psychophysical or psychological laws but once they are given a neurological, chemical or physical description, they fall under a particular law that is no deterministic in nature like the one that governs the physical properties and events. For instance,

> An event like an intention of raising my hand causes it to rise. Although every cause implies a law and no 'serious' law (that is,

> a scientific law on which one can base predictions, counterfactual etc.) exists that connects intentions and actions described as such. But there can be a neurological law that connect my intention and action under another description, i.e. the description of the intention as the firing of neurons in my brain cortex and the description of the action (raising of my hand) as the contraction of muscles in my arm. In other words, it is possible if one accepts the identity of intentions and other mental states with neurological events that is if one accepts physicalism (ibid.: 99).

Furthermore, Davidson (2001: 17) writes:

> The laws whose existence is required if reasons are causes of actions do not, we may be sure, deal in the concepts in which rationalizations must deal. If the causes of a class of events (actions) fall in a certain class (reasons) and there is a law to back each singular causal statement, it does not follow that there is any law connecting events classified as reasons with events classified as actions – the classifications may even be neurological, chemical, or physical.

To understand this better, we need to turn to Davidson's notion of anomalous monism that is intended to make his theory of causality meaningful and intelligible.

Davidson's anomalous monism is a physicalist monism but it holds nevertheless on the one hand – in property dualist terms – that physical events or phenomena have mental properties, and, on the other hand – in non-property dualist terms – that the mental is anomalous, that is, the mental is not governed by strict laws, and that while the mental needs to be physical via token identity to be causally efficacious, it does not mean we can point to the exact physical event in question.

To deal with the problem of determinism, Davidson (2002: 116) in his essay, "*Mental Events"*, argues for a token-token identity theory against type-type identity. Token-token identity holds that specific mental particulars can be identified with specific physical particulars. I.e., mental and physical tokens can be in a relation of identity to each other. Token identity theory offers Davidson an opportunity to deal with the problem of physical determinism and the autonomy of the mental, as for him, mental events such as perceiving, remembering, and decision makings are anomalous, i.e. they do not obey the laws that physical events or objects do (ibid.: 116). This is the case because he implies that mental events are not reducible to physical events, the latter being nomic in character as deterministic physical laws govern them. In other words, since Davidson claims there are no covering laws for the mental, and that all strict laws are physical, he postulates a token identity only between mental and physical events, since the "vocabulary" of the laws is not specified (Yalowitz 2014: 5).This leads to monism, while satisfying both his thesis on the nomological character of causality and his mental anomalism.

For some philosophers like Kim, it is a contradiction to argue that a mental event like desire to raise my hand is efficacious in a nomological physical world. However, for Davidson, a mental event is completely efficacious in the physical world (2002: 116). Davidson (ibid.) proposes three principles that he takes to be true of mental events.

(1) The principle of causal interaction: Some mental events interact causally with physical events.

He (ibid.) offers an example of someone sinking the *Bismarck.* The sinking of the *Bismarck* by someone shows that various mental events such as perceiving, noting, calculating, judging, deciding, intentionally acting and changing of beliefs, all played a causal role in its sinking. Some of these mental events caused someone to move her body in certain ways that caused the *Bismarck* to sink. (Note that obviously, causality can also be from the physical to mental – the perception of a ship, a physical object, caused someone to believe that a ship is approaching). So there is causal interaction between the mental and the physical.

(2) The principle of the nomological character of causality: where there is causality, there must be a law. Events that are related as cause and effect fall under strict deterministic laws.

However, as we have seen, Davidson does not think that this is the case for the mental as there are no psychological or psychophysical laws governing the mental. And this leads to his third principle.

(3) The anomalous character of the mental: There are no strict deterministic laws with which mental events can be predicted or explained (ibid.).

On the one hand it seems mental properties must be ontologically different from physical properties. On the other, it could be paradoxical and contradictory to accept all three principles together, because the first two principles: causal interaction and the nomological character of causality imply that some mental events can be predicted or explained based on strict deterministic laws, but the third principle of the anomalism of the mental denies that (ibid.: 117). There are no strict psychophysical law, that is, laws that connect the mental and physical, because mental events feature a kind of rationality or normative constraint that does not obtain among physical events. The anomalism of the mental attests to human freedom in the physical world. Human beings would not be free if certain mental such as decisions, intentions fall under the deterministic laws of law nature that condition physical realties (ibid.).

Thus, for Davidson, the principles seem all three to be true. He shows this by proposing a view of the mental and the physical that clears the apparent contradiction. According to this view, some *token* mental events are identical with *token* physical events (token-token identity), and they are in nomic causal relations, however, mental *types* (kinds)[6] are "neither identical to, nor nomically co-extensive with physical *types*" (Horgan, 1995: 475). This implies event-monism. However, Davidson's view is sometimes wrongly described as type or property dualism as the anomalism of the mental (the third principle) avoids the identification of mental types with physical types (Robb, et.al, 2018: 9). What is left out of consideration on such views

[6] Events for Davidson are unrepeatable, dated individual happenings that are spatially and temporary confined. They are different from processes, states, properties and attributes (ibid.: 117).

is that for Davidson the token identity holds, but cannot be pinpointed in the sense that we do not need – and cannot – know what physical event a mental event is token identical to precisely because of Davidson's anomalism. And, moreover, we are not talking here of physical events marked off from others by having mental properties, but of the fact that mental and physical events differ because they can be differently described, thus it does not imply that mental properties are 'real' properties of things at all, which is in line with Davidson's anti-realism.

Anomalous monism resembles physicalism, because it holds that all events are physical but it denies the physicalist thesis that mental phenomena can be given purely physical explanations, thus it is a non-reductive physicalism. Anomalous monism thus, to recap, holds that some mental events are identical with some physical events, and that the mental is anomalous (meaning the mental is not governed by physical laws). However, given the token identity holding between some mental and physical events, physical laws cover the causal relation between the mental and the physical or the physical and the mental *in the case of this token-token identity* (see also Yalowitz 2014).[7]

While anomalous monism denies that there are strict psychophysical laws, nevertheless, it is consistent with the view that mental phenomena are dependent or supervenient on physical phenomena or characteristics. Supervenience implies that there cannot be two events alike in all physical respects but different in mental respects, and that any change in mental respects implies a change in physical respects, while a change in the physical respect may not imply a change in the mental respect. In other words, a particular mental event can be multiple realized in different physical states (Davidson, 2002: 119). However, supervenience of the mental on the physical does not mean that there are psycho-physical laws or that mental terms may be given a complete physical description (contrary to materialism), or that mental terms

[7] However, Davidson's version of token identity has been criticized. One familiar criticism is the fact that mental events are not spatio-temporal like physical events, so it will be hard to have a precise identification between them. How do I identify my desire to play soccer that begets other mental properties with some particular neural event or set of neural events (that are spatio-temporal) occurring in the brain (see Hornsby 1981 and Leder 1985).

can be translated into physical terms without breaching the meaning of the mental (contrary to logical positivism).

> … dependence or supervenience of this kind does not entail reducibility through law or definition; if it did, we could reduce moral properties to descriptive properties, and there is good reason to believe this cannot be done; and we might be able to reduce truth in a formal system to syntactical properties, and this we know cannot in general be done (ibid.).

Now there are critiques that claim that supervenience is not enough to save Davidson's view from epiphenomenalism (not that that is the main reason why Davidson supports it), as it does not speak to the causal or explanatory relevance of the mental as mental (Horgan 1995: 475). Jaegwon Kim, for instance, has a long history of investigating the claims of non-reductive physicalism, specifically in terms of mental causation and, notably the views of Donald Davidson[8]. Kim (2001a) argues that Davidson's anomalous monism fails to explain how mental properties are related to physical properties and thus he claims Davidson's view does not help us to understand the mind-body relation, nor does it enlighten us on the notion of mental causation. Furthermore, Kim says:

> The anti-reductive physicalist who wants to remain a mental realist, therefore, must give an account of how the mental cause and the physical cause of one and the same event are related to each other. Token physicalism, like Davidson's anomalous monism, is not enough, since the question ultimately involves the causal efficacy of mental properties and anti-reductionism precludes their reductive identification with physical properties … (ibid.: 37).

[8] Jaegwon Kim, http://www.informationphilosopher.com/solutions/philosophers/kim/.

In order to analyse the issue raised by Kim in relation to token identity's inability to account for the relationship between the mental and the physical, one can turn to Davidson's theory of supervenience, but one must also keep in mind that Davidson is a mental anti-realist – he does not believe that mental properties really exist, the most that he allows is that there are mental descriptions of events that are different form the physical descriptions of events.

There are different interpretations of supervenience. Davidson understood supervenience in relation to his token identity theory as follows: "a predicate p is supervenient on a set of properties S, if and only if p does not distinguish any entities that cannot be distinguished by S" (Davidson, 1993: 4). For Davidson, this implies monism (one substance) and this is consistent with token identity. Thus, the notion of supervenience supports Davidson's token identity theory.

Davidson's own explanation of supervenience runs like this:

> Although the position I describe denies there are physical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respects, or that an object cannot alter in some mental respect without altering in some physical respect (Davidson, 2001: 214).

Supervenience can be explained from three perspectives namely irreducibility, co-variation and dependence (Maslin, 2007: 154). Irreducibility is the idea that the supervenient phenomena or facts are irreducible analytically, by definition or ontologically to subvenient phenomena or facts. Supervenient phenomena exist over and above subvenient basal phenomena. (Thus, in relation to Davidson's token-token identity claim, mental properties cannot be reduced by definition, or ontologically, to physical properties contrary to what type-type identity theory maintains, and so all that is left is token-token identity – and this is not

deterministic token identity in the sense that multiple realisability holds but also above all in the sense of Davidson's anomalism.)

Secondly, "the supervenient phenomena are determined by, and co-vary with changes in the underlying subvenient base. This means that there can be changes in the supervenient phenomena, if and only if there are corresponding changes in the subvenient basal phenomena" (ibid.: 155). But changes only in the supervenient phenomena cannot bring about changes in the subvenient base. Multiple realizability thus becomes a reality, which again indicates nothing more than a token-token identity theory. Co-variation also means that two individual states cannot differ in their supervenient properties except if there is some difference in the subvenient features. Thus, if there is no difference in the physical features, there will be no difference in the mental features.

Thirdly, in terms of dependence, supervenient phenomena emerge from subvenient basal phenomena on which they depend for their existence. The basal phenomena also determine the supervenient phenomena's characteristics. This dependence is asymmetric because the supervenient phenomena do not determine the basal phenomena. According to token-token identity theory, the presence of mental features is determined by physical ones. However, the physical cannot reduce the mental to the physical or eliminate the mental in favour of the physical. This means token-token identity is a weak form of physicalism, as Kim (2000: 120) points out. This kind of supervenience implies non-reductive monism (physicalism) in the sense that only token identity holds; or property dualism, as while there are two distinct kinds of property, physical and mental, the mental exists because of the physical, and the mental cannot change independently of the physical (Maslin, ibid.: 156).

Supervenience has also been characterized as weak, or strong or global (Kim, 2003: 217-239). Weak supervenience implies that in the actual world the mental will always supervene on the physical, but the mental does not necessarily supervene on the physical in all possible worlds. For example, my physical duplicates in the actual world will have the same thoughts as I have; but my physical duplicates in other possible worlds may have different thoughts. It is weak because it does not hold in all possible worlds.

Strong supervenience holds that mental states supervene on physical states but this supervenience relation holds both in the actual world and in all other possible worlds. It means that any two people physically the same, in any two worlds, will also be mentally the same. It is said that strong supervenience entails weak supervenience because a strong supervenient relation holds in any one possible world. Lastly, global supervenience compares properties between all possible worlds in their entirety. Any two possible worlds that are the same physically at the base will be mentally the same at the supervenient level.

Returning to Davidson, as mentioned before, he (2001) presents anomalous monism as the opportunity to have two kinds of descriptions of an event, namely mental and physical descriptions. The mental and the physical descriptions describe the same event (they have the same extension by tokens) but they describe it in different ways or in different vocabularies. For example, a mental event like George's dislike of the smell of vodka could be given a completely physical description leaving out any mental property: as "rarefied alcohol molecules interact with George's olfactory sensors, data is transmitted from the olfactory sensors to George's brain, and specific neural patterns fire in George's brain" (Fitz and Gumm, 2010: 41). The mental description and physical description are two different descriptions of one event. "The mental description does not have to refer to something else. It is about the same physical phenomenon as the ordinary physical description, but it describes the situation from a 'higher' or more complex perspective" (Thompson, 2012: 31).

That one and the same event can be described physically or mentally shows that the mental and the physical are in a token identity relationship with each other. Anomalous monism supplemented by supervenience brings mental events into relation with physical events. This is how Davidson addresses the criticism that token identity does not account for the interaction between mental events and physical events as it fails to explain, according to Kim, how mental properties are related to the physical properties. However, it does not enlighten us on the notion of mental causation, although Davidson seems to think it does. I will get back to this in the next chapter.

8. The Strengths and Weaknesses of Non-reductivism and Functionalism

Dale Jacquette (2009: 23-24) enumerates some advantages of property dualism. Firstly, when compared to dualism, non-reductive physicalism seems to be more advantageous in that it does not have to deal with the causal interaction problem between material and immaterial substances as noted in dualism. The problem dualism has is to explain how a material substance that is spatio-temporal could interact with a non-material substance that is not extended in space. Non-reductive physicalism professes the reality of only one material substance that is extended in space, though with mental and physical properties.

Secondly, unlike idealism, non-reductive physicalism does not need to appeal to God, rather it appeals to supervenience to account for mind-body interaction or the reality of the perceived world, since it holds that there is only physical substance. And God being non-physical has no place in it. More so, it does not have to prove the existence of God, which is another difficult topic in philosophy.

In comparison to reductive physicalism, reductive physicalism reduces the mind to physics thereby professing that mental causation is dependent on the causal power of physics, which is against our common view of looking at reality. Non-reductive physicalism on the other hand, in taking the mental seriously by affirming mental properties, is in consonant with our natural way of looking at the reality of mind that it causes actions independently of physics as those mental properties are taken to 'really' exist. Thus, property dualists are mental realists.

Despite the strength of non-reductive physicalism, it has been criticised, especially by Kim (but also in its property dualist guise by anti-realists such as Davidson), and especially concerning supervenience. The question for Kim is, if a mental event supervenes on a physical one, does the mental play any causal role in the physical world?

> … the supervenience argument captures the essence of the difficulties involved. The fundamental problem of mental causation for us is to answer this question: How is it possible for

> the mind to exercise its causal powers in a world that is fundamentally physical? (Kim, 2001a:30)

Another fundamental question with regard to mental causation is this: "if mental properties are physically irreducible and remain outside the physical domain, then, given that the physical domain is causally closed, how they can exercise causal powers, or enjoy any kind of casual relevance in the physical domain?" (ibid.: 58). I will discuss this in detail in the next chapter.

Returning now first to a general discussion of the strengths and weaknesses of non-reductive physicalism, Maslin (2007: 137) enumerates the strengths of functionalism as follows: The functionalist view of mind does not only identify mental states with actual potential outward behaviour as behaviourism did, it also acknowledges the internal causal states of the mind. There is no problem of mental causation in functionalism. For example, pain is not seen as a state of an immaterial substance that causes pain behaviour. Rather pain itself is a functional state described in terms of inputs, outputs and relations to other mental states (ibid.). Thirdly, it is not chauvinistic in its approach like type-type identity, which holds that mental states are unique to human beings. According to functionalism, mental states can be realised in other animals including in artificial objects like the computer (ibid.). And fourthly, and most importantly, the functional nature of mental states answers the question of the relation between the mind and body.

In spite of its numerous strengths as seen as above, functionalism is still inadequate to account for mental states. This is because it cannot account for qualia and intentionality. I will now examine the works of Ned Block and John Searle to buttress this point.

8.1. Criticism of and Reflections on Functionalism

Ned Block (2002) in his article, "*Troubles with Functionalism"* argues that two systems might be functionally identical with one another and still only one system may have qualia or consciousness. This may imply that functionalism is too liberal in granting mentality to things that are not mental. Block compares a mental state to a "… Turing machine table of a certain

kind. Each mental state of the system is identical with one of the machine table states specified by the machine table" (ibid.: 96). Block considers the inputs and outputs to be specified by the neural impulses in sense organs and motor output neurons (ibid.).

Specifically, Block (ibid.) would like you to imagine a homunculi-headed robot of a body like yours externally, but different from you internally. The neurons of its sensory organs are connected to the bank of lights in its head with a set of buttons connecting to the motor output neuron. Inside the robot's head, there is a group of little men called *homunculi*. On the wall in the robot's head, hangs a bulletin board on which is posted a state card, which represents one of the squares specified on the machine table. Each square has a group of little men to run it. A group called 'G' men directs the G square. When the letter 'G' appears on the postcard with the input light 'I' on, one of the 'G' men will press the outward button '101' and change the state card from 'G' to 'M'. The letters 'G' and 'M' characterize the internal states of the machine. The entire system simulates your brain because its functional organization is trained to realize your brain. The activities of the little men make the system realize the same machine table as your brain. Thus, the system is functionally equivalent to you (ibid.).

Block (ibid.) then tells the story (referred to as the 'absent qualia' thought experiment) in a different way using the population of China. We imagine in China, each person in its one billion population has a two-way radio that connects her with another Chinese person and to the artificial body (the robot) as seen in the first story (ibid.). Here we have a radio transmitter and receiver instead of the little homunculi men. The radio transmitter and receiver are connected to the input and output neurons. The bulletin board on the wall is replaced with a satellite that displays the letters, which could be seen by all one billion Chinese people. Thus, the system is functionally equivalent to us human beings because it can realize mental states through its functional properties. The whole system constitutes a huge artificial external brain.

For Block, it is unreasonable to attribute experiences of consciousness or pain or any mental state to the homunculi head or the Chinese brain. They have a functional system like a human being but lack what the philosophers call qualitative states or qualia. Therefore, he claims functionalism is wrong to argue that any system whose functional organization is

equivalent to us will have the same mental states we have. Thus, functionalism is too liberal in attributing mental states to things that do not have mental states.

Another argument against functionalism is known as the *Inverted Qualia* (Block, 1990).The qualia inversion argument holds that it is possible that two people (Peter and Jane) have the same brain system, which are functionally the same in terms of input, output and causal relations with other mental states but that the contents of their visual experiences and beliefs are different from each other. For instance, the colour Peter experiences when he sees a marigold is the colour Jane experiences when she sees a violet, and vices versa. Peter's total colour experience is the inverse of Jane's own colour experience so that when Peter sees red, Jane sees blue. However, when Peter and Jane see a red object, both of them will say publicly that they are seeing a red object because they have been taught to call such a colour red, but their phenomenal or internal experience differs from each other. Internally, Peter is having a reddish experience while Jane is having a bluish experience. Peter and Jane have the same functional brain system but their qualia or qualitative experience is not the same. Thus, the functionalist view of mind does not account for qualitative aspects of mental states (see also e.g. Kind, 2017). I argue that it will be hard to know that one's subjective experience is the inverse of someone else's if a quale is indeed subjective and private. It could be that they are giving a false report of their experience as we do not have access to them as they are subjective to the experiencer.

Another strong critic of functionalism is John Searle. According to him, functionalism can also not account for another characteristic of mental states, namely intentionality. Intentional states have traditionally been described as those mental states (unlike physical states) that are directed at, or about, what is external to them (Brentano, 1874). Beliefs, desires, hopes, and wishes are all intentional states because there is always something that is desired, that is hoped for, or something that is wished for. Searle (2000), in his article, "*Do Minds compute?"* argues that a computer cannot possess a genuine thought and understanding in spite of being capable of instantiating a program or implementing a machine table. A genuine thought possesses

intentionality, in other words, it has the features of being about some states of affairs other than itself.

To argue for his point, Searle asks the question; can a computer simulate human cognitive capacities? In answering the question, Searle distinguishes between 'Strong' Artificial Intelligence (AI) and 'Weak' Artificial Intelligence (AI). Weak AI holds that a computer is a helpful tool in the study of the human mind while Strong AI holds that the computer is not only useful in the study of the human mind, but a properly programmed computer is the same as a human mind because with the right programs, the computer can understand and have other cognitive states (ibid.). Searle agrees that a computer is a useful tool in the study of the mind (Weak AI), but argues against Strong AI, that is, the computer's ability for understanding and having other cognitive states. Searle (ibid.), using a thought experiment known today as the 'Chinese Room Argument' demonstrates his position as follows.

Searle would like you, who do not understand Chinese language, to imagine that you are locked up in a room, and you are given a large batch of Chinese writing, supposing that you do not understand Chinese language and writing. Furthermore, you are given another batch of Chinese script and a set of rules in the English language (your home language) for pairing and matching the second batch with the first batch. The rules help you to match one set of Chinese symbols with another set of Chinese symbols by looking at the shapes of the symbols. You are now given a third batch of Chinese symbols with more instructions in English to help you match or pair them with the first and second batches and with the instruction that you give your response in terms of Chinese symbols. Unknown to you, the first batch you received is a script, the second batch is a story, the third batch is a set of questions on the story (input), the symbols that you give back are the answers to the questions (output), and the instructions in English are the programs (ibid.: 250).

Furthermore, you also receive the same stories and questions in English language that you speak fluently and you are expected to give answers in return to the questions in English. This is not a problem because you speak and understand English very well (ibid.). Suppose that after some time, you become good at manipulating Chinese symbols using the instructions in

English. Hence, the answers that you give are the correct answers that any Chinese person would give. Thus, you appear as a Chinese speaker to a native Chinese speaker outside the room in which you are. Moreover, your answers to the questions in English are equally good too.

Searle would argue that you do not, however, understand Chinese. In the Chinese case, unlike the English case, you gave the answers by manipulating the Chinese symbols, of which you do not understand their meaning. You simply behave like a computer; you are an instantiation of a computer program. A computer understands nothing of any stories or writings it computes. Computers and their programs only perform their functions (syntax) without any understanding (semantics) (ibid.: 251). Thus, the functionalist approach to the mind-body problem for Searle is false, as the computer-and you in the Chinese Room-will lack the mental state of understanding what is happening. Thus understanding the syntax of a language does not necessarily imply semantic understanding.

In another article, "*Is the Brain a Digital Computer*?" Searle (2000) continues his attack on the computational model of mind that holds that the mind is to the brain what a program is to the computer. He begins his argument by positing these "three questions:

1. Is the brain a digital computer?
2. Is the mind a computer program?
3. Can the operations of the brain be simulated on a digital computer?" (ibid.: 254).

For Searle, answering question 1, no physical structure is intrinsically a digital computer because its characterization as one is relative to the observer or agent who ascribes a syntactical meaning to the purely physical features of the system (ibid.: 259).

In answering question 2 he argues that mind is not a computer program because programs are defined purely formally or syntactically whereas minds have intrinsic mental content, thus, a program by itself cannot constitute the mind (ibid.: 254). The formal syntax (sentence structure) of the program does not by itself guarantee the meaning or semantic understanding of the syntax. The reason why Searle only accepts weak AI (Artificial Intelligence)

thus has to do with this difference between semantics and syntax (ibid.). How do we arrive at the semantics (understanding of the meaning of the sentence) as the programs deal with only syntax not semantics, if the mind is a digital computer as strong AI claims? The meaning of a sentence is relative to its reader. A computer could run the steps of the program for mental capacity such as understanding Chinese, without understanding the words of Chinese (ibid.). Thus, syntax is different from semantics (meaning of the sentence), and it is not adequate for semantics. (The point made in the Chinese room experiment too.)

The answer to question 3 is yes. This is what Weak AI stands for. A computer is useful in the studying of the human mind. Strong AI is the view that the mind is a computer program while cognitivism, implying Weak AI, is the view that the brain can be stimulated by a digital computer. Searle's focus is on cognitivism (the study of how the brain acquire information and understanding of things). Searle notes that cognitivism is intuitively appealing because it is associated with Alan Turing's attempt to relate human intelligence to computation in his classic paper of 1950 (ibid.: 255). Turing's classic paper states a universal Turing machine can do anything that a human can do algorithmically. The Turing machine, just like a mathematician, can do the calculation 2 +2= 4. The difference is that the mathematician does his algorithm consciously while that of the machine is a non-conscious activity (ibid.).

Thus, in summary, one thing common to criticisms against any type of physicalism, including functionalism, as noted earlier in this chapter, is the inability to account for subjective experience and qualia adequately. However, there are physicalists who argue that mental phenomena or phenomenal properties (subjective experience, intentionality, qualia, etc.) are simply physical in some way, or are non-existent. They are also antirealist of the mental. They argue their position by criticising any arguments for phenomenal properties. Let us look at their arguments in the next section.

8.2. Arguments against Phenomenal Properties

Daniel C. Dennett (2002), in the article, "*Quining Qualia"*, argues, contrary to Thomas Nagel, that conscious experience does not have the qualities of privacy, being intrinsic, and being ineffable (ibid.: 229)**.** Hence, there are no qualia for Dennett**.** To remind the reader, qualia are

terms for the ways things seem or look or taste or feel or smell or sound to individual persons. Thus, qualia are phenomenal qualities of experience (ibid.: 226). Dennett uses thought experiments, which he calls 'intuition pumps', to argue his points.

In the intuition pump called the 'inverted spectrum', Dennett (ibid.: 230) argues that qualia are not verifiable empirically. Hence, they cannot exist. (A reminder that I have explained the range of inverted spectrum thought experiments above). Dennett asks the question how we know that when we look at things we see the same colours subjectively. He argues that as we are taught about colour words by being shown public coloured objects, it is possible our public identification of coloured objects will match, even if we experience entirely different colours subjectively. We both may see the colour green in the external object but subjectively the external green colour appears red to me (ibid.). Since there is no empirical way of verifying qualia, we cannot have a matter of fact about them nor compare our subjective colour experiences (ibid.: 231).

Let us consider one more 'intuition pump'. In 'intuition pump seven', which is about two coffee tasters, Chase and Sanborn (ibid.), they have to ensure that the taste of Maxwell Coffee remains the same always. But at some point, both of them seem to have stopped enjoying the taste of coffee, which they once enjoyed. Chase shares his experience, "The coffee tastes just the same today as it tasted when I arrived. But, you know, I no longer like it. My tastes have changed. I have become a more sophisticated coffee drinker. I no longer like the taste anymore" (ibid.: 232). This implies that his experience of flavour or his quale of taste has changed, but his judgment has remained the same. Sanborn says the opposite, "But my tastes haven't changed; my tasters have changed … You other tasters all agree that the taste is the same, and I must admit that on a day to day basis I can detect no change either. So it must be my problem alone" (ibid.). Thus for Sanborn, the quale is the same, but his judgment of it has changed.

Since qualia are said to be private or subjective, both Chase and Sanborn are correct in their own taste experience (ibid.). How do we determine who is right or wrong in his own qualia? We need an empirical test that would confirm Chase and Sanborn's different tastes.

Passing such a test will confirm their authority or reliability, while failing the tests will undermine it (ibid.: 233). If we were to use an external test to confirm the reality of qualia, then the characteristic of qualia being private, or subjective, or being subject to a privileged access, is discredited (ibid.).

Chase might be subject to a blind tasting to reconfirm his claim that he knows that the coffee tastes the same after six years of working at Maxwell Coffee very well. If he does not identify coffee in the blind tasting of sipping wine, tea and coffee in quick succession of minute intervals between first and second sip, then his claim that his experience of the flavour has changed but his judgment is the same is flawed. But if he is able to identify it, the claim of being a sophisticated drinker can be supported. The truth of the matter is an outside agent or objective support is needed to ascertain the truthfulness of Chase's claim (ibid.).

Dennett uses another 'intuition pump: the experienced drinker' (ibid.: 136). It is believed that one comes to like the flavour of beer gradually after many drinks. One trains herself to come to like the taste of beer or to enjoy drinking after prolonged beer drinking (ibid.). The taste that is enjoyed by the beer drinker is not the first taste of the beer she had. Her drinking of beer has changed the first taste of her first sip of beer. Her constant drinking has changed the qualia of beer. If this is true, it does mean that the phenomenal or qualitative features are not intrinsic properties but rather extrinsic or relational properties as one taste is compared in relation to the other. Hence, there are no qualia (ibid.: 237).

I claim that Dennett is attacking memory and not qualia. Chase and Sanborn are recalling or remembering their experiences of coffee tasting, not the current, immediate and actual qualia – the raw feels. In addition, the 'experienced beer drinker' too is comparing her memory contents (remembrances of the taste of first drink and second drink). If a quale is the way something tastes or looks to me, then one can only talk about the contents of memory and not qualia. It is difficult for one to talk about the taste of a piece of beef she is eating or a colour she is seeing. For her to describe the way meat tastes, she has to stop eating and enter her memory bank to recall how the meat tasted in order to describe its taste. Thus, in this case, she will only be talking about her remembrance of how the meat has tasted to her, which is a

subjective experience. She is either talking about the meat taste (remembering the taste) or experiencing the meat taste. Thus, for me, if the memory content is subjective and private, the current or actual qualia being experienced at a particular moment will be deeper and more private and subjective as one does not explain without remembering. Qualia indeed are mental properties, in other words, metaphysical properties that are not subject to empirical verification or testing. Hence, Dennett does not reduce qualia into oblivion as he promised. I will say that Dennett does not undermine qualia as real, private and ineffable. Qualia are problems to physicalism.

Dennett also criticized the mental property called intentionality. To remind the reader, intentionality is the 'aboutness' of the content of our minds, the 'directedness' of our thinking towards something. We have desires, beliefs and hopes about something external to our minds. Brentano with his theory of intentionality created a gap between phenomenal and physical properties. For the physicalists, this gap can be removed by reducing the mind to the physical, that is, through reductionism. But this is problematic, as I have argued above. Dennett however thinks reductionism is not the only alternative theory, hence, he proposes a theory that is neither committed to reductionism nor to dualism, but talks about creatures whose behaviour is to be explained in rational terms (Dennett, 2000: 290).

He imagines a system whose behaviour can be explained or predicted by ascribing to a system beliefs, desires, hope, fears, intentions, etc. He writes (ibid.: 290):

> I will call such systems Intentional systems and such explanations and predictions Intentional explanations and predictions in virtue of the Intentionality of the idioms of belief and desire (and hope, fear, intention, hunch ...)

For Dennett, a system of thought "is an intentional system only in relation to the strategies of someone who is trying to explain and predict its behaviour" (ibid.: 291). He buttresses his point with the case of a chess-playing computer whereby its opponent adopts certain strategies to predict its moves in order to win the game. According to Dennett, there

are three different strategies or stances that one can use to confront the system namely the design stance, the physical stance and the intentional stance (ibid.).

Based on the knowledge of how the computer is designed, one can predict the movement or design response that the computer will make in response to one's strategy. This prediction will be a true one if the computer performs according to its design without any malfunctioning. The design stance is generally adopted when we are making predictions about the behaviour of mechanical objects. For instance, if one kick starts the ignition of a car, the engine will run. "The essential feature of the design stance is that we make predictions solely from knowledge or assumptions about the system's functional design, irrespective of the physical constitution or condition of the innards of the particular object" (ibid.). In the design stance, the physical constitution of the system is irrelevant, what matters most is the knowledge about the system's functional design from which predictions are made (ibid.).

In the case of the physical stance, predictions are made on the basis of the actual physical state of the particular system in relation the physical laws of nature. Physical stance predictions cover the malfunctioning of the system for instance in determining and preventing any breakdown in the system (ibid.). There is a predictive power gained from moving from the design stance to the physical stance. However, this predictive power is increased further through the intentional stance.

In the intentional stance, we treat the chess-playing computer as if it were a rational or intelligent being like us with intentional states such as beliefs, desires and hopes, whose behaviour is governed by the intentional states with the aim of winning. We treat the chess-computer like that in order to predict its rational moves with the aim of counteracting these moves with human intelligence so that we can defeat the chess-playing computer (ibid.:293).

The other stances – physical and design – will not help us in doing that as the nature or physical composition of the system is insignificant in this stance. When we take the intentional stance towards the chess-playing computer, we do not have to worry about the details of its physical constitution or the details of its program (i.e., its design). Rather, all we have to do is

determine the best legal move that the computer can make given the information possessed by it. Once we treat the computer as a rational agent with beliefs about the rules and strategies of chess and the locations of the pieces on the game board, plus the desire to win, it follows that the computer will make the best move available to it (ibid.: 292-293).

All Dennett is doing in the intentional stance is applying human intentional states to mechanical objects.

> It might first seem that this tactic unjustifiably imposes human categories and attributes (belief, desire, and so forth) on these alien entities. It is a sort of anthropomorphizing, to be sure, but it is conceptually innocent anthropomorphizing. We do not have to suppose these creatures share with us any peculiarly human inclinations, attitudes, hopes, foibles, pleasures, or outlooks; their actions may not include running, jumping, hiding, eating, sleeping, listening or copulating. All we transport from our world to theirs are the categories of rationality, perception ... The question of whether we can expect them to share any of our beliefs or desires is tricky, but there are a few points that can be made at this time; in virtue of their rationality, they can be supposed to share our belief in logical truths, and we cannot suppose that they normally desire their own destruction, for instance (ibid.: 294).

Dennett's theory of an intentional system is committed to neither dualism nor physicalism; nonetheless, it makes room for a physicalist theory of mind because for him the intentional stance can apply to any creature or mechanical object.

As a system or device, the intentional system needs to be modelled on a natural intentionality. I argue that a device that works like the heart is not the heart because the device may be of metal or plastic in structure while the heart is a biological organ with veins and

arteries for the flow of blood. The intentional stance is not directed towards intentionality itself. Intentionality is still a mark that differentiates the mental from the physical. Nevertheless, Dennett's intentional system will help us in studying intentionality. That is good thing about his theory.

The physical stance and design stance are stances we take with other mechanical objects towards things. We take them because they are natural to us and to the mechanical objects in question. Artificial objects have them, but lack intentionality, which they do not need for their environment. The intentional stance, I would say it is a human construct to help humans in understanding their environment and their interactions with other creatures, which bear the physical stance and design stance. We apply the intentional stance to the chess-computer for us to understand the computer and challenge it in the chess game. Even the chess-computer is programmed and that program is a product of man's intentionality. Intentionality is for us and about us humans. We can apply it to anything we wish.

Though Dennett uses his article "*Intentional Systems*" to attack the reality of mental states and mental causation, one can arguably use the stances to argue for the reality of mental states and mental causation by looking at or interpreting Dennett's three stances (physical, design and intentional) in terms of the three souls of Aristotle (vegetative, sensitive and rational), which are natural to humans and they are, according to Aristotle, responsible for mental behaviour in humans. In this sense, the physical stance would represent the vegetative soul, the design stance would represent the sensitive soul, while the intentional stance would represent the rational soul. I make this claim because Dennett holds that the physical stance and design stance can be taken by all beings including human beings, but only human beings have the intentional stance because they bother to explain and predict the behaviour of other. This is similar to Aristotle's idea that plants only have the vegetative soul, both the vegetative soul and sensitive soul are found in animals and human beings, while human beings are the only beings that have a rational soul. In this way, the significance of the intentional stance could be viewed as akin to the significance of the rational soul of Aristotle. I will discuss Aristotle's notion of the soul in detail in chapter four.

Another voice against phenomenal properties is that of Brian Loar. He (2000:403) argues in his article, "*Phenomenal States",* for physicalist functionalism against the objections raised by Kripke, Nagel, Jackson and Chalmers. He argues against the anti-physicalist's argument that conscious mental qualities are not identical with ordinary physical properties, which concludes that phenomenal concepts are conceptually irreducible to physical concepts (ibid.). Anti-physicalists hold that phenomenal concepts – the conceptions we have of phenomenal qualities or qualia – are neither *a priori* implied nor *a priori* entailed by physical or functional concepts (ibid.). Loar agrees that there is a metaphysical intuition that phenomenal concepts (conceptions we have of phenomenal qualities or qualia) are irreducible, but he claims this does not mean that the phenomenal properties are not identical with physical properties. He notes that the main problem of physicalism is the problem of the 'explanatory gap', the idea that we cannot explain the raw feel or qualia (the feeling) in terms of physical-functional properties in the same way that we can explain what makes a certain substance a liquid (ibid.: 404).

He argues against the 'explanatory gap' by making use of the knowledge argument because he thinks the knowledge argument addresses anti-physicalist concerns generally. The famous version of the knowledge argument is Frank Jackson's (1982) Mary. Mary is physiologically omniscient, in other words, she knows all the physical facts about colour but she is ignorant of what it means to see a red colour as she lived all her life in a white and black room. She later sees a red colour and becomes aware of what it means to experience red. For Jackson, Mary learns something new about the world. It "is a new piece of knowledge and hence she must have come to know some non-physical facts (since by hypothesis, she already knew of all the physical facts). Thus, not all knowledge about the conscious mind is physical knowledge" (Gennaro, 2017: 13). This knowledge argument is also implicit in Thomas Nagel's (2002) article *What it is like to be a bat.* As a reminder,

> Nagel imagines a future where we know everything physical there is to know about some other creature's mind such as a bat. However, it seems clear that we would still not know something crucial; namely, "what it is like to be a bat". It will not

> do to imagine what it is like for us to be a bat. We would still not know what it is like to be a bat from the bat's subjective or first person point of view (Gennaro, 2017: 13).

Loar, however, argues that Mary does not learn any new fact or truth because experiential properties are physical properties.

> ... Mary already knew of all our physical properties that we have – under their physical descriptions. What she lacked and then acquired, rather, was knowledge of certain such properties couched in experiential terms (Loar, 2000:405).

Thus, for the physicalists, there is only one physical fact about colour vision, which can be known either by employing neurophysiological concepts or by actually undergoing the relevant experience, thereby, employing so-called phenomenal concepts (Jackson as quoted in Gennaro, 2017: 14). This is similar to J.J.C. Smart's argument, in section 2.4.2, about the experiences of observing the 'Evening Star' and 'Morning Star' being two different ways of knowing the same heavenly object.

Loar (2000: 407) demonstrates in two ways how phenomenal concepts can refer to physical properties. The first shows how phenomenal concepts directly refer to whatever properties they refer to. The second shows the irreducibility of mental concepts to physical concepts. For Loar, phenomenal concepts belong to a wide class of concepts he calls 'recognitional concepts'.

> A recognitional concept need involve no reference to a past instance, or have the form 'is' of the same type as that (remembered) one ... Recognitional abilities depend on no consciously accessible analysis into component features; they can be irreducibly gestalt (ibid.).

Phenomenal concepts as recognitional concepts are type-demonstratives for they directly refer to objects. They imply expressions like, 'those', 'these', etc. In using terms like 'these' and 'those', one does not need to appeal to the individual constitutive properties of a class of species. When one speaks of 'those' oranges, she does not appeal to any particular property of an orange. And for Loar, since phenomenal concepts refer to phenomenal properties, they directly refer to physical properties because he, a physicalist, holds that phenomenal properties are physical properties, thus "...the property of *its being like this* to have a certain experience is nothing over and above a certain physical-functional property of the brain" (ibid.: 408). As recognitional concepts, phenomenal concepts do not refer to specific or particular properties (contingent or essential). They encompass the whole.

On the questions of the irreducibility of phenomenal concepts to physical concepts (conceptual independence of phenomenal and physical-functional concepts), Loar holds that the irreducibility is because recognitional concepts and theoretical concepts are conceptually independent of each other (ibid.).

> Concepts of the two sorts have quite different conceptual roles ... recognitional abilities do not depend on or get triggered by conscious scientific analysis. If phenomenal concepts reflect basic recognitions of internal physical-functional states, they should be conceptually independent of theoretical physical-functional descriptions. That is what you expect quite apart from issues concerning physicalism (ibid.: 409).

Phenomenal concepts as recognitional concepts allow us to explain why they are irreducible to physical concepts, and as direct demonstratives, they would allow us to explain how they could refer to physical things even though they are conceptually irreducible to physical concepts.

Another argument against physicalism is Chalmers' (1996) notion of the 'explanatory gap' mentioned in section 1.4.2.1, it holds that a conscious state cannot be a physical state because of the gap in our ability to explain the connection between phenomenal properties and

brain properties (see Gennaro, 2017: 13). There seems to be a problem in understanding the relationship between brain properties and phenomenal properties since one can be present without the other. Why or how does some brain process produces that particular taste or visual sensation? It is easy to explain why water is H_2O or why heat is mean molecular kinetic energy, but how a particular brain property is a mental property posits a bigger problem for materialism. Hence, materialism is false (ibid.).

As noted above, Loar believes that we can explain how a certain phenomenal property might be or not be identical with a certain physical-functional property. He uses the functional definition of liquidity to make his argument. He states that an explanation of liquidity in physical-functional terms is fully *a priori*. The concept 'liquidity' is analysed in terms of its functional description, and then it is shown that the physical theory of water implies, *a priori*, that the functional description of liquidity is realized. There is conceptual dependence between liquidity and the physical theory of water. But, there cannot be an *a priori* explanation of phenomenal qualities in physical functional terms, because of the conceptual independence of phenomenal concepts and functional concepts. But this explanatory gap, for Loar, is an epistemic or conceptual phenomenon that does not have any metaphysical effects. It would be wrong to draw a metaphysical conclusion from an epistemological conclusion; that we do not know the relationship between the mental and the physical, should not entitle us to conclude that the mental is not physical.

I argue otherwise that an epistemological inquiry may lead to a metaphysical conclusion as epistemology, although distinct from metaphysics, is not necessarily completely separable from metaphysics. Epistemology presupposes a knowing subject and a known object. The knowing subject (human person) and known object (a tree, a house, a car, etc.) are entities that are material objects studied by metaphysics. The knowing subject must exist to know the existing object. We can only know what exists as physical or mental. So from the point of view of knowledge, we can make inferences about the existence or non-existence of a particular phenomenon.

On the other hand, for Loar, the conceptual imbalance between phenomenal concepts (recognitional-type demonstratives) and physical-functional concepts (descriptive terms) does not entail an explanatory gap (ibid.: 413). He argues that there is no explanatory gap between phenomenal concepts and physical-functional concepts in that there is nothing missing between them; they are identical (ibid.: 414).

> ... What generates the problem is appreciating that there can be two conceptually independent 'direct grasps' of a single essence, that is grasping it demonstratively by experiencing it and grasping it in theoretical terms. The illusion is of *expected transparency:* a direct grasp of a property ought to reveal how it is internally constituted, and if it not revealed as physically constituted, then it is not so (ibid.).

I have shown earlier in chapter one that phenomenal concepts are not structurally constituted; they are about 'what is it to feel like', which cannot be captured structurally or physically. They are not objectively constituted as physical things describable in physical terms. Thus, there must be a difference between them and physical substances. I claim that the difference makes a phenomenal concept a phenomenal concept. And that difference is no other thing but qualia.

Further, Loar argues that Nagel has a correct observation about mental facts or concepts; they are accessible from one point of view (subjective). They are self-directed. However, he is wrong in his conclusion that these subjective and phenomenal concepts cannot generate properties that can be captured in an objective sense. Loar thinks that it is coherent to claim that the subjectivity of a phenomenal quality is an objective physical functional aspect of that property. And objective descriptions leave out subjective conceptions because they do not make use of them, not because they cannot account for them fully as psychological states (ibid.; 418). Loar here acknowledges the reality of mental phenomena that are subjective in character, but only *qua* physical objects.

In turn, Churchland (1991: 270) from the perspective of eliminative materialism, argues that phenomenal properties are not beyond physical science; yellowness is an objective property identical to a certain wavelength and sweetness is an objective property identical to micro-chemical properties of a substance. For him, qualia and subjective experience are explainable from the perspective of science.

Moreover, the physicalists argue that Nagel's argument does not show that a bat's mental states are not identical with its brain states. It is necessary that there should be a part of a bat's brain, which humans cannot understand, as human brains are different from the bat's brain (Gennaro, 2017: 13). However, Nagel's aim in his paper is not to show whether a bat's mental state is identical to its brain state, rather his primary aim in the paper is to show that we may have a thorough scientific knowledge about bat's brain physical structure and its functionality, but yet we cannot know 'what it is like to be a bat'. The knowledge of its phenomenal consciousness will always be missing (ibid.).

On functionalism, Loar (2002: 418) notes there are two functional theses namely that all phenomenal concepts are functional concepts, and that all mental properties are functional properties (ibid.). Loar disagrees with the first thesis that all phenomenal concepts are functional concepts because phenomenal concepts cannot be captured in purely functional terms as he has shown with his argument about liquidity and water, however, he accepts the second, that phenomenal properties are functional properties. He holds that anti-physicalists have denied the second thesis with 'inverted qualia' and 'absent qualia' arguments, which for him are not persuasive enough.

The 'inverted qualia' argument, as mentioned before, is the argument that there may be different experiences with the same functional role. John and Jane may have the same functional role but have different experiences. John may experience something as red while Jane experiences the same thing as green. If this is possible, then phenomenal qualities are not functional properties (ibid.). If they were functional properties, John and Jane would be seeing or experiencing the same colour.

Block (1978) advanced the 'absent qualia' argument, also mentioned above already; he supposed that the Chinese nation were organised to realise the psycho functional "organization of a person seeing green. Evidently, the Chinese nation would not collectively be seeing green or having any other sensation" (Loar, 2000: 419). Any psycho-functional property could in this way be realised without a given phenomenal quality and hence a phenomenal quality cannot be identical to a psycho-functional property (ibid.). Loar is agnostic about the anti-physicalists' arguments. For him, it will be absurd to think that the Chinese nation as whole would have phenomenal qualities. However, this absurdity, for Loar, can be only considered if we have some intuitive knowledge of the nature of phenomenal qualities to know if the Chinese cannot have them collectively.

For Loar, in the content of both arguments – inverted qualia and absent qualia – currently we are not endowed with the natural capacity or insight into physical states to know that one is seeing red is not a functional property. With our limited brain research and knowledge, it is difficult to conclude that seeing red is not a fine grained functional property, Loar argues (ibid: 418)**.** I argue that now our limited scientific knowledge is telling us that subjective experience or quale is not physical or a functional property because no physicalist theory can presently account for them. This is what we should hold as of now until our knowledge is matured. But will the matured scientific knowledge in the future necessarily tell us that seeing red is a fine grained functional property? The answer is 'no' as David Chalmers has criticized any neuron-scientific explanation of consciousness. For Chalmers, neurobiology cannot explain consciousness. He (1996: 115) writes:

> Neurobiological approaches to consciousness have recently become popular. Like cognitive models, these have much to offer in explaining psychological phenomena such as the varieties of awareness. They can also tell us something about the brain processes that are correlated with consciousness. But none of these accounts explains the correlation: we are not told why brain process should give rise to experience at all … How

> does one perform the experiments that detect a correlation between some neural process and consciousness?

It has been argued that future physics will explain consciousness even though the physics as we understand it today, which consists "in an arrangement of particles and fields in the spatio-temporal manifold, undergoing complex processes of causation and evolution" (ibid.:118), is incapable of explaining consciousness. The question is how will the new physics differ from the old? Physics has always dealt with the structures and dynamics of fields, waves, particles and the like. For Chalmers (ibid.), any new physical theory, as far as it is physics, will still deal with the structures and dynamics. The best the new physics can do is to give us more structures and dynamics, which will help us to have "satisfying explanations of all sorts of high-level structural and functional properties but conscious experience" (ibid.) will remain unexplained. This implies that no new facts about physical structure and dynamics can result in facts about phenomenology.

Chalmers further argues that a novel physical theory such as quantum mechanics that posits a problem to the causal closure principle (microphysics is causally closed), by acknowledging the measurement by a conscious observer does not even explain consciousness despite the fact that it assumes the existence of consciousness and uses it to explain certain physical phenomena (ibid., 120). Neurobiological and cognitive theories, just as physics based theories, are inadequate in giving any reductive explanation to consciousness (ibid.: 120).

Mel Thompson (2012: 79) agrees with Chalmers:

> In spite of the brilliant achievements of neuroscience to date, our understanding of how the brain works is in its relative infancy. Indeed, philosophers who want to argue that neuroscience will provide all the answers to what we experience as mind tend to speak of a 'perfect neuroscience' in the future, rather than our present state of knowledge … Nobody seriously doubts that the brain is the physical organ most immediately

> responsible for those things that we experience as mind. The fact that we can now 'see' the neural activity corresponding to consciousness and mind, only reinforces that point. But the brain is part of the physical world and subject to physical causation, and the fact does not change with the closer inspection of the brain that neuroscience provides. The old problem of how brain activity actually delivers the experience of consciousness, or how a mental operation (e.g. deciding to do something) can result in physical action, remains the same. Hence, if perfect neuroscience enabled us to observe the operation, over milliseconds, of the chemical exchanges and pulses of changed electrical potential that constitute the firing of each of the billions of neurons, that would not in itself solve the mind/body problem. We would still be left with the mystery of how all that detailed neutral activity relate to our conscious experience …

However, the functionalists have their own approach to address the problem of consciousness, which is called higher-order theories. Let us now, to end off this chapter, look at the theories to see if they still fall short of addressing phenomenal consciousness adequately.

8.3. Higher-order Perception States of Consciousness

Consciousness can be distinguished as creature-consciousness and mental-state consciousness. Creature consciousness is related to an individual person or organism being conscious, and mental-state consciousness is when one of the mental states of a creature is conscious (Carruthers, 2016: 1). Creature-consciousness is further distinguished into intransitive and transitive types. That a creature or organism is conscious intransitively means that it is 'awake' not 'asleep'. And to say that a creature or organism is conscious transitively means that it is perceiving something or is aware of something (ibid.: 1). We can say that John is conscious of his fears. Thus, creature-consciousness is described based on the English transitive and

intransitive verbs. Intransitive verbs have no object, while transitive verbs are always accompanied by an object.

In the case of mental-state consciousness, there is a distinction between phenomenal consciousness – the property of what it is like to feel, hear, touch , see and taste something (Nagel 1974) and access consciousness – that is functionally defined (Block 1995). (This distinction has already been introduced in section 1.4) Mental states that are access conscious without being phenomenally conscious such relate to thoughts or judgment. It is debatable today if mental states can be phenomenally conscious without being access-conscious, and whether phenomenal consciousness is functionally reducible to representational intentions (ibid.2).

The fact that mental-state consciousness plays a functional role is not problematic from a physical perspective on mental functions (ibid.). What could this functional role be however? Some contemporary philosophers hold that for a mental state to be conscious it must have an impact on the creature's decision-making processes; it must aid the individual in her cognitive exercises or functions (Kirk 1994; Dreske; 1995; Tye 1995), and more so, that the processes must be rational ones (Block 1995). And there are those (Armstrong, 1968, 1984; Rosenthal

, 1986) who "think that the relevant requirement for access-consciousness should be related to high-order representations-experiences and/or beliefs – of that very state" (ibid.).

However, phenomenal consciousness seems to be a problem to the functionalist notion of mental phenomena. The problem is whether phenomenal consciousness can be functionally explained. Cognitive or representational theories hold that it can be functionally analyzed or explained. Higher-order cognitive theories claim "that phenomenal consciousness can be reductively explained in terms of representations (either experiences or beliefs) that are high-order" (ibid.). However, I argue that the functionalist's explanation of phenomenal consciousness is not adequate to account for it.

There are two strategies by which the functionalists account for consciousness. The first strategy, which is called the "divide and conquer" strategy, makes a distinction between P-consciousness and A-consciousness. This strategy can only explain some aspect of

consciousness (ibid.: 4).The second strategy looks at consciousness from a functionalistic perspective. In this case, the functionalists take intentionality as more fundamental than consciousness and explain consciousness in terms of intentionality. “Intentionality is the “aboutness” of mental states, their property of referring to something, or representing some state of affairs” (ibid.). And it is possible that the thing referred to or represented may not exist.

The example of the first strategy is Block’s division of consciousness into access consciousness (A-consciousness) and phenomenal consciousness (P-consciousness). A mental state is phenomenally conscious through its experiential properties which involve properties of sensations, feelings, perceptions, thoughts, wants and emotions. A mental state is A-conscious when it is used for reasoning and rational direction of action or speech.

> … the non-phenomenal notion of consciousness that is most easily and dangerously conflated with P-consciousness: access-consciousness. A state is access conscious (A-conscious) if, in virtue of one's having the state, a representation of its content is (1) inferentially promiscuous (Stich 1978), that is, poised for use as a premise in reasoning, (2) poised for rational control of action, and (3) poised for rational control of speech. (I will speak of both states and their contents as A-conscious.) These three conditions are together sufficient, but not all necessary. I regard (3) as not necessary (and not independent of the others), because I want to allow that nonlinguistic animals, for example chimps, have A-conscious states. I see A-consciousness as a cluster concept, in which (3) - roughly, reportability - is the element of the cluster with the smallest weight, though (3) is often the best practical guide to A-consciousness. Although I make a firm distinction between them (Block, 1995: 231).

Mental states that are P-conscious are the so called “hard problem” of consciousness, and for the functionalists they are poised for neurophysiological reduction. A-conscious mental

states are understood in terms of their functional role as they assist in making the representations accessible to the brain. Block's A-consciousness and P-consciousness represent two kinds of consciousness rather than two aspects of it. (Chalmers's discussed the two kinds of consciousness in terms of the hard problem and the easy problem of consciousness – see section 1.4.2.1). Based on Blocks' categories of consciousness, it is conceptually possible that a creature may have A-conscious states without ever having any P-conscious states.

> What about A without P. In the target article I said that such cases were conceptually possible, but I knew of no actual ones. If it is so much easier to find P without A than A without P, that is a striking empirical fact (ibid.:232).

Thus, the distinction between A and P consciousness does not help the functionalists in explaining phenomenal consciousness in terms of the intentional.

On the second strategy, the functionalists explain consciousness in terms of intentionality. Phenomenal properties represent the properties of ordinary external, non-mental things. For them, conscious experience does not have any directly introspectible phenomenal properties of its own. What are often taken to be the phenomenal properties of conscious experience are actually the properties of sensible, external non-mental things that experience represents. "A conscious visual experience of a ripe tomato has no introspectible property of phenomenal redness; it merely represents, accurately or inaccurately, the tomato as red" (ibid.: 6). This is called the transparency thesis, an ally of first-representationism, the doctrine that holds "that the phenomenal character of consciousness is exhausted by – supervenes on and depends on – certain of its firs-order representational properties". "Our experience does not reveal the existence of any qualia, for our experience is transparent – when we attend to our experiences, our attention goes right through their objects" (Kind, 2007: 1). This is a representationalist view of consciousness according to which the qualitative content of experience supervenes on or even reduces to, the intentional content of experience (ibid.).

A first-order representational property represents external, non-mental objects or states. (A representational property that purported to represent a mental state or event would be termed higher-order.) (ibid.). "…that when one turns one's attention away from, say, the blue sky and onto one's experience itself, one is still only aware of the blueness of the sky. The experience itself is not blue; rather, one 'sees right through' to its representational properties, and there is nothing else to one's experience over and above such properties" (Gennaro, 20). The representational property does not account for phenomenal consciousness – the property of what it is like to be in a state of perceiving blueness in the first instance. What is obtainable is the functionalist interpretation of the mind, which is the relation of input and output in generating access consciousness of blueness.

According to the first order representational theory, not all the first-order representational properties contribute to the phenomenal character of a mental state. How do we know the mental properties that are qualified to be conscious mental states? According to M. Tye's first order representationist theory of consciousness, the phenomenal character of consciousness has these features that he has abbreviated as PANIC. The mental state consciousness is poised (P), abstract (A) (the content is not about any concrete thing), non-conceptual (N) (may not be a concept for the represented object or state), intentional content (IC) (conscious state must have a content). The functional role of mental-state consciousness is captured in terms of being poised; the phenomenal properties directly impact general cognition and action, assuming attention is properly focused and certain concepts are possessed (ibid.). The important thing is "the experiences and feelings … stand ready and available to make a direct impact on beliefs and/or desires. For example … feeling hungry … has an immediate cognitive effect, namely the desire to eat" (ibid.).

Tye's theory is criticized to be inadequate to address the hard problem of phenomenal consciousness,

> This is partly because what really seems to be doing most of the work on Tye's PANIC account is the very functional sounding "poised" notion, which is perhaps closer to Block's

> consciousness … and is therefore not necessarily able to explain phenomenal consciousness. It does not cover all kinds of conscious states. Some conscious states seem not to be "about" anything, such as pains, anxiety, or after images, and so would be non-representational conscious states. If so, then conscious experience cannot generally be explained in terms of representational properties (ibid.:20).

Higher-order representationism (HOR) also aims to explain consciousness, but unlike its first order counterpart, it takes mental states to be conscious in virtue of them being the intentional objects of other mental states. In this case, a mental state is conscious because it is the object of some kind of higher order representation (HOR) (Gennaro, ibid.: 20). For example, my desire to write a good encyclopedia entry becomes conscious when I am (non-inferentially) "aware" of the desire. Intuitively, it seems that conscious states as opposed to unconscious ones, are mental states that I am "aware of" (ibid.). High order theories try to explain consciousness in mentalist terms, that is, by reference to such notions as "thoughts" and "awareness". Conscious mental states arise when two unconscious mental states are related in a certain specific way; namely that one of them (the HOR) is directed at the other (M) (ibid.). Higher order (HO) theorists believe that their approach to consciousness offers a better explanation of consciousness than any purely first-order representationism (FOR) theory, which has significant difficulty in explaining the difference between unconscious and conscious mental sates (ibid.).

David Armstrong (2000) also defends functionalism against the criticism that it cannot account for consciousness or qualia. He holds that a physicalist account of mind can account for consciousness. He bases his arguments on the close relationship between sense perception and selective behaviour. For him, perceptions are inner states or events apt for causing certain sorts of selective behaviour towards our environment (ibid.: 143). For Armstrong, to perceive is the same as obtaining a key to a door; you are not obliged to use the key always but the key becomes essential if you want to open the door (ibid.). A blind person is a person who does not

have certain keys and as a result, she is not able to operate in her environment like a sighted person (ibid.).

Consciousness for Armstrong "is nothing but perception or awareness of the state of our mind"**.** (ibid.). For instance, the absent-minded driver is temporarily unaware of the road. His perception of the road is on and off; if it were not, the car would skip off the road, but he is intermittently aware of the road as he drives. He perceives the road but he does not perceive his perceiving (ibid.). He is not conscious of his experience of the road continuously. Armstrong holds that consciousness of our own mental states is a perception of our own mental states. It is, therefore, an inner state with the capacity for selective behaviour towards our mental states. And this inner state (akin to higher-order properties in the sense of the above discussion) is physical in nature. Armstrong (ibid.) writes:

> If we are convinced, on general scientific grounds, that a purely physical account of man is likely to be the true one, then there seems to be no bar to our identifying these inner states with purely physical states of the central nervous system. Therefore, consciousness of our own mental state becomes simply the scanning of one part of our central nervous system by another. Consciousness is a self-scanning mechanism in the central nervous system.

I think that what Armstrong has in mind when he talks about consciousness is access consciousness and that is what he argues for in his article as reflected in his quote "... And so consciousness of our own mental state becomes simply the scanning of one part of our central nervous system by another. Consciousness is a self-scanning mechanism in the central nervous system" (ibid.). He still seems to ignore the phenomenal consciousness, the qualia.

There is a higher-order thought (HOT) theory that takes the higher-order state (HOR) to be an occurrent thought (to be a thought of some kind) rather than a perception. For D.Rosenthal (2005), the proponent of this theory, the first-order mental state M is the

intentional object of a second-order thought. Even the second order thought may become the intentional object of a third-order state; for instance, in the case where the subject engages in conscious introspection, which does not occur in everyday consciousness (ibid.). There is no phenomenal quality about the second-order state, because it is a thought, not a perception. According to functionalism, a mental token belongs to its mental type by the token's relational properties (ibid.).

Higher-order theories of consciousness thus try to explain the distinctive properties of consciousness in terms of some relation obtaining between the conscious state in question and a higher-order representation of some sort (either a higher-order perception of that state, or a higher-order thought or belief about it). The most challenging properties to explain however remain those involved in phenomenal consciousness – the sort of state that has a subjective dimension, that has 'feel', or that it is like something to undergo.

9. Conclusion

I have observed that reductive physicalism is also known as the identity theory or mind-brain identity theory or central sate materialism.

Behaviourism is a form of reductive physicalism that holds that we know the minds of others through their behaviour. It arose as a criticism to Cartesian dualism that separates the mind from the body. Ryle offers a counter theory of mind to Descartes because for him every mental phenomena is a disposition to behave or act in certain way. Ryle offers a behaviourist view of mind called logical behaviourism that deals with meaning of the mental terms. It has been criticized that it cannot account for mental state and how they relate with each other as the focus is on external behaviour.

I have also discussed identity theory that comes in two kinds of views namely type-identity and token-identity. It was criticism of type identity that led to token identity. Token identity identifies an instance of mental state is an instance of physical state. Type identity (type physicalism) holds that pain in general is theoretically reducible to the C. fibers firing. U. T. Place and J. J. C. Smart are proponents of type-identity theory. One of the advantages of

identity theory is that there is no problem of interaction between the mental and the physical unlike dualism. The causal efficacy of the mental is dependent on physical properties. The eliminativism materialist like Richard Rorty and Paul Churchland argue that mental properties do not exist what exist is only physical property. Eliminativism is seen to be contrary to our common sense view of human beings as we do think that the mind causes actions in us.

Non-reductive physicalism holds that there is only one physical event but mental properties are different from physical properties. Under non-reductive physicalism we have property dualism, emergentism, epiphenomenalism, double aspect theory and anomalous monism.

Functionalism, a non-reductive physicalism theory, holds that mental state can be functionally realized in many ways. Functionalism holds that mental states have functional or causal roles in the brain. Mental state or mind can be understood without knowing the particular detail of their physical –chemical brain states or structure which realizes the mind. A mental sates is characterized by its causal relation to sensory inputs, other mental states and behaviour. We have different types of functionalism such as psychophysical functionalism, machine functionalism, analytical functionalism and role and realizer functionalism. Davidson is a role functionalist and Kim is realizer functionalism. For all functionalists, pain is associated with the causal role of being typically caused by tissue damage and typically causing aversive reaction.

I also looked at anomalous monism another form of non-reductive physicalism. Mental anomalism is central to anomalous monism and it holds that mental states and properties are not nomic. It tends to establish the autonomy of the mental in its being anomalous, that is, the mental is not governed by any deterministic laws. Anomalous monism is rooted in the theory of causality of Davidson. Davidson has argued that reasons are causes; they cause actions in humans. Reasons are seen as causal explanation. While anomalous monism denies that there are strict psychophysical laws nevertheless it is consistent with the view that mental phenomena are dependent or supervenient on physical phenomena.

Davidson is a mental anti realist; he does not believe that mental properties really exist, but he holds there are mental description of events that are different from the physical description of events. Notion of supervenience supports Davidson's token identity. Non-reductive physicalism cannot account for mental sates especially qualia or phenomenal consciousness.

But the functionalists have their own approach to address the problem of consciousness which is called the high order theories. Through these theories, they consider consciousness from a functional perspective. They take intentionality to be more fundamental than consciousness and explain consciousness in terms of intentionality. Intentionality is the "aboutness" of mental states. For them, phenomenal consciousness does not have any directly introspectible phenomenal properties of its own. Phenomena of consciousness experience are actually the properties of sensible, external non-mental things that experience represent. For Armstrong, consciousness of our own mental states is a perception of our own mental state. Functionalism comes closer to account for consciousness but it only succeeds in accounting for access consciousness and fails woefully to account for phenomenal consciousness.

In the next chapter, I will discuss mental causation in modern theories with the major focus on Kim and Davidson with the aim of showing that Davidson does not salvage non-reductive physicalism from Kim's criticism.

Chapter 3: Modern Discussion of Mental Causation – Kim and Davidson

1. Introduction

In chapter one, I defined the playing ground for my thesis by defining some key terms such as mental realism, which is the view that mental properties have independent causal power contrary to Kim's notion that the causal power of the mental resides in their physical properties. I indicated that the causal efficacy of mental phenomena under investigation is that of phenomenal consciousness. For me, phenomenal consciousness can, *qua* mental, cause both physical and mental events. In chapter two, I showed that one of the main problems of the physicalistic theory is the inability to account for mental causation, as it has no place for phenomenal consciousness. Moreover, it was argued that even functionalism does not account for mental causation.

In this chapter, I will examine the notion of mental causation in the context of Donald Davidson and Jaegwon Kim's discussions. Davidson is a role functionalist while Kim may be categorized as a realizer functionalist. Davidson is a proponent of the theory called non-reductive physicalism that appreciates the reality of the mental by anchoring it on the physical yet does not reduce the mental to the physical. This is the view that mental properties is a different ontological class of properties from the class of physical properties, and that mental events cannot be reduced to physical events, although the physical is somehow prior to the mental. In this sense, this view acknowledges that the mental supervenes or emerges from the physical. In this way, it creates a midway between Cartesian dualism and physicalism. I will consider its response to the issue of mental causation and argue (mostly in the next chapter, although I will touch on the argument here) that in the form of property dualism it may in some respects be similar to Aristotle's theory of hylomorphism, a theory that every living thing and non-living thing in nature are explained from the principles of matter and form.

Donald Davidson argues in favour of mental efficacy (the causal power of the mental to cause or affect mental and physical events) in order to preserve the autonomy of intentional and other mental states in the physical world. He focuses on the notion of supervenience of the

mental on the physical in order to show that there is a correlation between the mind and the brain though he committed to there being no laws linking mental and physical expression. This view is contrary to Jaegwon Kim's reductivist account in which he argues that mental properties are epiphenomenal (causally inert) in a world that is physically closed, in the sense that physical causes are enough to account for every physical effect. On such a view, there is no autonomous mental causation. I will show that Davidson's argument for mental causation offers only a rational explanation of events (whereby mental predicates or concepts are used to explain or describe other mental and physical events), and not a causal explanation.

Davidson argues for mental causation based on his view entitled anomalous monism (discussed in section 2.7), but because of Kim's supervenience argument against mental causation in terms of the completeness of the physical world, I argue that supervenience is not the right principle to save mental causation and non-reductivist argument. I argue that it is Aristotle's hylomorphism and his theory of causality that offer a possible route to saving mental causation. Mental causation can – and even should – not be saved by any principle, such as anomalous monism, that gives prime importance to physics.

I will first look at the history of the modern discussion on mental causation, after which I will look at the standard theory of action as mental causation entails action or agency. Then I will recall Davidson's causal theory that is cushioned on anomalous monism, and then I will discuss Kim's argument against supervenience in terms of mental causation and his functional reduction, having discussed his notion of mental causation in section 1.3.2. This will lead me to give criticism and reflection on mental causation in the context of Davidson and Kim with the sole aim to show that Kim's supervenience argument undermines Davidson's effort to save mental causation and the non-reductivist argument.

Let us now briefly look at the history of modern discussion mental causation.

2. Modern Discussion of Mental Causation

Kim (2001a: 57) claims that Descartes is responsible for both the problem of mental causation and the mind-body problem, since Descartes propounded the ontology of two radically

separate types of substances, namely material bodies that are spatial and temporal and conscious minds that are immaterial and non-spatial. The question that was asked by Descartes' contemporaries was: How could such disparate substances, one extended in space and the other essentially lacking in spatial properties, causally influence one another, or "intermingle" as Descartes said, to form a "union" that we call a human being?(ibid.: 57)

Kim (2001a) claims however that in recent years the mind-body problem has evolved into a new form that is different from the Cartesian version. This new form of the mind-body problem seems now to be a threat to physicalists such as property dualists who want to take the mental seriously. It was actually Davidson's 'anomalous monism' that ignited the current worries about mental causation by conceiving the idea that mental properties or events are constitutively distinct from physical properties or events. The mental, for Davidson (2001), is essentially normative and regulated by the principles of rationality, which has no place in the physical realm. From this constitutive difference between the mental and the physical he argues for his 'anomalism' of the mental, according to which he claims there are no strict psychophysical laws connecting the mental and the physical. In spite of this fundamental difference between the mental and the physical, like Descartes, Davidson still wants to argue for causal interaction between the mental and physical, and thus, he argues for anomalous monism as another fundamental aspect of his theory of the mind (Kim, 2001a: 58).

Therefore, the current debate concerning the problem of mental causation is based on Davidson's anomalous monism and is about the causal efficacy or relevance of mental properties in a world that is fundamentally characterized or defined by physical properties (ibid.). This new problem applies to all forms of non-reductive physicalism, such as property dualism, that has become a popular view with the 'death' of the mind-brain identity theory (that is reductive 'type' physicalism, and discussed in section 2.4). The question is about the causal relevance of irreducible mental properties in a world that is physically closed (the world or universe being physically closed implies that reality is composed of micro particles or is reducible to micro-particles) (ibid.).

Jaegwon Kim and other physicalists (anti-dualists) deny the causal efficacy or power of mental properties. For them, it is unthinkable that a mental event that is not in the same spatio-temporal realm as a physical event can interact with, or cause, the physical event.

We discuss mental causation because of the vital role it plays in our lives. It propels humans into action. Kim (2001a:31), in his book *Mind in a Physical World*: *An Essay on the Mind-Body Problem and Mental Causation*, enumerates some reasons why he argues that mental causation is real. These reasons could be: (1) without the causal efficacy of mental states, human agency is an illusion. We are responsible for our actions because our desires, beliefs, intentions, decisions, and hope cause us to move our physical bodies to perform actions in the physical world. (2) In reasoning, mental states cause one another. Acquisition of new beliefs or knowledge may cause me to change my previously held beliefs and knowledge, which, without mental causation, will not be possible. (3) Perception, memory and reasoning are possible due to mental causation. Moreover, without them, there will be no human knowledge.

Thus, mental causation plays a very important role in causing human action. There are many ways of examining human action in philosophy but the kind of theory that Davidson worked with is termed as the standard theory of action.

3. The Standard Theory of Action

Donald Davidson has been credited as one of the main contributors to both the standard theory of action and the standard conception of action – "The contributions of Anscombe and Davidson have established a standard conception of action, and Davidson's work has provided the groundwork for a standard theory of action" (Schlosser, 2019: 2).

The standard conception of action explains action in terms of intentionality while the standard theory of action "explains the intentionality of action in terms of causation by the agent's mental states and events" (ibid.: 1). And these two approaches gave rise to what is called "a standard conception and a standard theory of agency". Both hold that agency is a term used to denote the performance of intentional action (ibid.:2).

However, the standard conception makes two claims namely (1) the concept of intentional action is more fundamental than the notion of action. In particular, "action is to be explained in terms of the intentionality of intentional action" (ibid.). (2) Intentionality is closely connected with the idea of acting for a reason (ibid.). The first claim is about two different views of individuating our actions. An action may be intentional under some description and unintentional under others (Anscombe 1957; Davidson 1963). Imagine that you disturb a thief by switching on the light in your house, and you switch on the light to see (this is one event that is intentional under description), not to disturb the thief. Nevertheless disturbing the thief is still an event because it is an intentional action under some description. But on the second view of the first claim, "something is an action either if it is identical with or "generated by" an intentional action". (ibid.). This view holds then, disturbing the thief is your action either if it is an intentional action or if it is generated by an intentional action (your switching in the light, in this instance). If it is merely generated by an intentional action, it is an unintentional action of yours.

The second claim of standard conception holds that there is a close connection between acting intentionally and acting for a reason. Anscombe and Davidson think that this 'close connection' is identity. In the footsteps of Aristotle,

> they both held the view that to act intentionally is to act for a reason, and that to act for a reason is to act in a way that can be rationalized by the premises of a sound practical syllogism, which consists, typically, of a major premise that corresponds to the agent's goal and minor premise that corresponds to the agent's take on how to attain the goal (ibid.).

And for Davidson, to have an intention means to have a desire and a belief that are in line with the major and the minor premise of the relevant syllogism (ibid.). The holders of the second claim believe that "intentions play a very important and irreducible role in practical reasoning, long-term planning, and in the initiation and guidance of action" (ibid.:3).

Furthermore, the standard conception is distinguished from the standard theory, in the sense that it is not associated with any theory of what it is to act intentionally and for a reason, nor is it associated with any particular theory of the nature of reason explanations. However, on the other hand, the standard theory gives "a causal account of intentional action and reason explanation for it holds that something is an intentional action and done for a reason just in case it is caused by the right mental states and events in the right way" (ibid.). The 'right' mental states and events are states and events that rationalize the action from the agent's point of view (such as desires, beliefs, and intentions). This is often called "the causal theory of action".

The standard conception is compatible with non-causal theories of intentional action and reason explanation. It is generally agreed that a reason explanation of an action usually renders the action intelligible by revealing the agent's goal or intention (ibid.). According to non-causal theories, having the relevant goals or intentions does not consist in the possession of causally efficacious mental states or events (ibid.). Davidson's theory of action is known as a causal theory of action.

4. Davidson's View of Mental Causation

Donald Davidson's action theory is a causal theory of action which I have discussed in section 2.7. We can remind ourselves that Davidson holds that reason explains an agent's action in as much as the reasons are causes for their action.

To argue for mental causation, Davidson argues that mental events are physical events and their interactions fall under physical laws, which implies monism. However, due to the rationality and coherence of mental properties, mental properties do not fall under physical laws. Hence, mental types or properties are anomalous. Thus, we have what he calls anomalous monism (discussed in section 2.7). Anomalous monism implies a token identity relationship between mental events and physical events; mental events are physical events but mental types are not physical types (see section 2.7). Anomalous monism is anchored on four principles. It will be good to mention the principles again:

(1) The principle of causal interaction: some mental events causally interact with physical events.

(2) The principle of the nomological character of causality; where there is causality, there must be a law. Events that are related as cause and effects fall under deterministic law.

(3) The anomalous nature of the mental: there are no strict deterministic laws with which mental events can be predicted or explained (there are no strict laws that govern mental phenomena). This implies the irreducibility of the mental to physical phenomena and the autonomy of the mental.

(4) Each mental event supervenes or depends on a physical event.

The third principle appears to conflict with the first two as it denies their position. Davidson reconciles the principles by arguing that each mental event has a physical description and so each mental event is a physical event. Thus one event for Davidson can be described or characterized both in mental terms and physical terms. Davidson's main aim in terms of his anomalous monism is to show that mental properties are autonomous and there is mental causation in the physical world. That is, he wants to show that our reason can cause us to perform some actions.

Davidson's anomalous monism arguments implies that:

> If we are (or can be) rational and coherent it follows that we should believe that we are (or can be) free to respond to requirements of logic and evidence. This seems to capture a deep and fundamental intuition about who we are and the nature of mind: rationality implies autonomy, and autonomy implies freedom to respond to the normative principles of logic and evidence, objective moral standard, and so forth … Davidson represents this freedom to respond in terms of anomalousness

> … Davidson's core motivation [is] to provide a realistic and workable account of autonomy (Scharf, 2010: 344).

However, Kim argues that anomalous monism and Davidson's supervenience render mental properties epiphenomenal.

Kim's main argument is that mental events should be causally efficacious based on mental properties not on physical properties but anomalous monism does not show that, but rather that mental causation is dependent on physical properties. Kim, as a physicalist, believes that everything in the universe is physical or reducible to the physical. Mental entities cannot have causal influence in the physical world.

5. Kim and Davidson's Discussion of Mental Causation

Kim's notion of mental causation, which has been discussed in section 1.3.2, is overtly incorporated in his argument against anomalous monism (2001a:31) in his book, *Mind in a Physical World*: *An Essay on the Mind-Body Problem and Mental Causation*. For Kim, Davidson's anomalous monism is not a mind-body theory. Recall Davidson's argument. Causality in the physical world is nomic – so a mental event must instantiate a law to be causally efficacious in the physical world by being given physical descriptions as there are no strict psychological laws regulating mental properties. Hence, Davidson argues that mental events are physical events. This is the monism in his anomalous monism. His argument thus implies that for any event to be part of a causal relation it must be part of the physical domain. Mental events are causally efficacious therefore only because they are identical with causally efficacious physical events (Kim, ibid.: 33).

In objection to Davidson's argument, Kim argues that anomalous monism cannot account for mental causation. He (ibid.: 34) argues as follows: Assuming mental event *m* causes physical event *e*. For Davidson, this causal relation happens because *m* and *e* both instantiate physical laws. Thus, *m* must be given physical descriptions so that it falls under a certain physical kind (perhaps, neural) kind *N*, *e* falls under a physical kind *P,* and an appropriate causal law connects events of kind *N* with events of kind *P.* Kim (ibid.) argues that the causal relevance

of the mentality is in danger here because *m* as a mental event seems to have no role in determining the kind of causal relation it enters into. Event *m'*s causal relations are fixed, wholly and exclusively, by the totality of its physical properties, and the mental property instantiated by *m's* does not do any causal work. Thus, the mental property is epiphenomenal.

For Kim, due to the anomalism of the mental supposed by Davidson, Davidson needs to answer this question: How can anomalous properties be causal properties (ibid.: 33)? Kim holds that a solution to this problem is to show in contrast to Davidson's position, that mental properties are not anomalous or the anomaly of mentality in Davidson's sense is no barrier to being causally efficacious. For Kim,

> The trouble is that once we begin talking about correlations and dependencies between specific psychological and physical properties, we are in effect talking about psychological laws and these laws raise the spectre of unwanted physical reductionism. Where there are psychological laws, there is always the threat or promise of psychophysical reduction (Kim, 1989:42).

Another possible solution is to relax or circumvent the nomological requirement on causal relations. Thus, some philosophers such as Child (1992, 224) and the Macdonalds (1986) have come to the defense of Davidson (Yalowitz 2014: 49). They argue that the existence of a strict supervenience laws is not against mental anomalism in so far as we cannot comprehend such laws or use them to predict and explain actual mental events (ibid). Consequently, anomalous monism becomes a contingent epistemological position instead of the necessary metaphysical doctrine it claims to be (ibid.).

Davidson (2003) in his article, "*Thinking Causes"*, responds to Kim's criticism of his anomalous monism. As we have seen, Kim argues that Davidson's supervenience is inconsistent with the first premise of anomalous monism – the principle of causal interaction – (specifically, the claim that mental events cause or are caused physical events). Kim writes:

> The fact is under Davidson's anomalous monism, mentality does no causal work. Remember: on anomalous monism, events are causes only as they instantiate physical laws, and this means that an event's mental properties make no causal difference. And to suppose that altering an event's metal properties would also alter its physical properties and thereby affect its causal relation is to suppose that psychophysical anomalism, a cardinal tenet of anomalism is false (Kim, 1989: 6).

In response to Kim's criticism, Davidson (2003: 136) makes the distinction (mentioned above) between causation and causal explanation. For Davidson, causation is a relationship between individual events (spatio-temporally bound entities) no matter how those events are described. In terms of causation, one event causes another ('A' event causes 'B' event), while in terms of causal explanation we explain *why* one event causes another. Explanation is dependent on *how events are described* (ibid.). Physical events are described using terms in physics and mental events are described using terms in psychology. The events have the power to change or cause things, not their properties, but properties are helpful and needed in the explanation or description of the events. For Davidson, events *per se* do not cause one another in virtue of their properties contrary to Kim. Properties help in description and explanation. Physical properties (size, colour, weight, etc.) provide a physical description or explanation that are subject to physical laws while mental properties (belief, intention, choice, etc.) provide a rational explanation of causality which is not subject to any physical law but based on the principles of rationality and coherence (ibid.).

For Davidson (ibid.: 140), the fact that all mental events can be given physical descriptions makes them fully part of the physical world. Thus, for Davidson, mental events have causal relations with physical events and anomalous monism does not lead to epiphenomenalism.

> Suppose Magellan notices that there are rocks ahead, an event, that through the intervening events such as uttering orders to

> the helmsman causes the ship to alter course. Magellan's noticing is a mental event, and it is causally efficacious. That event is also is a physical event, a change in Magellan's body, and describable in the vocabulary of physics. As long as the predicates used to describe the mental event are not strictly reducible to the predicates of physics, all this is in accord with AM (ibid.: 141).

On supervenience, Davidson (ibid.: 137) holds that supervenience is about token events; not about the event types as Kim thinks. As token events, mental properties make a difference to what physical properties individual events would have. And if physical properties have causal power, then mental properties have the same causal power.

> Though supervenience entails that any change in a mental property *p* of a particular event *e* will be accompanied by a change in the physical properties of *e*, it does not entail that a change in *p* (mental property) in other events will be accompanied by an identical change in the physical properties of those events. Only the later entailment would conflict with *AM* ... Rather, supervenience helps in showing not only that *AM* ... is consistent, but also there is a version of *AM* ... that gives a possible relation between the mental and the physical... (ibid.).

According to anomalous monism, causation or causality entails deterministic laws and there is no psychological law between the mental and physical. Davidson makes the distinction between strict law (homonic) and non-strict law (heternomic). Strict laws are found in physics and non-strict laws cover mental events, and the non-strict laws cannot be reduced to strict laws because features of mental events are different from the features of physical events.

Davidson (2003: 139) argues that Kim gives no reason anywhere why all laws at issue must be strict while he himself has numerous times argued for the distinction between strict

and non-strict laws. The former is due to the nature of physical events and the latter due to the holistic and rational nature of mental events. He argues that Kim (1989), Fodor (1989) and Dretske (1989) all fail to notice the difference between strict laws (laws of ideal physics) and non-strict laws (generalizations that we might call laws). Davidson believes that Kim, Fodor and Drestke are wrong to think that without psychophysical laws, the mental is causally inefficacious or inert. Davidson, like the trio, believes there must be laws linking the mental and physical events but holds that such laws are not strict, and mental events are not reducible by definition or reduction to physical events. Thus, their arguments against the causal efficacy of the mental (ibid.: 141) do not hold.

However, Kim argues that the mental plays no causal role because under anomalous monism, events are causes and effects only if they instantiate physical laws and it is then only under physical descriptions that mental events can have causal relations with physical or mental events. The mental as mental does not cause anything (Kim, 1984: 267, Davidson, ibid.: 141). In response, Davidson says "... It is events that have causes and effects. Given this extensionalist view of causal relations, it makes no sense ... to speak of an event causing something as mental or by virtue of its mental properties or as described in one way or the other" (ibid.). For Davidson, causation is extensional – it holds between events regardless of how they are described – while explanation is intensional as it is sensitive to how events are described (Yalowitz, 2014: 16). He invites Kim and other critics to prove the impotence of mental properties, and for him it is impossible to do that because under supervenience mental properties do make a difference to other events. That mental properties supervene on physical properties guarantees the explanatory relevance of the mental properties.

> ... Supervenience ... implies that if two events differ in their psychological properties, they differ in their physical properties. If supervenience holds, psychological properties make a difference to the causal relations of an event, for they matter to physical properties, and the physical properties matter to causal relations (ibid.: 142).

For Kim (1989:39-41), it is difficult to comprehend Davidson's notion of supervenience because it is implausible to reconcile the irreducibility of the mental to the physical with the dependence of mental on it. Davidson notes that the whole confusion about anomalism and supervenience is a failure of understanding the distinction between event particulars (token identity) and event types (type identity). Moreover, this failure results in mixing up singular causal relations with causal laws, which subsequently results in ignoring the difference between causality (simply stating that a causal relation holds between two events) and causal explanation (explaining an event whereby laws and properties are used) (Davidson, 2003 : 143).

In the article, "*Can Supervenience and 'Non-strict Laws' Save Anomalous Monism?"*, Kim (2012) continues to argue that anomalous monism is a form of epiphenomenalism as it does not attribute any causal role to mental properties; it says a lot about physical properties but little or nothing about mental properties. Hence, it entails for the epiphenomenalism of mental properties. Moreover, Kim argues that the fact that Davidson in proving the causal efficacy of the mental supplements anomalous monism with supervenience and non-strict laws, shows that within the context of anomalous monism mental properties have no causal role (Kim, 2012: 236).

Kim here (ibid.) accepts anomalous monism's distinction between causality and causal explanation, but argues that it seems that still the mental cannot causally explain physical events. Thus, again the causal efficacy of mental properties is in question. For Kim, anomalous monism is about a metaphysical causal relation, and causal explanation is more than a metaphysical relation. Causal explanation demands to know in virtue of what property one event causes another. Moreover, Kim argues that Davidson denies the mental a role in explanation if it is not in virtue of a mental property that one event causes another. Hence, mental properties are causally inert. The causally efficacious properties (physical properties) in causal explanation will exclude the other inert properties (mental properties) (ibid.: 241). This is what Kim terms "the problem of causal-explanatory exclusion".

Davidson (2003: 144) argues that mental causes are not independent of physical causes since mental events are physical events under the supervenience thesis. However, mental causal explanation can be logically independent of physical explanation because mental concepts are not reducible to physical concepts. At the same time, physical explanation cannot exclude mental explanation. Each of the kinds of explanation logically has its own area of interest, as the mental sphere is different from the physical sphere. However, physical explanation can only exclude mental explanation when causal relations, which hold only between particular events, nomologically are taken as synonymous with causal explanation. In this case, one can accept the principle of 'causal-explanatory exclusion'.

Kim continues to argue in his essay, *"Can Supervenience and Non-Strict Laws Save Anomalous Monism?*", that the supervenience thesis supports the causal relevance of mental properties, and not their 'causal efficacy' (Kim, 2012: 239). Mental properties are 'relevant' to causal explanations since what mental properties an event has, affects what physical properties an event has in terms of supervenience; nevertheless, it is the physical properties of events that are causally efficacious. Kim (ibid.) emphasizes that for Davidson to defend his position in "*Actions, Reason and Causes"*, he will need to prove the causal efficacy of the mental rather than its causal relevance, which may be too weak to defend the 'because' in rationalizing explanation. In addition, those who argue for mental causation would work with causal efficacy instead of causal relevance of the mental.

6. Kim's Argument against Supervenience in terms of Mental Causation and his Functional Model of Reduction

For Kim, the issue of causal exclusion is another problem for mental causation. Kim – like most physicalists – holds that the physical universe is causally closed. In other words, every physical event has a complete, independent cause in terms of another physical event or a single, complete and independent causal explanation. Hence, for a mental event *m* to cause a physical event at *t* would be a clear violation of the causal closure of the physical domain, a relapse into

Cartesian interactionist dualism that mixes physical and non-physical causes in every single causal chain (Kim, 2001a: 37).

Moreover, there is also the principle of exclusivity or non-overdetermination to deal with – one event cannot have both mental and physical causes. And if *p* has also a physical cause p^x at *t*, then the question is what causal work *m* is doing, since *p* has a physical cause p^x. (ibid.) "The physical cause therefore threatens to exclude, and preempt, the mental cause. This is the problem of causal exclusion"(ibid.). The work of the antireductive physicalist who wants to remain a mental realist, is to explain how the mental cause and the physical cause of one and the same event are related to each other. For Kim, token physicalism, like Davidson's anomalous monism, is inadequate to do that as the ultimate question is about the causal efficacy of mental properties, and AM holds that mental properties are not reducible to physical properties (ibid). "…the problem of causal exclusion is to answer the question: Given that every physical event that has a cause has a physical cause, how is a mental cause also possible?" (Kim, ibid.: 38). Kim actually holds that the exclusion problem faults mind-body supervenience, and while supervenience and physical realization (mental properties are realized in or by physical properties) can create difficulties for mental causation, physical realization as a functional reduction is the solution to mental causation. Kim takes supervenience, physical realisationism and emergentism to basically mean the same thing:

> In my first lecture, I argued that mind-body supervenience could usefully be thought of as defining minimal physicalism – that it is the minimal commitment that anyone who calls herself a physicalist should be willing to accept. We saw also that mind-body supervenience is entailed by physical realizationism, the thesis that mental properties are instantiated in virtue of being realized by physical properties in physical systems. Moreover, emergentism, too, is arguably committed to mind-body supervenience: if two systems are wholly alike physically, we

> should expect the same mental properties to emerge or fail to emerge, in each (ibid.).

Be that as it may, the real source of the problem of mental causation for Kim is the supervenience of the mental on the physical. Mind-body supervenience, in the context of causal closure and the principle of exclusivity, makes mental causation unintelligible. Kim's mental causation argument against mind-body supervenience is in the form of two horns of a dilemma:

(1) Either mind-body supervenience holds, or it fails (ibid.: 39).

Kim (ibid.), before proceeding with his arguments, restates the mind-body supervenience thesis:

"Mental properties supervene on physical properties in the sense that if something instantiates any mental property *M* at *t,* there is a physical base property *P* such that the thing has *P* at *t*, and necessarily anything with *P* at a time has *M* at that time" (ibid.).

On the second horn of the dilemma, "if the mind-body supervenience fails", there is no visible way of understanding the possibility of mental causation" (ibid).

Kim notes that Jerry Fodor argues that without supervenience, mental causation is unintelligible, while Horgan holds that physical supervenience of qualia makes qualia causally efficacious. The mechanism of supervenience is required in non-reductive physicalist approaches in order to circumvent the causal closure of the physical domain, the thesis that no causal chain will ever cross the boundary between the physical and the nonphysical. Supervenience is brought in, to give mental properties physical characteristics that are at least nomologically sufficient for them to be causally efficacious in a world which is fundamentally physical. Thus, every mental event or sensation or intentional state must have a physical basis; it occurs because a suitable physical base is present (Kim, ibid.: 40). However, if mind-body supervenience fails, the mental domain floats freely, unanchored in the physical domain, and

then causation from the mental to the physical will violate the principle of physical causal closure (ibid.). Thus, without supervenience mental causation has no chance to make sense.

For Kim (ibid.), nevertheless, mind-body supervenience brings mental phenomena within the domain of the physical for the physical determines the mental and as such the mental is not independent of the physical. The question is, does mind-body supervenience bring the mental close enough to the physical so that the mental can avoid the causal closure of the physical domain? If the answer is negative, then mind-body supervenience is inadequate to be a solution of the causal exclusion of the mental by the physical. And mind-body supervenience instead of being a solution, may be the elephant in the room to the disappointment of Fodor and Hogan. For Kim, this is possible for the following reason, as explained already:

> Suppose that an instantiation of mental property *M* at *t* causes another mental property M^x to be instantiated an instant later, at t^x . Given Supervenience, M^x has a physical supervenience base Px such that P^x is instantiated at t^x, and given this, M^x must of necessity be instantiated at t^x. If P^x is there, M^x must be there no matter what has preceded the instantiation of M^x. Moreover, without P^x, or an alternative physical base, M^x could not be instantiated at t^x. This threatens M's claim to be the cause of M^x*'s* instantiation at t^x; for even if *M* had not occurred, M^x would still have occurred as long as its supervenience base P^x was present at t^x. (This argument can be formulated in terms of realization, with "X supervenes on Y" replaced everywhere with "X is physically realized by y".)(Kim, 2007:411).

This argument rests on the principle of causal exclusion. Moreover, for Kim, causal over-determination is possible, that is the thesis where an event has two different equal causes. However, if an event has two causes – a mental cause and a physical cause – one cause must go. In this case, the mental cause must go, because according to the causal closure of the

physical, an event must have only one physical cause. Hence, there cannot be a mental cause for a physical event. And again, M_1 and P_1 cannot both be genuine causes of M_2. For Kim, "*M*-*M** causation … is rather like a series of shadows cast by a moving car: there is no causal connection between the shadow of the car at one instant and its shadow an instant later, each being an effect of the moving car. The moving car represents a genuine causal process, but the series of shadows it casts, however regular and law-like it may be, does not constitute a causal process" (2000.: 45). We thus have this scenario: the M_1 - to - M_2 and M_1- to - P_2 causal relations are only apparent, arising out of the genuine causal process from P_1 to P_2.

Thus, again, if we combine the principle of causal closure and exclusivity with supervenience, we get the following scenario (see Kim 2001b: 271-283): if we assume again that an event, which is an instantiation of a property K, causes an instantiation of property K*, can the one event be a cause of the other? It seems not so bad until we bring in the argument for supervenience, the principle of causal closure and the principle of exclusivity or the rule against causal over-determination: According to supervenience we cannot have the event instantiating K* without an event instantiating property P* on which the K* event supervenes. So, if we have event P*, we must have event K*. But then the K* event is caused both by the K event and the P* event and that is against exclusivity. Did perhaps then the K event cause the P* event to cause the K* event? This would imply that mental to mental causation needs mental to physical causation but this cannot be right because of the causal closure of the physical world. So, the mental is causally impotent. Thus, if the argument for supervenience holds, mental causation cannot be explained or defended.

For Kim, then, both the failure or success of the mind-body supervenience argument makes mental causation unintelligible (Kim, ibid.: 46). Kim also holds that mind-body supervenience cannot account for how our mentality is related to our physical nature, in other words, mind-body supervenience is not a philosophical theory of how minds are related to bodies. Kim has two related reasons for his position. First, mind-body supervenience is consistent with the classic positions on the mind-body problem such as emergentism and physical realizationism (that the mind is physically realized), as already mentioned.

Emergentism is a form of dualism that takes mental properties to be nonphysical with intrinsic causal powers, whereas for Kim physical realizationism is monistic physicalism. Moreover, epiphenomenalism that is in contention with physicalism is also committed to mind-body supervenience. If mind-body supervenience is consistent with these diverse and conflicting mind-body theories, that all have difficulties in dealing with the mind-body problem, then surely supervenience will somehow reflect these difficulties (ibid.: 12).

Moreover, for Kim, mind-body supervenience does not show *why* mental properties supervene on the physical properties; hence, it is not an explanatory theory either. Rather, it merely states a pattern of property co-variation and a dependence relation between the mental and the physical. However, it is silent on the nature of the dependence relation between the two, which can explain why the mental supervenes on the physical (ibid.: 14).

Thus, according to Kim, mind-body supervenience deals with a phenomenological relation as it is about the relation of a particular mental event with a particular physical event instead of the relationship obtained in type-type identity. So it deals with the patterns that are manifestations of some deeper dependence relations covered by causal dependence or reductive dependence. Consequently, for Kim (ibid.), the mind-body supervenience argument does not solve, but rather states the mind-body problem. The proponents of non-reductive physicalism must seek elsewhere the metaphysical ground or basis for their mind-body theory; supervenience cannot provide that. Any mind-body supervenience relation must be grounded on a metaphysics that explains the mind-body supervenience (ibid.).

I agree with Kim's challenge to the proponents of non-reductive physicalism to look elsewhere for metaphysical grounds for their theory. Supervenience, just like reductive physicalism, cannot account for the metaphysical grounds or origin of mental properties. One thing that Davidson has made abundantly clear is that mental events because of their holistic, normative and rational nature are quite different from physical properties. If they are quite different, how can a physical property cause the rise of a non-physical property that has different qualities from it? Why these new properties are not called physical properties? Davidson did not state the principle of supervenience within the framework of the first three

principles (interaction between mental events and physical events; where is there a causal relation, there must be a deterministic law; and there is no psychophysical law between the mental and the physical); it was brought in later. But he does not show how supervenience is derived from the first three principles or its consistency with the principle of anomalous monism. And if supervenience is dependent on physical properties, it gives no primacy to mental properties, but to physical properties. It makes causality at the micro-level more fundamental, and causality at the macro-level dependent on the one at the micro-level. Hence, the mental will be governed by physical laws. This casts doubts on the efficacy and autonomy of the mental. Thus, the supervenience argument cannot account for the metaphysical grounds of the mental properties**.** It is true that Davidson's distinction between causality and causal explanation takes away some of the sting of the problem, but I argue we need more metaphysical meat than making interpretative distinctions between *descriptions* of properties. In the next chapter, I will show that the solution to accounting for such metaphysical grounds for mental properties in relation to physical properties is Aristotle's theory of hylomorphism. Aristotle's notion of 'form' might be the 'non' in the non-reductive physicalism.

In terms of supervenience and mental causation, Kim concludes that mind-body supervenience does not solve mental causation; it makes the mental epiphenomenal (lack of causal efficacy). And for Kim (2007: 412), to argue that the mental is epiphenomenal is against our general belief about ourselves that we are cognizers and agents in the physical world.

> To be an agent means being someone who, in light of his beliefs, and desires and intentions, can causally act on the world. If our preferences, beliefs, and intentions have no causal role in moving our limbs and thereby cause things around us to be rearranged, how can we view ourselves as agents, people who can do things…

Kim holds that an epiphenomenalist conclusion could be avoided if we accept mind-body reduction (physical reductionism), that is, allowing the mental causes or properties or instantiations to be physically reduced to the physical (ibid.), thereby making it occur within the

physical domain, and as such then physical causal closure will not be a problem. But for Kim, this is not an argument for reductionism rather it is a hope that reductionism is correct, i.e., that the mental can indeed be physically reducible and that is the only way to vindicate the causal efficacy of the mental (ibid.).

More specifically, Kim argues that mental properties can only be causally efficacious through functional reduction. Let us now consider Kim's notion of a functional model of reduction. The functional model of reduction is an aspect of functionalization. And functionalization states that things are defined in terms of the function or role they play. That mental properties are functional, implies that their essence or nature or importance is comprehended from the perspective of the role they play within a system as we noted earlier in the previous chapter. There are various versions of functionalization and they are distinguished according to their roles or functions. A property's role might be defined based on causal input-output relations or in terms of its relation to computational states (computational functionalism) or in terms of its role in relation to a goal-directed process or system (teleological functionalism) (Van Riel and Van Gulick, 2019: 32). One of the demerits of functionalism, like any other physicalism, is that it is inadequate to account for the reality of qualia or subjective experience.

The functional model reduction holds, "… reduction as consisting of a relation between a functional property *P* and a structural or substrate property P^x that plays the role determined by the functional property *P*, or at least does in a restricted class of systems or contexts" (ibid.). According to Ned Block (forthcoming: 1), in terms of functional reduction, the first step is to functionally characterize the property to be reduced, and the second step is to discover the physical property that plays that role. This kind of reduction explanation in one form or another (notably as identity reduction) is noticeable in the writings of Smart, Armstrong, Lewis, Kim, Levine, Chalmers and Jackson (ibid.). We noted earlier in section 2.6.2 that functionalization is of two types namely realizer functionalization (it identifies the mental state-type-pain with whatever occupies the causal role) and role functionalism identifies pain with the role itself. Kim is identified as realizer functionalist like Armstrong and Lewis.

Kim (2001a), in discussing the functionalization of the mental, responds to the criticism that his supervenience argument against mental causation should apply to all other special science properties. This is what is termed the generalization argument. If the mental lacks causal power even if it supervenes on the physical base that constitutes the causal power, then, special sciences such as biology that supervene on micro particles should also be epiphenomenal as their causal powers are derived from a subvenient or realizer physical base. For Kim, a mental event *M* can retain its causal status if it is reducible to a physical event *P*; a given instantiation of *M* must be reductively identifiable with the instantiation, at that particular time, of its subvenient or realization base.

This kind of reduction for Kim is not Ernest Nagel's model of reduction that defines reduction as an inter-theoretic relation through bi-conditional[9] bridge laws "… which provide the essential reductive links between the vocabulary of the theory targeted for reduction and that of the base theory, and thereby enable the derivation of the target theory from its reduce" (Kim ibid.: 90). Putnam and Fodor's multiple realization argument shows that Nagelian bridge laws cannot be available. If a higher order property *P* has multiple realizers in lower order properties $Q_1, Q_2, Q_3, \ldots$; it is not possible to provide *P* with a single lower-order correlate *Q* to yield a bi-conditional bridge law (ibid.: 92).

There is also an explanatory question to be taken into account, i.e., "can we understand why we experience pain when our C. fibers are firing, and not when our A .fibers are firing "(ibid.: 95)? Why does a certain mental state occur at the occurring of certain neuronal activity? Kim takes the explanatory question seriously, and comments that "it isn't that on physicalism every phenomenon must be physically explainable; there may well be physical phenomena that aren't physically explainable..." (ibid.: 96). Furthermore, Davidson's anomalism argument also shows that the Nagelian model of reduction is not possible. In Davidson's terms, the mental is anomalous; there are no psychophysical laws connecting the mental with the physical.

[9] The bridge laws take a bi-conditional form, providing for each primitive predicate of the theory to be reduced with a nomologically co-extensive predicate in the base theory (Kim 2001a.: 90).

Thus, based on the arguments such as Davidson's anomalism of the mental and multiple realization, Kim notes that there "is no a single neural substrate for a mental state … it is not possible to reductively identify the mental state with a neural state or mechanism. Since pain has multiple neural realizers, N1, N2, …, it is not possible to pick one of these, say N1, and claim that pain = N1. (Kim, 2007: 412). As Kim is keen on saving the mental from epiphenomenalism, he proposes his own reduction that is different from identity reduction. "But identity reduction is not the only kind of reduction; there is also functional which, if it can be carried out for the mental vis-á-vis the physical/neural, could vindicate the causal role of mentality" (Kim, 2007: 413).

Hence, Kim argues that Nagel's bridge laws are not necessary or sufficient for proper reductions. He (2001a: 97) goes on to propose a functional reduction model instead. One might be surprised that Kim is using the theory of functionalism that is generally taken to be anti-reductionist to solve mental causation on reductive physicalist terms. He appeals to functionalism however because, according to him, it shares the same metaphysical base with physicalism. (This is denied by writers such as Block (forthcoming`: 13). Kim writes:

> ... If the functionalist conception of the mental is correct – correct for all mental properties – then mind-body reduction is in principle possible, if not physically practically feasible. This is contrary to one piece of current philosophical wisdom, the claim that functionalism, as distinguished from classic type physicalism, is a form – in fact the principal contemporary form – of mind-body anti-reductionism. What I am urging here is the exact opposite – that the functionalist conception of mental properties is required for mind reduction. In fact, it is necessary and sufficient for reducibility. If this is right, mind-body reductionism and the functionalist approach to mentality stand or fall together; they share the same metaphysical fate (Kim, 2001a: 101).

For Kim, in terms of the functionalization of the mental, a mental property is first explained as a second-order functional property in terms of its functional property causal role (functional definition). This is followed by finding its realizers – the first-order (physical) properties that play the causal role as stipulated in the functional definition – and then there must be an explanation of how that realizer of the property actually performs the specified causal tasks. The functional model of reduction shows that a functional reduction of the mental can explain mental causation on physicalist terms. According to Kim (ibid.: 98-99), thus the functional model is as follows:

> … to reduce a property M to a domain of base properties, we must first "prime" M for reduction by construing, or reconstruing , *it relationally or extrinsically*. This turns M into a relational/extrinsic property. For functional reduction, we construe M as second-order property defined by its causal role – that is, by a causal specification H describing its (typical) causes and effects. So M is now the property of having a property with such and such causal potentials, and it turns out that property P is exactly the property that fits the causal specification. And this grounds the identification of M with P. M is the property of having some property that meets specification H, and P is the property that meets H. So M is the property of having P. But in general the property of having property Q = property Q.

The second-order properties are functional properties but predicates, expressions or concepts do not bring into existence new properties; rather, they are a new way of talking about already existing properties (ibid.: 110). For Kim, mental properties are causally efficacious if they are functionally reduced; "the functionalization of mental properties enables them to escape the supervenient argument" (ibid.: 116). This is because functionalized mental properties do not bring in a new causal power in the system beyond those of their first-order realizers. In Kim's model, functional properties are nothing over and above the physical

properties. Kim thus advocates physical monism as the solution to the problem of mental causation.

For Kim, mental properties can only be causally efficacious with regard to physical properties or mental properties if they are reduced to physical properties. So, to some extent, this is also an indirect way in which to assign causal efficacy to mental properties, via the physical, as Davidson also does, albeit in a totally different way. However, in providing a solution to mental causation, Kim holds that a further step should be taken to show that mental properties are actually physically reducible. "This means that the question "Is the mental physically reducible?" is transformed into this one "Is the mental functionalizable?" – that is, Can mental properties be given functional characterizations" (Kim, 2007: 414). And he thinks they can.

Kim (ibid.) notes that mental phenomena are normally divided into two, namely, intentional phenomena and sensory/qualitative phenomena. Intentional phenomena are also known as propositional attitudes (beliefs, desires, and intentions), while the sensory/qualitative phenomena are the states and what it is like to be in them. According to Kim, the division is not watertight in that there are content-carrying states that are also qualitative in nature such as feelings and emotions (ibid.). So for Kim "The question about the causal efficacy of consciousness, then, is a question concerning the phenomena of the second category. But it is also a question about those intentional states with a qualitative aspect" (ibid.). Kim holds that through functional reduction cognitive and intentional properties such as belief, desire, and perception are functionally reducible. For Kim (ibid.: 415):

> Intentional states are functional states specified by their job descriptions. To believe that it is going to rain later today is to be in a state *S* such that if you are in *S* and are asked "What the weather going to be like today?" S will cause you to answer "It's going to rain"; moreover, if you are in *S* then, given a desire not to get wet, *S* will probably cause you to take an umbrella to work, and so on. We can be sure that there will not likely be a

> full and complete functional definition of believing that it is going to rain, but given the supervenience of this belief on behaviour, there can be no fact about this belief that goes beyond actual and possible behaviour. To begin a functional reduction of this belief, we can start looking for a neural mechanism that does the causal work so far specified. In scientific terms, belief may be nothing more than the storage of information, in an accessible and retrievable form; which can be used for inference and the rational guidance of behaviour. This is a functional conception, a conception of belief in terms of its work in the cognitive /psychological economy of a psychological subject.

However, qualia (sensory qualities of conscious experience) are not reducible since they resist functionalization because what a qualitative state such as pain (though caused by certain stimuli and it, in turn, causing a broadly definable set of behaviours) is, cannot be its job description/functional role, but is the fact that you can feel or experience it, " …what makes pain, pain, is the fact that it hurts, not how the pain mechanism is wired to inputs and outputs, then you will deny that pain, and other qualia, can be given a functional characterization. If this is correct, qualia are not functionalizable and hence, functionally irreducible" (ibid.:416).

For Kim, properties that are not physically reducible are epiphenomenal. There are two ways of thinking about reduction, namely, identity reduction and functional reduction. But none can reduce the qualitative aspect of consciousness (phenomenal consciousness). Is phenomenal consciousness epiphenomenal then, as it cannot be functionalized? Kim holds that qualia as absolute intrinsic qualities are irreducible while qualia similarities and differences appear functionally characterizable and hence reducible (ibid.).

> The qualia similarities and differences are behaviourally manifest and this opens the door to their functional characterization. This means that what is really important about

> our perceptual experience – that is, cognitive role – can be functionally characterizable and hence shown to be causally efficacious. The conclusion, therefore, is that epiphenomenalism cannot claim a total victory over qualia; we can save from the epiphenomenalist threat a crucially important aspect of qualia – that aspect of qualia that makes a difference to our cognition and capacity to cope with the world. (ibid.)

Through functional reduction, intentional and cognitive properties are causally efficacious (basically by reducing the functional causal role of the mental property to its physical realiser), though Kim admits that functionalization of qualia or the causal efficacy of the phenomenal character of experience remains a problem. Kim proposes that we can just live with epiphenomenalism of qualia at present until their causal efficacy becomes explainable (2001a**:** 102) as science progresses.

If phenomenal properties resist functionalization, it means that they are irreducible to the physical and, hence, they are epiphenomenal, because they lack causal power that resides in physical properties, according to Kim. The question is what is the ontological foundation of these phenomenal properties? How do the physicalists explain their metaphysical ground? Physics seems to fail to account for the origin of the phenomenal properties. Thus, there is a fact that physics cannot explain, and it faults the causal closure principle (the universe is not physically closed) for if indeed, the universe is physically closed nothing can resist functional reduction. Kim does not agree, as he is confident future science will allow us to find ways to functionally reduce also phenomenal properties.

Having considered Kim and Davidson's discussion, let us now consider some criticisms of their discussion. I will first look at criticism and reflection in the context of Davidson's theory of mental causation and in the next section do the same for Kim.

7. Criticism and Reflection on Mental Causation in the Context of Davidson and Kim

7.1. Criticism and Reflection on Davidson's Notion of Mental Causation

Daniel Hutto (1999:381), in the article, "*A Cause for Concern: Reasons, Causes and Explanations",* argues against Davidson's theory of causalism or causality that holds that reasons are causes of action. Hutto aims to establish the inadequacy of Davidson's theory of causalism in the study of reason-explanation and find a replacement understanding of 'mental causation'**.** Recall that reason-explanation is a species of causal explanation that designates a reason as a rational cause (Davidson, 1980:233). Firstly, Hutto looks at Davidson's claims that reason-explanation is a species of causal explanation and notes that Davidson's view does not sufficiently address how we determine for which reason we act. Secondly, he examines anomalous monism in terms of epiphenomenalism and establishes that the former offers no solution to the mysterious connection between reason and action. Thirdly, he argues that due to the incompatibility of reductive physicalism and special features of reason explanation, we should eliminate causalism about reason and reason explanations. He proposes a new way of rereading causation and for the elimination of the mysterious connection problem. For my purposes I will briefly consider Hutto's second point.

In terms of epiphenomenalism, Hutto (ibid.: 383) notes that for some philosophers, psychological phenomena cause and are caused by physical phenomena, and reasons are causally efficacious and warrant causal forms of explanation. He points out that Davidson in his article *Actions, Reason and Causes* argues that reasons could be causes and that reason-explanation must be understood as a species of causal explanation. Davidson notes that there is a difference between citing *a* reason (primary reason) and citing *the* reason for action. In executing an action, one may have more than one primary reason for acting, but often only one reason is the main reason for acting. For us to explain why a person acted, we have to choose a belief/desire pair out of many alternative belief/desire pairs, which could make sense for the act (Davidson, 1987: 42).

The reason for acting designates the mental events (belief/desire pair) that actually were responsible for or caused the particular action; the mental events that were efficacious in producing the action (Davidson, 1980: 232).

Hutto (ibid.) furthermore notes that Davidson's primary motivation for a causalist treatment of 'because' in reasons-explanations is derived from his admiration of Aristotle's theory of action. More accurately, perhaps one can just agree that Davidson is simply agreeing with Aristotle that reasons are desires rather than accepting the modern tendency of viewing reasons in terms of intentions. Davidson notes that it was Aristotle who introduced the concept of wanting as a causal factor (Davidson, 2001: 11). And, in that sense, "Aristotle's account ... alone promises to give an account of the 'mysterious connection' between reasons and actions" (Davidson, ibid.: 11).

Davidson's causal theory which we have noted as a causal theory that falls under the standard action theory implies that action is an action only if it is caused by desires, beliefs and intentions in the right way. Actions are explained in terms of the agent's reasons (as desire/belief pairs) or in terms of rationalizing mental states and events, and we recognize or identify something as an action only under its intentional description (Davidson, 1963), which can be functionalised.

Schlosser (2019) claims that the standard theory fails to capture human agency. For on the standard action theory, "… a being has the capacity to act intentionally just in case it has the right functional organization: just in case the instantiation of certain mental states and events (such as desires, beliefs, and intentions) would cause the right events (such as certain movements) in the right way. According to this standard theory of agency, the exercise of agency consists in the instantiation of the right causal relations between agent- involving states and events" (Schlosser, 2019: 3). It fails to capture agency, because it reduces actions to mere happenings. These happenings are the events of physical elements in the brain.

> I think that the standard story is flawed in several respects. The flaw … is that the story fails to include an agent – or, more

> precisely, fails to cast the agent in his proper role. In this story, reasons cause an intention, an intention causes bodily movements, but nobody – that is, no person – does anything. Psychological and physiological events take place, but the person serves merely as the arena for these events; he takes no active part (Velleman, 1992: 1).

It seems the issue is that the phenomenal aspect of consciousness of the agent is not involved in the action, hence the agent does not know what it means to be in any state or event of the action. The agent is not *engaged* in his action as action is something happening to him. In this case, it will be difficult for the agent to have a future goal or be goal oriented in planning action. One has to acknowledge against these critiques however that Davidson does address agency, despite his role functionalism, in the sense of his argument that reasons are desires. Perhaps what is needed is to explore the relation between desire and belief on the one hand, and phenomenal consciousness on the other.

7.2. Criticism and Reflection on Kim's Notion of Mental Causation

Jacob (2002) in the article "*Some Problems for Reductive Physicalism*", notes that Kim's book, *Mind in a Physical World*, is a powerful critique of non-reductive physicalism. Non-reductivists are physicalists, thus they reject Cartesian substance dualism. However, they are property dualists as they reject identification of mental properties with physical properties. They profess the autonomy of the mental (ibid. 648). Functionalism teaches that mental properties are higher-order in that they specify roles for properties that are multiple realizable in lower-order physical properties. According to functionalism, once a mental property is functionalised, it possesses all the qualities of the lower properties that make it causally efficacious. Putnam used the multiple realization argument to criticize reductive physicalism, as pointed out above. Putnam argues that mental properties are not (type) identical to lower properties that reduce them, because mental properties can be realized in different physical properties (ibid.: 648). Thus, orthodox functionalism is a kind of non-reductive physicalism that is committed to the causal efficacy of the mental. But Kim argues in the book that orthodox functionalism cannot

profess that mental properties are irreducible (or autonomous) and causally efficacious at the same time.

Jacob (2002: 649) notes that "Kim's argument is directed against anti-reductionism, not against the functionalist conception of mental properties". Through functionalism, a functional property shares in the causal powers of the physical property that realizes it. In other words, functionalism saves mental realism. Mental properties lack causal power except that of the causal power of its physical realizer. Thus, there is no autonomy of mental properties (mental realism is saved) (ibid.: 649).

Jacob however argues that Kim does not make an impressive argument against the autonomy of mental properties. And he doubts the possibility of Kim's reductionism vindicating mental realism. For Kim, mental properties enjoy causal efficacy through functionalization, because once they are functionalized they inherit the physical properties of the lower physical realizers on which they supervene; if they are not functionalized they become epiphenomenal or eliminated. Can this be called mental realism? The mental properties are not consistent in their causal efficacy; they are causally efficacious now and then depending on the nature of the mental states and events and particular causal work involved.

Jacob further argues that if physical properties are adequate on their own to cause an effect within a system, without interference from any mental properties, then there is no need for the causal exclusion argument, and then Kim's supervenience argument is undermined (ibid.: 654). Kim (2002: 673), in response to Jacob's review, praises him for correctly outlining his basic arguments against non-reductivism particularly the supervenience argument. He notes that Jacob agrees to most of his ideas, but not all.

Kim (ibid.) holds that mental to mental (M-M^x) causation (the same level causation) implies downward causation from M-to-P^x. However, Jacob argues that the relation between P^x and M^x is not a proper causal relation, consequently there is no causal competition between the M-*to*-M^x and the P^x-*to*-M^x relation. For Kim (ibid.: 673), Jacob is correct that the P^x-*to*-M^x relation is not a causal type. And his claim that for that reason the causal exclusion principle

does not apply is correct too. Based on that, we cannot use it to argue that the two relations exclude each other. But Kim disagrees with Jacob on the fact why there is still tension about the origination of M^x; is the M^x instance there because of M or because of P^x? Kim holds that his argument that same-level causation implies downward causation, is not obliterated. There is a genuine tension between the two claims, one causal and one based on supervenience of *M* because that does not remove the tension about the origination of M^x, a tension Kim believes needs to be resolved.

8. Conclusion

Kim is a reductive physicalist. His four principles of mental causation imply mental properties have to be reduced to physical properties for them to be causal efficacious in the physical world. The mental cannot cause things *qua* mental. As a realizer functionalist, he made effort through his functional reduction model to functionalize certain mental states (desire, beliefs and intention) to be causally efficacious in causing action and cognition in humans. However, he shares in the weakness of functionalism to account for phenomenal consciousness adequately.

Davidson's utmost aim is to argue for the mental causal efficacy in the physical world, and for Kim, mind has no place in the physical world unless it is reduced to physics. Action, in functionalist terms, consists of input as causal relation among the mental states and events and production of behaviour as an output in the agent. There is no involvement of the agent in the sense of what it is like to be that agent, even given Davidson's view that reasons, as desires, are causes for action. To my mind, any true and adequate solution to the mind-body problem should be able to account for the reality of phenomenal consciousness as causal *qua* mental that carries the agent along. I will argue in the next chapter that Aristotle's theory of hylomorphism may have the answer to doing just that.

The failure of modern views of mental causation to account for human action in a full blooded manner adequately in that it ignores the phenomenal consciousness and human agency may be a clarion call for a new understanding of mental causation. Hutto and Yablo are two writers who seem to have hinted something like this. Yablo (2002: 190) writes:

> But maybe we are not misled to think that outcomes effected by the physical antecedents are neither speech nor action, nor expression of any sort of human agency. Maybe the mistake was to think that outcomes of the kind normally credited to human agency are caused by the physical antecedents (ibid.).

Hutto (1999.: 395) notes that the way Aristotle would deal with the mysterious problem of finding the connection between reason and action is quite different from any Humean based version of causation.

> To ally oneself with Aristotle on the issue of the way in which reasons are causes is to re-enforce that 'reasons' gain their explanatorily force by consideration of the agent and not by any reconciliation with impersonal causal perspectives. It is useful to remember that Aristotelian causes and their modes of explanation are not causal in the modern sense of the term (ibid.).

Our knowledge of causality is derived firstly from our capacity to project our human imagination and thinking on to the world. A first attempt to extend the notion of cause to inanimate objects would have, at least initially, involved a projection of humanlike agency onto those objects. Thus, the origin of our notion of causality does have an anthropomorphic basis (Strawson, 1985: 124-125). The reason for action or reason-explanations should thus not be treated as on par with other forms of causal explanation used in the natural sciences (Hutto, 1999: 397). Melden, A. I. (1961 : 208) writes: "It is certainly true that we use 'cause' in speaking about the action of agents, but we can no more infer from this verbal consideration that actions are the Humean effects of an event than we can from the etymological derivation of the term 'motive". This implies that Davidson's theory of action that is Humean and physically oriented cannot explain human agency.

I will discuss Aristotle's theory of hylomorphism and mental causation in the next chapter.

Chapter 4: Ancient Discussion of Mental Causation - Aristotle

1. Introduction

In chapter one, I defined the playing ground for my thesis by defining some key terms such as mental realism, which is the view that mental properties have independent causal power contrary to Kim's notion that the causal power of the mental resides in their physical properties. I indicated that the causal efficacy of mental phenomena under investigation in this thesis is that of phenomenal consciousness. For me, phenomenal consciousness can, *qua* mental, cause both physical and mental events. In chapter two, I showed that one of the main problems of the physicalistic theory is the inability to account for mental causation as it has no place for phenomenal consciousness. Moreover, it was argued that even functionalism does not account for mental causation.

In chapter three, I examined the notion of mental causation in the context of Donald Davidson and Jaegwon Kim's discussions. Davidson is a role functionalist while Kim may be categorized as a realizer functionalist. Davidson and Kim as functionalists in their respective ways fail to account for phenomenal consciousness in human action, thereby not leaving the agent have any place in their actions. It seems neither view can adequately deal with human agency. I hold that any true and adequate solution to the mind-body problem should be able to account for the reality of phenomenal consciousness as causal *qua* mental, and that it is phenomenal consciousness in this sense that carries the agent along.

I will argue in this chapter that Aristotle's theory of hylomorphism may have the answer to doing just that. Its capability to account for phenomenal consciousness in terms of mental realism implies that it accords the agent a place in their action. The agent is not foreign to their action; they own their action because they carry it out. To achieve my aim, I will also employ the ideas of medieval philosophers such as Thomas Aquinas and other neo Aristotelians.

I have argued in chapter three that Donald Davidson does not rescue mental causation and non-reductivism because his theory of anomalous monism through supervenience is

physically oriented, hence there is no independent mental causation, thus siding with Kim on Davidson. In his turn, Kim argues that if it is in virtue of physical properties that the mental properties are causally efficacious, then the mental properties are rendered epiphenomenal. The mental has no independent causal power. I will argue that it is Aristotle's hylomorphism and theory of causes that can save mental causation and any non-reductivist theory of mind from Kim's criticism.

It will be proper to point out that Aristotle is a non-reductivist and that is why I am appealing to his metaphysics to 'save' the non-reductivist theory of mind from Kim. I will explain my reasons for considering Aristotle a non-reductivist, after I have unpacked his theory of hylomorphism and mind. Consequently, I will consider whether he is a functionalist or not.

My focus is not on offering a full analysis or close text reading or critique of the *Metaphysics*, *De Anima*, or any specific Aristotelian text I refer to. Rather, my focus is on showing that Aristotle's hylomorphic metaphysics (bolstered up by the work of medieval writers such as Aquinas) offers a solution to the mind-body problem in general, and for the problem of mental causation specifically, that modern and contemporary accounts of these problems do not offer.

My project is this chapter is set out as follows. I will first discuss hylomorphism and the mind, then I will discuss the mind as immaterial phenomenon by giving Aristotle and Aquinas's view on the immaterial nature of the mind. Then, I will argue that this immaterial mind is the phenomenal consciousness that is causally efficacious in the physical world, and consequently I will discuss Aristotle's response to mental causation by giving his theory of causes and mental causation. Thereafter, I will set out to discuss the causal efficacy of phenomenal consciousness in the context of self-consciousness (self-knowledge) and its causal efficacy in causing humans into action in relation to the theory of functionalism. After, I will show how phenomenal consciousness can cause one to be virtuous. The last task will be discussing the criticism of hylomorphism.

Let us now first consider Aristotle's theory of hylomorphism and the mind.

2. The Hylomorphic Analysis of the Mind

In his *Metaphysics* (M), Aristotle, discussing the concept of substance, states that substance undergirds all the other categories ('attributes') and other categories would not exist without it. For Aristotle, substance is primary in its definition; the definition of the other beings is related to the definition of substance, but the definition of substance does not include the definition of any other category (M 1029a 28).

Aristotle holds that the notion of substance implies "separability and individuality. Substance is capable of being separated; it is independent and needs nothing else for existence. That substance implies individuality means, it is a 'this', i.e. an individual, particular, and numerically one thing" (Vella, 2008: 54-55). There are four notions that explain what a (primary) substance is, namely essence, the universal, genus and substratum (M 1028b34-35). Substance is the sensible or particular things we see around us. Substance as substratum means it is the underlying thing; is that of which other things are predicated but is not predicated of anything else (M 1028b 36). Substratum or substance is characterized in three ways as matter, form and the compound of matter and form.

> And so we must first determine the nature of this; for that which underlies a thing primarily is thought to be in the truest sense its substance. And in one sense matter is said to be of the nature of substratum, in another, shape and in a third, the compound of these. (By the matter, I mean for instance, the bronze, by the shape the pattern of its form, and by compound of these, the concrete whole). Therefore, if form is prior to the matter, more real it will be prior also to the compound of both, for the same reason (M 1029 a1-5).

An individual thing is a composite or unity of matter and form. The metaphysical doctrine that holds that a substance consists of two principles – matter and form – is called hylomorphism. It is derived from two Greek words '*hyle*' (wood, matter) and '*morphe*' (form) (Ainsworth, 2016:1). For Aristotle, forms are not disembodied immaterial things as Plato would

argue; they are embodied in individual things (living organisms or artefacts). Forms are the shapes and plan that matter takes in order to be anything at all. On Aristotle's account, forms cannot exist without the matter and matter cannot exist without a form. The compound of matter and form is just the individual thing, i.e. the bronze statue (Vella, ibid.: 56).

I will acknowledge here that Aristotle's explanation of the term, substance, is sometimes hard to comprehend. And many writers do state the difficulties surrounding comprehending Aristotle's concept of substance:

> He devotes the central books of Met to primary entities or substances … It is not merely that the detail of his arguments is often uncertain; the general drift of his thought, the general thesis or theses towards which he was tending, the overall metaphysical position which he was inclined to accept – these things themselves are subject to scholarly disputes (Barnes, 1995 : 90).

However, in analyzing substance, Aristotle argues that matter, and the compound of matter and form cannot be the substratum of a thing, the thing that makes the thing what it is. The compound of matter and form cannot be the substratum or explain the substance of a thing because the substance compounded of both matter and form is the result of the combination of matter and form themselves. It is not a substance in the primary sense; its explanation comes after matter and form have combined – "the substance compounded of both, i.e. of matter and shape may be dismissed; for it is posterior and its nature is obvious" (M 1029a30).

Matter cannot be a substratum of a thing because matter exists in reference to something. It is a relative term. It is impossible for matter to be a substance because matter is never separable from a thing, however both separability and individuality belong chiefly to substance (M 1029a27-28). Thus, matter is required certainly for anything to exist but it is not the matter that makes something what it is. The same stone or wood for example could be shaped to make a house or a pile of rubbish.

For Aristotle, the substance of a thing explains what makes that thing the definite thing that it is. He holds that it is the form or essence of a thing that explains its being in terms of substance as substratum. ”The essence of each thing is what it is said to be *propter se*. For being you is not being musical, since you are not by your very nature musical. What, then, you are by your very nature is your essence” (M 1029b 13). And it is form or essence that explains the role of substance in this sense.

The form or essence is not merely the sensible shape of a thing; it must also include the function and purpose of that thing. And essence entails an explanation of what that thing is. For instance, the essence of a house includes the bricks, cement and stones and the purpose of putting these material things together to have the form of a house. In the case of a house, the function is that it serves as a shelter for human beings (M 1041b 5-10). The sensible shape is thus part of the form or essence, but it is not entirely a form or essence.

Aristotle holds that the form or essence is not an element, unlike how the pre-Socratic philosophers would argue. Essence is not a fraction of a thing, but rather essence is a principle of a thing. According to Aristotle:

> The syllable is not its elements, *ba* is not the same as *b* and *a*, nor is flesh fire and earth. (For when these are separated, the wholes, i.e. the flesh and the syllable, no longer exist, but the elements of the syllable exist, and so do fire and earth). The syllable, then, is something, not only its elements (the vowel and the consonant) but also something else, and the flesh is not only fire and earth or the hot and the cold, but also something else … But it would seem that this ‘other’ is something, and not an element, and that it is the cause which makes this thing flesh and that a syllable. And similarly in all other cases. And this is the substance of each thing (for this is the primary cause of its being); and since, while some things are not substances, as many as are substances are formed in accordance with a nature of

> their own and by process of nature, their substance would seem to be this kind of 'nature' which is not an element but a principle. An element, on the other hand, is that into which a thing is divided and which is present in it as matter; e.g. *a* and *b* are the elements of syllable (M 1041b 13-19, 25 – 29).

Thus, for Aristotle, constituent elements of a thing are not its essence. A thing can be broken into its constituent elements, but essence is not one of these elements.

Aristotle thus argues that we cannot have an adequate explanation of a thing by analyzing its matter alone, an adequate explanation is obtained when we understand a thing is matter structured and organized in a certain way. It is form (essence), the principle, that binds all material elements, or parts of a thing, together to form an individual object (living or artefact). It is the substance of that thing.

Thus, for Aristotle, form or essence is prior to the matter and 'more' real (M. 1029a6). Form is actual or actual being whereas matter is potential. Matter and form or essence combine to produce a unified determinate substance, like a man or a horse or a statue. Further, these substances are ontologically basic beings, all other accidental beings inhere in them. Thus, the intrinsic quality of ordinary everyday physical substance is analyzed in terms of matter and form and that is the hylomorphic theory of individual substance.

The important feature of individual substances is that they change. Substances undergo a variety of changes or alterations. To understand this, let us first consider Aristotle's notions of potentiality and actuality. The concepts of potentiality and actuality help in understanding the nature of reality and the matter-form relationship, as well as Aristotle's concept of change and the unity of matter and form. Matter is the potentiality while the form is the actuality of a substance. Potentiality is a capacity to become something else while actuality is the capacity of acting. Every substance is a combination of potentiality and actuality. Every actual being is a being in potentiality (M 1045b 20-30). Actuality always precedes potentiality. For instance, a seed is an actual being as a seed but a plant in potency. In the case of producing a statue,

bronze is actually a bronze and potentially a statue, but it will be an actual statue when and only when it is informed with the form of a statue.

Of course, before being made into a statue, the bronze was in potency to become many other things such as a cannon, a steam engine, or a goal on a football pitch, but it was not in potency to become a paper or football. Thus to say "that x is potentially F is to say that x has already actual features in virtue of which it might be made to be F by imposition of an F form upon it. Genetically form and matter can be defined as follows: form is that which makes some matter that is potentially f actually f, matter is that which persists and which is, for some range of f's potentially f" (Shields, 2016:11). Potentiality and actuality underlie the process of change or remaining unchanged. Motion or change is the actualization of potentialities; it is the fulfilment of what is potential as actual. For instance, bronze has potentiality to be made into a statue, it will only become a statue when the potentiality is actualized by being acted upon by an agent. Once the potentiality is acted upon or actualized, it will change into a statue.

Potentiality and actuality are also seen as capacity and the exercise of the capacity. A carpenter is potentially a carpenter (he has the capacity to make furniture) but he is an actual carpenter when he is making furniture (he is exercising that capacity). Wood has the potential to become ash after a fire. Thus, the actual being brings about the fulfillment of the potentiality in another being. The fire, an actual being, can cause the fire to realize its potentiality of being ash. For Aristotle, change simply means the fulfillment or actualization of a particular thing's potentiality so that it becomes what it was potentially before.

In his *Physics* (Phy), Aristotle says that a thing that has a principle of change within it is a substance. To have this principle of change within makes a thing a natural substance or implies that it has a nature. "Things 'have a nature' which have a principle of this kind. Each of them is a substance; for it is a subject, and nature always implies a subject in which it inheres" (Phy 192 b). Thus, nature has two senses, namely form and matter.

Aristotle uses matter and form, in the *Physics*, to account for change in the natural world. Changes are of two kinds namely accidental change and substantial change. Accidental changes occur when concrete particulars or substances gain or lose a property (see Phy 1:7).

Accidental changes can be qualitative or quantitative. Qualitative change occurs when James uses bleaching cream and his complexion becomes brighter and quantitative change occurs when he gains weight by eating too many carbohydrates. However, substantial change occurs when a substance is corrupted or generated (when a substance comes into or passes out of existence) through the matter of the substance acquiring a new substantive form. For instance, a lump of clay that becomes a statue, or a piece of wood becoming a chair.

In general, Aristotle argues that in every change, there is something lost or something gained within the substance. As stated, substantial change occurs when a substance comes into or goes out of existence. In cases of substantive change, there is also something that persists. "To take an example favorable to Aristotle, in the case of the generation of a statue, the bronze persists, but it comes to acquire a new form, a substantial rather than accidental form. In all cases, whether substantial or accidental change … there is something that remains the same and something that is gained or lost" (Shields, 2016: 21).

For Aristotle, change implies three things: (1) something which persists through the change (2) a 'lack' which is one of a pair of opposites (3) a form acquired through the course of the change (Phy 190a13-191a22). In the accidental change, there is no change of form. Accidental change is when the house is painted white. However, in substantial change, in generation or destruction of a substance such as when a human being dies, there is some matter that remains through the change, if not so, there will be creation or generation of a thing out of nothingness or destruction of something into nothing **(**Phy 191a23-b17). This will be against his hylomorphic (form and matter) theory that holds that there is always a matter waiting to receive a form. I will analyze Aristotle's concept of change in more depth when I discuss his theory of causations. Let us now turn to Aristotle's theory of mind and the immaterial nature of mind.

3. Aristotle's Theory of Mind and the Immaterial Nature of the Mind

I am going to consider Aristotle's *De Anima (DA)* as translated by Hugh Lawson-Tancred (1986) to set out and discuss the immateriality of the mind. *De Anima* is Aristotle's work that is most

focussed on the mind-body problem. However, he does refer to his *Metaphysics*, *Physics* and *Posterior Analytics* in his mind-body argument. Lawson-Tancred (1986), in the introduction of the translation, points out the primary aim of Aristotle in *De Anima*.

He notes that Aristotle agrees with other philosophers that there is a principle in the human body called the 'soul'. However, the notion of 'soul' has a specific meaning for Aristotle (ibid.: 11). The Greek word '*psyche'* that means the 'soul' in English has had different meanings for different philosophers, right from Ancient to the Modern time. What the term meant for Aristotle is different from its English translation. For Aristotle, '*psyche'*, is that by virtue of which something is alive and thus '*psyche'* would be best translated in English as the principle of life or the principle of animation. Thus, '*psyche'* for Aristotle is wider in meaning than what the English word 'soul' or 'mind' or 'consciousness' entails (ibid.: 11-12). Richard Sorabji (1974: 65) notes there might be an attempt to translate Aristotle's concept of 'soul' as 'mind' in the modern period:

> The word 'soul' may sound archaic to some modern ears, and people may be tempted to substitute the word 'mind'. But then they are likely to confine the functions of the soul to what we call mental acts, and this will take them away from Aristotle's conception of soul.

Lawson-Tancred (1986) notes that for Descartes, the founder of modern philosophy, the features of mental life (mind) are more important than the features of biological life, but that for Aristotle, the features of biological life are more important and bigger than the mind (mental features or properties). '*Psyche'* deals with all aspects of a human being; it pertains to life in general, it is not limited to the mind or consciousness (ibid.: 13). For Aristotle, the soul, then, is that by which we perceive, feel, think and act (DA 411a26-30), since these are all activities peculiar to living things. To this extent, Aristotle's use of "soul" is quite similar to our use of "mind". We commonly say we perform these activities "with our minds" without implying any Cartesian view about dualism or privacy (Caston, 2006: 317). Thus, it can be

argued that Aristotle was somehow preoccupied with those issues that preoccupied Descartes, as a subset of his notion of the concept '*psyche*'.

In *De Anima*, Aristotle applies his theory of matter and form (hylomorphism), which helped him to explain that change is real contrary to Plato's position that change is not real, to explain the nature of life in living things. Living things in nature can be analysed from their properties of matter and form but in the case of living things, their form is the soul. The form of an individual substance entails a close link between the arrangement and functioning of that individual substance. The form of a living thing determines the way it functions (Lawson-Tancred, 1986: 15).

The theory of the soul as the form of living things is central to Aristotle's psychology in *De Anima*. The soul or the form enables the living being to perform its functions. The capacity for performing a particular function can be dormant or active at any particular time. In discussing the concept of form or the soul, Aristotle stated that every composite has a 'first actuality' and a 'second actuality'. When the capacity for performing a particular action is dormant, it is known as the first actuality of the body and when it is put into use, it is called the second actuality of the body. The soul could then be viewed as the first actuality of the living body. The idea of the soul as the first actuality of the body brings out the similarity and differences among living things and offers Aristotle the chance to discuss sensation, perception, imagination and animal functioning (ibid.: 15).

Thus, the first actuality is the physical and structural arrangement of the composite so that the composite is able to perform or carry out the second actualization, which is the characteristic functioning of the composite.

> … a composite particular may have its first and second actualization. It is in virtue of the first of these, that its Matter is so arranged as to render it capable of performing its characteristic function and it is in virtue of the second that it 'actually' performs them. For example, the Matter of a radio is converted by the man who makes it into the exercise of the first

> actuality of a radio and by the man who switches it on into the second actuality (Lawson- Tancred, 1986: 70).

Lawson-Tancred (ibid.: 24) notes that during the pre-Socratic period, the soul was taken as a phenomenon that was invented to describe a certain aspect of the world. More so, it was seen as a material substance. Some pre-Socratic philosophers were substantialist or materialist psychologists; some not. There were many kinds of psychological theories in ancient times. Some of the theories were of the view that the soul is a physical substance, while there were others that argued that it is a spiritual substance. These two divisions later gave rise to many other divisions among which were those that argued that both the body and the soul are spiritual or the body can be spiritual and the soul physical, or the body can be physical and the soul spiritual or both body and soul can be physical. The first view represents the position of idealists, which also was the position of Plato. No known philosopher ever really taught the second position, while the Cartesian dualists held the third stance that the soul is spiritual and the body physical. The fourth view conceives the soul as a physical substance (ibid.: 24).

As mentioned, many of the pre-Socratic philosophers were of the view that the soul is a physical substance, but Aristotle did not support this. For Aristotle, every good account of the soul must account for three aspects namely (1) How living bodies generate physical processes or spontaneous movement; (2) How living bodies could have cognition and mental awareness of their surroundings, and (3) How physical and mental processes could form a unit in certain living bodies. Aristotle believes that his predecessors were unable to account for the latter, the unity of physical processes and mental processes in some animals. He explains this at length in Book 1 (ibid.: 27). Aristotle's position shows that an aspect of the soul could be taken to be the modern notion of the mind, since it is by the soul that a human being performs mental and conscious activities. Victor Caston (2006: 317) writes:

> The soul, then, is that by which we perceive, feel, think and act, since these are all activities to our use of "mind". To this extent, Aristotle's use of "soul" is quite similar to our use of mind. We commonly say that we perform these activities "with our

> minds", without implying anything Cartesian about dualism or privacy.

Among the pre-Socratic philosophers, as mentioned already, we have materialist psychologists who taught that the soul is of a single material stuff. Moreover, this material stuff was viewed to be one of the four elements namely water, fire, earth and air. On these views, it is the particular element in terms of its special features or characteristics that is responsible for the psychic features in the living body. Another group of the material psychologist are the 'many material stuff' believers. Empedocles was a prominent member of this group. Empedocles held that the soul, like any other thing in nature, is composed of the four elements. The third group of materialist psychologists was called atomists and they held that the soul is not made of the four elements, but rather it is made of small particles called atoms. The proponent of this theory is Democritus. He held that the soul is made of the smallest and most mobile of the atoms that has many properties similar to those of fire. Aristotle directs his criticism of materialist psychology against the theory of Democritus. Aristotle also attacks Empedocles' view that the soul is made of many stuffs (ibid.: 28-29).

Aristotle rejects both Democritus' account of psychic motion and Empedocles' account of cognition thereby rejecting materialist philosophy of the pre-Socratic era. This rejection marks out clearly the difference between Aristotelian psychology and the materialist psychologists. For Aristotle, both Democritus and Empedocles gave a wrong account of psychic functions because they explained the soul only in terms of its material components (ibid.: 30-32). Their accounts are reductive in nature because they held that every natural phenomenon can be accounted for based on its material particulars. Contemporary reductivists like Jaegwon Kim have, in that sense, continued the tradition of Democritus and Empedocles albeit in another format.

Aristotle refutes Democritus' account of cognition by arguing that the soul of an animal cannot bring about the cognition of the objects of its environment on account of having the same material constitution as these objects. Aristotle believed that the soul must have something superior to the objects of its environment. He also argues that if perception is 'like

by like[10]' how would the soul perceive and understand or acquire numerical and abstract knowledge that is not physical in nature (ibid.: 33-34).

Aristotle argues that biological activity cannot be explained by the accidental presence of a foreign matter within the physical constitutions of the animal body. Any animal or human activity must be explained from both perspectives of body and soul, which combine to offer one explanation of a single animal activity. "Any activity of a living body will be explicable in one way in terms of the matter that composes the body itself and in another way in terms of the soul, but it will be the same activity that is explained in both occasions" (ibid.: 37). For the physicalists the soul is nothing but the effects of the activities of the organs in the body. Thus, physicalism professes attributes of the soul. The activities of the body warrant us to say that there is soul, but no role is given to the soul in explaining the activities of the living body.

The presentation of the soul as a general form of a living body is a very important theory in Aristotle's psychology and central to this theory is the application of his matter and form dichotomy (hylomorphism). It is interesting that Aristotle's theory of mind that is anchored on the unity of form and matter could arguably be likened to David Chalmers's theory of consciousness for Chalmers's type f-monism (Chalmers, 2002)holds that physical properties and phenomenal properties are united. Type F monism is the view that consciousness and physical properties are intertwined because there is no separation between phenomenal properties and properties; the two are found together in every substance just like the matter and form of Aristotle constitute every individual substance.

Aristotle believed that any object including every living thing is made of matter and form. The form guarantees the consistent functioning of any living thing. Aristotle held that the soul is the first entelechy of a natural body with organs. It means that every living thing is alive in virtue of the soul, its form. Without the soul, the living thing will not be able to live or carry

[10] Like by like (like affects like) implies that in the act of perception, the organs of perception only perceive objects of perception that are like them in terms of being made up of the same elemental material. However, Aristotle argues against the 'like by like'. For Aristotle, it is correct to say that the organ of perception perceive something unlike it, and in perceiving, the organ of perception becomes like perceptible object. For him, in perception unlike becomes like; unlike perceptible organ in perceiving becomes like the object of perception (See Vella, 2008: 101-012).

out those functions that ensure the realisation of its intrinsic end or purpose. The inherent capacity that is the soul enables the living thing to reach the goal of his existence (Lawson-Tancred, 1986: 59).

Now, in Book 11 of *De Anima* (DA), Aristotle shows that forms are substances. If forms are substances then the soul is a substance as form. In classifying the substantiality of the soul, Aristotle appeals to his metaphysical theory that every substance is substance in one of three aspects namely substance as matter, substance as form, substance as the composite of matter and form. Aristotle would dismiss soul as matter or composite of matter and form. Matter (body) is capable of taking any form, it is associated with potentiality. For him, soul as substance is form. Thus, the soul is the formal cause (that which makes a thing to be what it is), the efficient cause (that which makes a thing to actually exist) and the final cause of any living being (the end for which a thing is made) – I will discuss Aristotle's full theory of causes later in this chapter. A particular thing is alive and functioning because of the soul. The soul is more than the first attributes of the body. It must also from another related perspective, be the reason why a particular thing is in existence. It should be the cause of changes that the particular thing undergoes or experiences in relation to its environment. It must aid the object in undergoing these changes and it represents the purpose and goal of that particular thing. The soul is the primary actuality of the body, providing the body with its essential character and therefore is inseparable from it. That is why the soul is an actuality of the first kind of a natural body having the potential for life in it, and is an actuality of the second kind of a natural organized body' (DA 412 a27, 412b5). The soul is the form of the natural body, and it is form in the sense of first actuality.

Now, it may sound as if we are moving towards substance dualism, if Aristotle claims the soul to be a substance in terms of its being form (and for the sake of argument also that matter is substance), but we are concerned here with how matter and form exist *within* a living organism. Within the living organism, body and soul (including the subset of properties such as perception, which we would today describe as 'mind', as mentioned above) are inseparable, and no individual living organism can exist without being a hylomorphic combination of the

substances of matter and form. I thus argue that, while Aristotle acknowledges that there are three ways in which substance 'can be said', as mentioned above, namely as form (soul), matter (body), and individual substance (combination of matter and form), it is the latter that is at issue when one considers his philosophy of mind, as human beings are substances in the latter case. Therefore, Aristotle is decidedly not a substance dualist, but rather a non-reductive physicalist, perhaps best viewed in terms of property dualism. I come back to this below.

Aristotle uses examples of natural substances such as an axe (artefact) and an eye (part of animal) to illustrate his definition of soul.

> In the same way, if some tool, say an axe, were a natural body, its substance would be being an axe, and this then would be its soul. And if this were separated from it, it would not continue to be an axe, except homonymously, whereas as it is it is an axe. But it is not of that sort of body that soul is the being what it was and the account, but a natural body of the right kind, having itself the principle of movement and rest (DA 412b12-14).

Thus, that the axe is a natural body is because of its 'soul'; being an axe (*axenes*), which is the capacity to cut, once the axe loses this capacity (its 'soul' or form); it is unable to cut, it is no longer an axe except in name or homonymously. It will be like a toy axe that cannot cut because it lacks the essence of its being, its form or 'soul'.

For a human example such as an eye, Aristotle states that "for if the eye were an animal, then sight would be its soul, being the substance of the eye that is in accordance with the account of it. And the eye is the matter of sight, then when sight leaves it, it is no longer an eye except homonymously, in the way of a stone or painted eye" (DA 412b17-21). Thus, a painted or artificial eye lacks the soul or capacity of seeing; it is an eye only in name.

The soul is not a different entity from the body. The soul and body together form an individual substance, just like matter and form is explained to do in the hylomorphic theory. The soul is the capacity or collection of capacities of the body, hence it is not separated from

the body. For Aristotle, the soul is the actuality or form of the bodily organs or parts, it is not possible for the soul to be separated from the body. However, the capacities of the soul are separable when they are not actualities of some organ:

> It is quite clear then that the soul is not separable from the body, or that some parts of it are not, if it is its nature to have parts. For with some of the parts of the soul the actuality is of bodily parts themselves. Not that there are some parts that nothing prevents from being separable, through their not being the actualities of any body (DA 413a3-6).

The capacities such as sight and hearing are actualities of bodily organs, namely the eye and ear respectively; they cannot occur in the absence of an eye and an ear. Thus actualisations are actualities of certain organs. Aristotle regards thinking or intelligence as the only capacity of the soul that is not an actuality of any bodily organ. Here we may encounter a new sense of doubt as to whether this does not after all imply some form of substance dualism, but I will argue in the remainder of the chapter that all that this really imply is at most property dualism:

> But nothing is yet clear on the subject of the intellect and the contemplative faculty. However, it seems to be another kind of soul, and this alone admits of being separate, as that which is eternal from that which is perishable, while it is clear from these remarks that the other parts of the soul are not separable, as some assert them to be, though it is obvious that they are conceptually distinct (DA 413b 17-19).

Aristotle's theory of the soul can also help us to understand how life in plants and animals manifests itself and the interdependence and interconnection between plants and animals. The soul, as the first entelechy, is made of a series of psychic faculties, which appear in hierarchic order in the animate part of nature. Plants, animals and man are on different positions on the hierarchal order of living things but each superior one shares a particular

feature/s in common with the one below it. It is the task of the theory of the soul, or psychic faculties, to explain the commonalities and differences according to these grades.

In *De Anima*, Aristotle shows how nutrition and reproduction make up the first faculty (vegetative soul) that occurs in plants, animals and man, and how sensation, perception and motivation (desire), make up the second faculty (sensitive soul) that occurs in animals and men but not in plants, and how thought, the third faculty (rational soul) occurs only in men, which is what enables men to be the only species that is capable of possessing knowledge (DA 414a20-25). "…soul is that by which primarily we live and perceive and think…" (DA 414a 12). For Aristotle, for living things, living is existing.

For Aristotle, the capacity of thought is quite different from other capacities of the soul as it is not an actuality of any bodily organ like other capacities, as mentioned in the quote above. Hence, soul in terms of thought is separable from the body, unlike other capacities such as nutrition, movement and perception that are capacities of some certain bodily organs. Thought seems to be an independent substance implanted within us and to be incapable of being destroyed (DA 408b19-30). There is a lot of debate about how exactly to interpret the latter point. For my purposes, I will take it *not* to be against non-reductivism, should the intellect remain after death, and thus not in favour of full-out substance dualism, as it is very clear that Aristotle, in terms of his hylomorphic theory and what it means to be an individual human being, is not a substance dualist.

The intellect is that by which the soul thinks and supposes. Since thought thinks all things it is appropriate then that it is not mixed with the body. The eye sees only visible things while the intellect thinks all things (Vella, 2008: 106-7). For Aristotle, "before the intellect thinks, it is not actually any existing thing. The intellect exists as pure potential; it is potentially any intelligible form" (ibid.: 107).But, intellect in its activity (thinking) may parallel a perceptual activity. "Just as perception involves the reception of a sensible form by a suitably qualified sensory faculty, so thinking involves the reception of an intelligible form by as suitably qualified intellectual faculty" (Shields, 2016: 13). Thus in a schematic outline, this is what happens when one thinks about an object *O:*

> S thinks O if and only if: (i) S has the capacity requisite for receiving O's intelligible form; (ii) O acts upon that capacity by enforming it; and; as a result, (iii) S's relevant capacity becomes isomorphic with that form. (Shields, ibid.: 13)

However, intellectual activity or thinking occurs at an abstract or higher level than perception, and it encompasses all things as the intellect thinks all things, without perceiving them. The intellect, unlike perception, is not affected or becomes like what it thinks; it abstracts only the forms (ibid.: 14).

How can the intellect take on an intelligible form while itself remaining unaffected? In *De Anima* 3.5, Aristotle sets out to address the difficulty by introducing a logical distinction between the active intellect and the passive intellect. He uses his theory of four causes to explain the difference between the passive intellect and active intellect. (The theory of the four causes is fully explained in a subsequent section. Aristotle's metaphysics is so complex, it is difficult to explain all concepts at once.)

The passive intellect is a sheer potentiality to receive an intelligible form and as a potentiality, it needs an actuality to cause it to actualization or bring it into full existence as no potentiality brings itself to actualization, the bronze in a potential state cannot cause itself to be a statue, it needs an actual agent, the efficient cause to do that. Thus, the passive intellect as a potentiality needs an efficient cause to cause it to actualization, and this is the active intellect. Thus, the mind in its potential state is the passive intellect and in its capacity as an actuality (form as efficient cause, formal and final cause) is the active intellect.

> Now in all nature there is for each genus something that is its matter (and it is this that is all those things in potentiality), and something else that is their cause, productive of them in virtue of bringing them all about – as, for instance, a skill stands towards the matter it uses. No less in the soul, then, must these different features occur. And indeed there is an intellect characterized by the capacity to become all things, and an

> intellect characterized by that to bring all things about, and to bring them about in just the way that a state, like light does. (For in a way, light also makes things that are potentially colours in actuality.) Now this latter intellect is separate, unaffected and unmixed, being in substance activity. For in all cases that which acts is superior to that which is affected, and the principle to its matter (DA 430a10-19).

For Aristotle, the active intellect thus produces all intelligible forms while the passive intellect can potentially become any intelligible form produced by the active intellect. He further characterizes the active intellect as unmixed and unaffected, thereby refining his earlier claim that the intellect must be unaffected and unmixed (DA429a18-29). The passive intellect is affected and mixed. It becomes the intelligible forms and so is affected; it is also the case that the passive intellect perishes with the body, and so it is mixed with the body. The passive intellect is a corresponding matter to the form, which is the active intellect. The passive intellect is the matter of the faculty of intellect; it is the organ that is enformed by the intelligible form. On the other hand, the active intellect is what is immaterial, immortal and eternal; its essence is activity, and this cannot cease. It is in this sense, also reminding of what has been argued before in terms of the modern notion of mind being a subset of the soul, then that I argue that the mind is immaterial for Aristotle and it is real as it causes things to happen *qua* mental. This much we can take from Aristotle, while what exactly he meant with the active intellect persisting after death is up for debate (see e.g. Sorabji (1990) and Gerson (2005)). The notion of the immortality of the active intellect however does not necessarily negate a non-reductive physicalist interpretation of Aristotle's philosophy of mind, and thus for my purposes this is sufficient, as it is clear that the soul as the active intellect causes action, is immaterial, and cannot be mixed with the body. In terms of the hylomorphic theory this makes for a philosophy of mind that is non-reductivist in nature. It is only if we do not bring the hylomorphic theory into consideration that we wonder whether the fact that Aristotle also seems to be arguing for the immortality of the active intellect (e.g. Gerson 2005) doesn't make for substance dualism. Here I would argue that we should not think of the active intellect as a

substance but as a principle, but I will not further discuss this concern here. I will rather turn towards medieval philosophy to support my arguments here and provide more clarity.

Medieval philosophers, especially Thomas Aquinas, have followed Aristotle to hold the idea that the mind is immaterial.

4. Medieval Philosophy and the Immaterial Nature of the Mind.

Medieval philosophers, especially Thomas Aquinas, agree with Aristotle that material objects are composed of form (*morphe*) and matter (*hyle*). Form and matter are real (metaphysical entities) within material objects, but not separate substances. The immaterial form shapes indeterminate matter to produce and intelligible material reality (Eardley, et.al. 2010: 29-30).

But for the medieval philosophers, the immaterial form of human beings have a double function: "For later medieval philosophers, writing under the influence of Aristotle's natural philosophy and metaphysics, the human soul plays two quite different roles, serving as both a substantial form and a mind" (Pasnau, 2012: 486). Thus, for them, the human soul has both biological and other psychological functions. The human soul is fundamentally a mind and a form of body.

The soul gives rise to the human mind without actualizing matter at all – the human soul by itself, quite independently of the body, is responsible for the thinking and willing that are the characteristic operations of mind. Thomas Aquinas tries to explain this:

> It is necessary to say that the principle of intellectual operation, which we call the soul of a human being, is a non-bodily and subsistent principle. For (1) it is clear that through the intellect a human being can cognize the natures of all bodies. But (2a) that which can cognize certain things must have none of those things in its own nature, because that which exists in it naturally would impede its cognition of other things. In this way we see that a sick person's tongue, infected with a jaundiced and bitter humour, cannot perceive sweet; rather, all things seem bitter to

> that person. Therefore if the intellectual principle were to contain within itself the nature of any body, it could cognize all bodies. But everybody has some determinate nature. Therefore [3a] it is impossible for the intellectual principle to be a body. [3b] It is likewise impossible for it to operate through a bodily organ, because [2b] the determinate nature even of that bodily organ would prevent the cognition of all bodies. Analogously, a determinate colour not just in the pupil, but even in a glass vase, makes liquid poured into that vase seem to be of the same colour. Therefore [4] this intellectual principle, which is called mind or intellect, has an operation of its own that the body does not share in (Aquinas as quoted in Pasnau, 2012: 494).

Christian Brugger (2008: 113) has reformulated Aquinas's argument as follows:

> 1. Cognition includes receiving information pertaining to particular entities not only as particulars, but also as kinds (that is, it includes knowing the natures of things). 2. The act by which we know a thing's nature is the act of forming universalized concepts. 3. Now, a particular entity can only be perceived as a particular because it is materially instantiated (that is, it is perceived as a particular by way of its materiality). 4. A universalized concept is universal precisely because it is abstracted from particularity and hence materiality. 5. That which has cognition of a materially instantiated particular can have it only by receiving information into itself pertaining to the principle of the entity's particularity, namely, its materiality (sensation). 6. The power by which a cognitive power receives particular information pertaining to a thing's materiality must be a materially instantiated act. 7. The power by which a thing

> carries out materially instantiated acts must itself be materially instantiated. 8. The act by which a cognitive power receives information free from everything material must itself be free from everything material (intellection). 9. In the formation of universalized concepts the intellect receives information free from everything material. 10. The act of concept formation therefore is free from everything material. 11. Something able to act free from everything material must itself be free from everything material. 12. The intellect, as capable of forming universalized concepts, is free from everything material.

Thus the intellect can think about anything, which would be impossible if the intellect were to be a body or operate through a bodily organ. "If the human soul is responsible for the operations of the mind, and if the mind operates independently of the body, then the human soul is certainly a very special kind of form" (Pasnau, 2012: 494). For, Aquinas, the soul, in relation to the mind, is responsible for a feature of human being that is not a body and does not use body in its operation. Thus, for Aristotle and other medieval Aristotelians, the human soul is immaterial. It is not immaterial because it is a form nor is it immaterial because it does not form a body, and it is not located spatially. It does inform a body as every other form and is inhered where the body is. The human soul is immaterial because it operates independently of the body, as it thinks all by itself (ibid.: 495).

Following Aristotle, Aquinas divides the immaterial soul (human intellect) into two principles namely passive and active intellects:

> … the human intellect – unlike the divine intellect – begins in a state of potentiality towards everything intelligible, though actually knowing nothing. The human intellect is passive because it must receive something in order to be actualized, much as matter needs to receive form in order to be actualized. Our intellects do not begin in an already actualized state, nor can

> they actualize themselves…Since our minds are designed to know material things, which turn out not to be intelligible in their own right, the human intellect must also have an active power, specifically a "power to make things actually intelligible". Following a long Aristotelian tradition, Aquinas will call this active power the agent intellect, while he will call the passive power the potential or possible intellect. (Eardley, et.al. 2010: 42-43)

Some philosophers, like William Ockham (1991), have argued in response to Aquinas that we cannot know through experience or reason that thinking is the operation of an immaterial substance, the soul. "… all we experience are the acts themselves of the mind – acts of thinking, willing and so on. There is nothing in experience that points towards any sort of distinctive, non-physical origin" (Ockham as quoted in Pasnau, 2012: 496). One can argue in contrast that mental experiences such as the phenomenal or perceptual experience of seeing colour or hearing a sound is beyond any physical or biological explanation as they belong to the mind. I will discuss this phenomenal experience further under consciousness later in the chapter.

Picking up the concern around Cartesian dualism discussed in section 1.3.1, it is suggested that the Aristotelian stand that the soul can exist and function apart from the body supports Cartesian dualism as Cartesian dualism holds that the mind is independent from the body; the mind is immaterial and non-extended while body is material and extended. In this sense, Aristotle is accused of not being faithful to his hylomorphic project with regard to thought. In terms of thought or intellect, hylomorphism seems to collapse. Hylomorphism professes unity of form and matter in the substance, in other words, it professes that matter and form are inseparable. However, when it comes to thought or intellect, we have two kinds of intellect namely the passive intellect and the active intellect. The active intellect is separable from the body; it is not attached to anybody. It is unmixed and thinks about everything. Hence,

it is argued that this shows Aristotle's acceptance of substance dualism (Shields, ibid: 15, Vella, 2008: 114).

A neo-Aristotelian may respond to the criticism in this way; the active intellect may be separated from the body, however, it works with the body. It is a soul in an appropriate kind of body; a body that is capable of imagination and perception. For it to think, it needs images, and images are products of imagination and perception, which are actualities of the organs of the body.

> … Although Aristotle may consider the possibility of thought independent of body, his conclusions concerning the relationship between sensation, imagination, and thinking preclude accepting it. For Aristotle, "the soul never thinks without an image" and all imagination is "movement resulting from an actual exercise of a power of sense" (DE 431a17; DE429a2). As a result all thinking "requires a body as a condition of its existence" (DE403a7-10). Mind cannot exist independent of the body (Hsieh, 2002: 4).

It makes sense to speak of an active intellect as it provides the passive intellect with intelligible forms. And these intelligible forms are used in actual thinking by the passive intellect. The soul thinks with the passive intellect, and by receiving forms from the active intellect, it is capable of analyzing and judging, even critiquing its thought by the forms provided by the active intellect. The active intellect becomes a mirror to the passive intellect. Thus, because the active intellect works in collaboration with the passive intellect and organs of perception, Aristotle cannot be accused of giving in to dualism.

I quite agree that Aristotle cannot be accused of dualism because the intellect is a faculty of the soul in a human person, which is a composite of matter and form (body and soul), and the intellect and the other faculties of soul (vegetative and sensitive souls) make up the human person. The intellect cannot work in isolation to the other two, even to the human environment and society. The intellect, with the vegetative and sensitive souls, helps humans to

live a rationalized life in society. Humans use the intellect to navigate their ways and plan for their lives, for their human flourishing and survival, hence it is related to other souls.

Having explained Aristotle's theory of hylomorphism, the soul and mind, let us consider briefly why I take him to be a non-reductivist, and then I will consider his response to mental causation.

5. Aristotle as Non-reductivist

Aristotle is a non-reductive physicalist because his theory of hylomorphism presupposes that there is such a thing as the soul *(psyche*) and a fundamental irreducible distinction between the soul and the body, and more generally between their formal and material contributions in the analysis of psychological phenomena (Caston, 2006: 319). The soul is necessarily the form of the body. Thus, psychological phenomena have both a formal and a material aspect.

Aristotle would reject "type identity" (the theory that holds that mental kinds are reducibly identifiable with physical kinds), but would accept "token identity" (the theory that holds that a particular mental event is reducible to a particular physical event, but not that a set of mental properties are reducible to a set of physical properties) (ibid.) The soul and body need each other to exist but the soul is not reducible to the body. There is no direct identification of the soul with the body. An individual living thing is both animate (soul) and embodied (body), so every living individual instantiates both the soul and the body (ibid.) – in terms of form and matter.

6. Aristotle's Response to Mental Causation

6.1. Aristotle's Theory of Four Causes

Aristotle's theory of form and matter, potentiality and actuality, plays a very significant role in his doctrine of four causes or theory of causality. "The theory of causes (doctrine of causation) explains the operative causes of real being or the cause of all things in nature. It explains the principles of causation that produce change" (Sreekumar, 2008:3).

Aristotle argues that we have knowledge of a thing when we have grasped its cause, which entails investigating and answering the question, 'why' about that thing. And the why-question is a call for an explanation. Aristotle thinks of a cause as a kind of explanation, a scientific explanation for that matter (Falcon, 2019: 2). Aristotle's concept of cause is presented in *Physics,* 3.2. His notion of cause (*aitia*) does not connote the modern notion of cause. For Aristotle, '*aitia'* can mean cause, reason or explanation.

> … Now that we have established these distinctions, we must proceed to consider causes, their character and number. Knowledge is the object of our inquiry, and men do not think they know a thing till they have grasped the 'why' of it (which is to grasp its primary cause). So clearly we too must do this as regards to both coming to be and passing away and every kind of physical change, in order that, knowing their principles, we may try to refer to these principles each of our problems (Phy 194 b16-20).

Thus, Aristotle's four causes are supposed to explain any natural change. While Aristotle is not saying that any natural change has four separate causes, he holds that for any natural change there are four explanations that must be given. For us to have complete knowledge of anything we must seek for and investigate its four causes or explanations (Vella, 2008: 76). Aristotle outlines the four causes in this manner:

> In one way, that out of which a thing comes to be and which persists, is called a cause, e.g. the bronze of the statue, the silver of the bowl, and the genera of which the bronze and silver are species [material cause].
>
> In another way, the form of the archetype, i.e. the definition of the essence, and its genera, are called causes (e.g. of the octave, the relation of 2:1, and generally number, and parts in the definition) [formal cause].

> Again, the primary source of the change or rest; e.g. the man who deliberated is a cause, the father is the cause of the child, and generally what makes of what is made and what changes of what is changed [efficient cause].
>
> Again in the sense of end or that for the sake of which a thing is done, e.g. health is the cause of walking about (why is he walking about? We say, 'to be healthy', and, having said that, we think we have signed the cause) [final cause]. (Phy 194b23-35)

John A. Vella (2008) explains the four causes in this way; the four causes are identified as 1) the material cause, 2) the formal cause, 3) the efficient cause and 4) the final cause. Aristotle converts two of his principles of change namely matter and form (potentiality and actuality) into causes namely the material and formal causes and these account for material and formal explanations of any change or no change. Then there are the efficient and final causes. The efficient cause is the primary source of change, while the final cause accounts for the goal or aim of change.

Let us consider the four causes in relation to a builder building a house. The formal cause is the idea in the mind of the builder that shapes the house. The material cause is the bricks, stones, zinc, etc. of which the house is built, the efficient cause is the builder himself including the tools used by him. The final cause is the finished building, the house; it is for the sake of which it is built.

The material cause and formal cause are inside the individual substance undergoing change or remaining unchanged, the efficient cause is external to the individual substance. The builder who causes the bricks and stones to change to a house is external to the materials. The efficient cause, in the case of artefacts, may be similar to Hume's concept of cause as according to Hume a cause precedes its effect; while the final cause is teleological, it is for the sake of which something happens. It is similar in meaning to the Greek word *telos*. The final cause means the end or aim or purpose of any change or staying unchanged (ibid.: 77).

The causes are not only applicable in the case of artefacts made by humans, but also in natural organisms. Natural substances according to Aristotle have within them the efficient cause, which is the source or principle of change or staying unchanged. An acorn seed grows and matures into an oak tree according to its own internal source; it contains within itself its efficient cause. Aristotle's teleology means that there is a purpose in nature; every natural process aims at some good for the natural substance involved. This good is inherent in a natural substance, which is the source of change for the substance. The final cause provides an explanation of the end or goal of the change (ibid.: 78) – Jane does physical exercise in order to achieve good health. The material cause is the physical body, and the formal cause is the soul. (I will come back to this later on in this section.)

Every instance of causation, either natural or artificial, requires the four causes or explanations. The four causes work together to give a complete explanation of any change or remaining unchanged. None of the four causes alone can give a complete explanation of any change (ibid.: 78). Motion or change can take place in a material substance only when the efficient cause acts on the matter to bring out the potential form in the matter, which also accounts for the aim or goal of the change (Phy 198a 21-23).

And, obviously, the four causes are closely related to Aristotle's theory of matter and form. He sees form in terms of the efficient, formal and final cause: "The last three (*formal, efficient and final*) often coincide (*especially in natural objects*); for the what and that for the sake of which are one, while the primary source of motion is the same in species as these" (Phy 198a25-26). The words in italics are mine.

In his further explanation of Aristotle's theory of four causes, Vella (2008: 79) remarks that Aristotle identifies the formal cause with the final cause in terms of living organisms, because the final cause is in a sense contained in the formal cause. When a human infant acquires or reaches a matured human form she has attained her final cause. As a human infant, she is an adult human person potentially, and to become an adult person is the goal of her growth (change) and she reaches that goal once the potential has been actualized by acquiring the form of an adult human person.

Vella (ibid.: 79) further remarks that Aristotle also holds that the efficient cause is in the formal cause and final cause in the case of living organisms. The efficient cause is a function performed by a something with a matured form. The efficient cause of a human infant is the father; for he gives human form to the infant, and he can only do this because he embodies the matured form. He is the efficient cause because he has the form (formal cause), which is identical with the final cause. The same relationship is noted in the production of artefacts such as a statue from bronze. The artisan (efficient cause) imparts the form of the statue present in her mind to the bronze, which is both the formal and final cause. The efficient and final causes are merged in the formal cause. So, form is efficient, formal and final cause.

Having introduced Aristotle's theory of hylomorphism and four causes, let us now turn to an exposition of his notion of causation in the context of the mind-body problem.

6.2. Aristotle's Account of Causation

For Aristotle, as we noted in section 1, all individual substances consist of matter and form; matter is the potentiality and form is the actuality, and change is actualization of the potentiality. Thus, in any substantial change, matter acquires a new form. The matter is always in potentiality to receive a new form.

The matter and form theory (hylomorphism) plays a very important role in Aristotle's theory of causes. In his theory of causes, he converts matter and form into the material cause and formal cause. The other two causes he identifies are the efficient cause and final cause. He considers the form as the formal, efficient and final causes. Via the hylomorphic theory the matter and form distinction is applied to his theory of mind. Matter and form are understood as the body and soul distinction. The body (matter) is potentiality and soul (form) is the actuality. As soul is the form; it is also the efficient, formal and final causes and they are inherent or internal in the living thing. The physical body is thus simply the material cause of the organism. Body and soul are integrated together to form the body.

It will be good to note here that as the soul is the efficient, formal and final causes in Aristotle's theory of causes of natural substance, and is inherent in the body, its operation must be different to the operation of the efficient cause of non-living things or artefacts, which is

external to them. Hence, Hume's concept of cause (cause and effect are spatiotemporal, and cause always precedes its effect) should not be applied to living things that have the efficient cause within them in an Aristotelian context. Hence, in this context, Davidson (1980), Yablo (2002), and Hutto's (1990) invitation to rethink mental causation with regard to human agency seems a good way to go. I will come back to this later in section 6.5, where I will discuss the mental faculty called 'desire' that together with practical reason and imagination causes humans into action. These mental faculties are inherent in human beings, not external to human beings.

The soul, for Aristotle, is a real cause, and it can explain the behavior of all living organisms. Soul as a real cause is responsible for the directed growth of any living organism. The soul is also the real cause of the unity of different parts of an organism. Aristotle applies his hylomorphic account and his theory of four causes to explain and analyse psychical states such as perception, desire and thought. Moreover, it is in his theory of mental causation that biological functionalism becomes explicit in that the soul as the efficient cause (formal cause and final cause) is the functional organization of the body (material cause) that brings about perception, thought and desire in the right kind of bodily component. The functional role that the soul plays in the bodily organisms has created a debate whether Aristotle is a functionalist or not. I will engage in the debate later in section 8.

6.3. Perception

Aristotle treats perception in causal terms. Perception or sensation is a type of change; it passes from potentiality to actuality. For him, sense perception involves the body being affected or influenced in a particular way; there is always an external object to be perceived (seen, heard, touched, smelt, tasted), which is the object of perception that affects our sense organs (the eye, the ear, the nose, the tongue, the skin).

The senses do not perceive themselves, but rather they perceive some object external to the senses (DA 417a2-5). According to Aristotle, the fact that the senses can only perceive when there is some external object present, shows that the senses exist by way of potentiality rather than by way of actuality (DA 417a5). The senses perceive when there is an external

object; without an external object there is no perception. However, the senses will be in potentiality to perceive in the absence of an external object. Aristotle aligns his discussion of perception or sensation to the discussion of the soul as first actuality and second actuality. Perception, as with all capacities of soul, exists as a first actuality of a body with organs. Perception thus requires the first actuality of the organs of perception as well as an external object that affects the organs in some way. When an object is present to the organ, the organ achieves the second actuality (DA417a2-5). Thus, entelechism is also noted in the act of sense where entelechy refers to the realization of potential in the sense of referring to the supposed vital principle that guides the development and functioning of an organism.

Thus, every perceptual activity in a living being is made possible by the soul, the inner principle of change or rest. The soul is the efficient cause that initiates the motion or being affected or acted upon.

> To be 'acted upon' has more than one meaning; it may mean either the extinction of one of two contraries by the other, or the maintenance of what is potential by the agency of what is actual and already like what is acted upon, as actual to potential (DA 417b2-5).

Thus, for Aristotle, being affected implies two metaphysical aspects namely extinction and maintenance (preservation). There is an occurrence of extinction, in that, the content of the organ of perception or sensation is annihilated before perceptual activity; and it takes on the form of the object of perception. During perception the organ of perception abandons its form and takes on the form of the object of perception that is contrary to it. Hence, the distinction is by contrary. However, the aspect of maintenance or preservation is effected too because the form of the external object perceived is preserved by the sense organ as it is enformed by it (Vella, 2008: 100).

In *De Anima*, Aristotle is not only interested in the act of sense perception but also he intends to resolve the problem raised by the pre-Socratic philosophers which is whether perception is by 'like by like' or 'unlike by unlike'. For Aristotle it is 'like by like' and 'unlike by

unlike' (Law-Tancred, 1986: 77). This is contrary to the pre-Socratic philosophers who maintained that it is 'like by like'. The 'like by like' theory holds that the organs of perception are only affected by objects of perception that are like the organs, that is, objects are composed of the same elemental material as the organs. Thus, the organs of perception are themselves made of matter. The pupil of the eye for example is made of water according to some pre-Socratic thinkers. However, for Aristotle, sense organs are potentially any perceptible form. The sense organs can become anything that we can perceive by taking on the perceptible form of the object of perception. The perceptible form can be the sensible shape, sound, colour or flavour of anything. The 'unlike by unlike' view of Aristotle implies that the perceptive organ can move from a state of actual dissimilarity to actual similarity with the sense object. This is transition from first actuality to second actuality (ibid.: 77).

And, in the way I have set out above, the sense faculty is like the actual sense-object (the sense faculty is potentially like what the perceived object is actually), it is affected as being unlike, but on being affected, it becomes like and is such as what acts on it (DA418a 4-6). "The perceptive faculty is in potentiality such as the object of perception already is in actuality" and when something is affected by an object of perception, "it is made like it and is such as that thing is" (DA 418a3-6).

> Aristotle is happy to speak of an affected thing as receiving the form of the agent which affects it and of the change consisting in the affected thing "becoming like" the agent. So there is in both cases a hylomorphic model of alteration involving *enforming*, that is, a model according to which change is explained by the acquisition of a form by something capable of receiving it (Shields, 2016: 11).

Aristotle holds that we gain information about the world both through the acts of sensation and perception. He explains this by showing the difference between sensation and perception. Sensation is an act of alteration of a body organ in order to bring it in greater similarity with the sensed object, while perception is closer to an act of judgement and helps us

to conclude that the sensed object is a black cup or white cup. Thus, the word *aesthesis* could mean 'act of alteration' or 'act of making a judgement'. Perception is an awareness of sensible form.

This way of presenting a sense-perception act causes people to think that Aristotle is talking about consciousness, but Aristotle nowhere in *De Anima* discusses consciousness *per se*. Hence, some writers argue, the modern idea of consciousness cannot be imposed on Aristotle (Lawson-Tancred, ibid.: 78). I would argue however, that Aristotle's idea of perception could be described in modern terms as a simple form of consciousness since awareness of sensible form is entailed. After all, to be aware of something (perceive something) is to be conscious of it, at least in a minimal way without necessarily being conscious of all the facts about that thing.

Perception or perceptual activity is then a case of mental causation because as every sense organ is made of matter (potentiality or material or physical cause) and form (actuality or formal or mental cause), the sense organ is in potency to receive a new form (actuality). In perception or perceptual activity, the soul (formal and efficient cause), initiates the perceptual activity and causes the sense organ to take sensible form of the object of perception. For example, the soul of the eye, in the presence of a green car, will cause the eye that is in potency to see any colour, to abandon its form and take the form of the green colour of the car. In this case, the eyes sees a green car.

The capacity of thought is another area where Aristotle's concept of mental causation is exhibited.

6.4. Thought (Intellect)

Aristotle holds that thought or intellect or some part of it is separable from the body. But, intellect in its activity (thinking) may parallel a perceptual activity. "Just as perception involves the reception of a sensible form by a suitably qualified sensory faculty, so thinking involves the reception of an intelligible form by a suitably qualified intellectual faculty" (Shields, 2016: 13). Thus in a schematic outline, this is what happens when one thinks about an object *O:*

> S thinks O if and only if: (i) S has the capacity requisite for receiving O's intelligible form; (ii) O acts upon that capacity by

> enforming it; and; as a result, (iii) S's relevant capacity becomes isomorphic with that form. (Shields, ibid.: 13)

However, intellectual activity or thinking occurs at an abstract or higher level than perception, and it encompasses all things as intellect thinks all things, without perceiving them. The intellect, unlike perception, is not affected or becomes like what it thinks; it abstracts only the forms (ibid.: 14).

How can the intellect take on an intelligible form while itself remaining unaffected? In *De Anima* 3.5, as explained already, Aristotle sets out to address the difficulty by introducing a logical distinction between the active intellect and the passive intellect. He also uses his theory of four causes to explain the difference between the passive intellect and active intellect.

The passive intellect is a sheer potentiality to receive an intelligible form and as a potentiality, it needs an actuality to cause it to realize itself as no potentiality brings itself to actualization; the bronze in a potential state cannot cause itself to be a statue, it needs an actual agent, the efficient cause to do that. Thus, the passive intellect as a potentiality needs an efficient cause to cause it to actualization, and this is the active intellect. Thus, the mind in its potential state is the passive intellect and in its capacity as an actuality (form as efficient cause, formal and final cause) is the active intellect.

For Aristotle, the active intellect thus produces all intelligible forms while the passive intellect can potentially become any intelligible form produced by the active intellect. He further characterizes the active intellect as unmixed and unaffected, thereby refining his earlier claim that the intellect must be unaffected and unmixed (DA429a18-29). The passive intellect is affected and mixed. It becomes the intelligible forms and so is affected; it is also the case that the passive intellect perishes with the body, and so it is mixed with the body. The passive intellect is a corresponding matter to the form, which is the active intellect. The passive intellect is the matter of the faculty of intellect; it is the organ that is enformed by the intelligible form. On the other hand, the active intellect is what is immortal and eternal; its essence is activity, and this cannot cease. Aristotle compares the active intellect to a light in that as we need light to see colours, so we need the active intellect to know the intelligible

forms; the active intellect is a kind of 'light' which operates and actualizes our potential to know and to think, and brings either to fulfilment (DA 430A24-25). Thus, Aristotle applies his theory of hylomorphism and the four causes to cognition (or what we may refer to in modern terms as the mind), whereby he pronounces the passive intellect as the matter and the active intellect as the form that causes the passive intellect to know.

The intellectual activity (thinking) is an instance of mental causation as the intellect is composite of passive intellect (matter – the brain) and the active intellect (the form), and it is the active intellect that causes the passive intellect to be actualized by providing it with intelligible forms that it uses in its intellectual activities such as thinking, analysing, judgement, critiquing, decisions making, etc. The active intellect as part of the soul which is still situated in the body even if it may or may not be viewed to be immortal, through its causal work, actualizes our potential to know or critique and bring it to full actualization and fulfilment. And the intelligible forms supplied to the passive intellect by the active intellect are forms of the sensible objects and non-sensible objects for intellect thinks all things. The intellect operates in conjunction with the perceptible sense organs because through them the intellect perceives the outside world, and abstract intelligible forms.

Another area where Aristotle shows the reality of mental causation is desire, which is the basis of his action theory.

6.5. Desire

Aristotle notes that perception and thinking are faculties or capacities of the soul, while perception helps us to perceive things or objects in our surrounding, and thinking helps us in thinking about the objects of thought and making judgements about them, however, neither perception nor thought can account for the fact that human beings and other animals act on their object of desire. The soul is involved in such movement because every ensouled being (being with a soul) has within it a principle of motion and rest (DA 405b11; 409b 19- 24). It is the characteristics of living beings to be able to move themselves in search of their objects of desire and run away from their predators and dangers (Shields, 2016: 16).

Aristotle argues that all animals are capable of motion or movement, but none of the three faculties of the soul, namely nutrition, perception and thought, can account for desire-initiated movement or goal directed behaviour. The nutritive faculty cannot account for such movement because plants, as living beings, with that nutritive faculty or power cannot move themselves for a goal-directed action. Perception cannot do that either since some animals have this faculty without ever moving themselves at all. And goal-directed behaviour cannot be the product of reason because reason is contemplative, it is not directed towards any action, though reason can approve an action, but it is insufficient on its own to put a human being into action, it relies on appetite to initiate a goal-directed movement (DA 432b14-33a5).

Reason, in Aristotle's philosophy, operates in two distinct ways namely, contemplative reason and practical reason. So, contemplative reason and practical reason are not two kinds of reason, but one reason working in two different ways (Hammond, 1902 1xiv). Practical reason is goal oriented for it engages in calculation for the sake of something, while contemplative reason is not goal-oriented, it engages in theoretical activity. "… Intellect and desire, are productive of locomotion, and the intellect in question is that which reasons for a purpose and has to do with action and which is distinct in its end from the contemplative intellect. All desire is also purpose directed. The object of desire is the point of departure of action" (DA 432b14-33a5). Hammond (1902: 1xiv) explains the distinction between the practical from the contemplative: "the function of theoretic [i.e. contemplative] reason is to discriminate between the true and the false; the function of practical reason is to discriminate between the good and the bad". The practical reason, primarily, deals with "deliberation and conduct" (ibid.).

For Aristotle, a faculty cannot explain a purposive action if its activity is insufficient to engender movement, hence even desire itself cannot initiate such a movement or be responsible for action. "For the self-controlled, though experiencing desire and appetite, yet do not do the things that they desire, but defer to the intellect" (DA 433a6-8). Therefore, Aristotle argues that desire alone, as single faculty in the soul, cannot account for a purposive action completely, rather desire and practical reason are two faculties of the soul implicated in animal movement (DA 433a17-19). However, the two do not work in isolation from one another.

Rather, practical reason considers an object of desire as something desirable while desired objects prick the practical intellect and set it in motion (DA 433a17-2). Thus, for Aristotle, it is the faculty of desire, which through its activities and objects is responsible for initiating end-directed motion in animals. What animals seek in action is some object of desire which is or seems to be good.

Thus, for Aristotle, the faculty of desire is involved in initiating human action, but desire is not the only faculty involved in the explanation of purposive human behaviour; practical reason, and also imagination, play indispensable roles in human action. Humans and other animals use the faculty of imagination for storing and recalling images for their cognitive activities and movements. And images stored in the imagination can motivate and guide in their activities (DA 429a4-7).

Mental causation is exhibited through desire because the soul is implicated in motion as it is a principle of motion and rest in every living being. So, living beings are capable of moving themselves for the good they desire or to run away from the danger that threatens them. However, it is the three faculties of the soul namely, desire, practical reason and imagination (formal or efficient cause) that cause living beings, especially human beings, to move themselves in ways likely to result in their survival and flourishing. The desire together with practical reason and imagination cause human beings to act.

I have explained Aristotle's account of mental causation through his accounts of perception, thought and desire. But the question one might still ask is whether the soul is the phenomenal consciousness and as such is causally efficacious, which the functionalists like Kim and Davidson were unable to account for adequately in their theories of mental causation. I will show in the next section that the soul or immaterial intellect is what I take to be the phenomenal consciousness, and argue that it is causally efficacious in causing human action.

7. The Immaterial Intellect as the Modern Notion of Phenomenal Consciousness and the Causal Efficacy of Self-Consciousness

I stated in chapter 1 that in this thesis I am going to argue that the phenomenal consciousness is causally efficacious. I have discussed Aristotle's theory of hylomorphism and his theory of the four causes. We saw that for Aristotle the soul as form, is the efficient cause, the formal, the final cause. It is also the immaterial active intellect.

The question one might ask now is whether the notion of the soul as the immaterial active intellect, is the same as the modern notion of phenomenal consciousness, the qualitative aspect of consciousness. It is generally held that Aristotle did not specifically talk about consciousness, let alone phenomenal consciousness. I argue that the immaterial active intellect is the phenomenal consciousness which is known among the classical philosophers as self-knowledge or, to use modern terms, as self-consciousness. I will engage with some neo-Aristotelian philosophers to argue that the rational soul (the intellect) is the phenomenal consciousness. Aristotle might not have used the term phenomenal consciousness, but I argue the classical philosophers have something equivalent to it, which is self-knowledge.

I indicated earlier that the way Aristotle presented a sense-perception act causes people to think that Aristotle is talking about consciousness, but Aristotle nowhere in *De Anima* discusses consciousness *per se*. Hence, philosophers such as Tancred claim the modern idea of consciousness cannot be imposed on Aristotle (Lawson-Tancred, 1986: 78). I argue however, that Aristotle's idea of perception could be described as a simple form of consciousness since awareness of sensible form is entailed.

However, Victor Caston (2002:451) notes that Aristotle analyses perceptual awareness in terms of high-order intentional states, such as perceiving that we see or that we hear. But unlike those theories he denies that this is due to a second token mental state being directed at the first. Instead a perception is directed at itself as well as to a perceptible quality in the world. It possesses high order content in addition to its original first-order content. So while this awareness is high-order and relational, it is also intrinsic to the act of perception, because it involves a reflexive relation. An 'inner sense' is thus not required for this purpose either.

Sanguineti (2013) in his article, "*The Ontological Account of Self-Consciousness in Aristotle and Aquinas"* argues that notions of consciousness and self-consciousness are modern terms, but their equivalent in classical philosophy is self-knowledge. He notes that consciousness relates more to psychological states, while self-knowledge is wider in meaning and is more concerned with knowing the truth about one's being. He argues that self-consciousness is viewed fundamentally in Aristotle as a strong form of being, which means self-possession in an ontological sense.

According to Aristotle, the five senses – sight, hearing, taste, touch and smelling – deal with sensible objects of the external world (ibid.312), but the sentient subject, through a capacity Aristotle called 'common sense', can feel or perceive the operations of the five senses simultaneously (DA 425b12-27). As perception or sensation is a nonintellectual activity, Juan calls such perception of sensory operation 'sensitive consciousness' (ibid.313), while Juha Sihvola (2007) calls it 'perceptual consciousness. But Aristotle writes:

> Now every sense has both a special function of its own and something shared with the rest. The special function, e.g., of the visual sense is seeing, that of the auditory, hearing, and similarly with the rest; but there is also a common faculty associated with them all, whereby one is conscious that one sees and hears (for it is not by sight that one is aware that one sees; and one judges and is capable of judging that sweet is different from white not by taste nor by sight, nor by combination of the two, but by some part which is common to all the sense organs…). (On Sleep and Waking, 455a13-21).

Sanguineti notes that modern philosophical and scientific writings do not make a distinction between sensitive and intellectual knowledge, as there is no clear distinction between the senses and the intellect. However, with Aristotle's concept of common sense that makes the perception of sensitive intentional operations possible, perceiving that one sees or hears is different from understanding that one sees or hears. Thus, a distinction is made

between sensitive and intellectual knowledge. Consequently, in line with Aristotle epistemology and metaphysics, sensitive consciousness is different from intellectual consciousness. Hence, animals can experience sensitive consciousness, but not intellectual consciousness, as they are not rational. But they are asleep and awake (ibid.).

Sanguineti points out that modern philosophers and scientists associate perception with sensation unlike the classical philosophers. However, Aristotle's *aesthesis* is either translated as perception or sensation, although for him, perception includes judgment as pointed about above, while sensation does not. But in Aquinas, the verb "to perceive" (*percipere*) is applied to the sense and the intellect (ibid.). In that when one sees a person, "one both perceives him with one's eyes and intellectually recognizes him as a person" (ibid.: 314). Sanguineti (ibid.) argues that the sensitive consciousness connotes that consciousness and intentionality are mutually associated at the sensible level as perception of the external world is inseparable from the perception of one's body as a passive subject receiving data from the outside, and as an active subject endowed with the capacity of modifying the objects displayed in the physical world (ibid.). Thus, one is in action when one perceives.

Sanguineti notes too that Aristotle acknowledges intellectual consciousness in his ethical writings, for he holds that a person can be aware that he acts, thinks, and exists (ibid.:315). Thomas Aquinas, elaborating Aristotle's account of human self-consciousness, looks at it from the point of view of self-knowledge. For Aquinas consciousness as an act, knowing that one is thinking, is a second-order cognitive operation, which is absent in animals. Following Aristotle's account of consciousness (with all the disclaimers above), sensitive consciousness is the first step of intellectual operations (ibid.: 317). The second-order operation is reflection. This level of consciousness occurs when the intellect moves from the external world to internal subjectivity, thereby reflecting on its actions: "our intellect operates upon itself producing a reflection" (Aquinas as quoted in Sanguineti, ibid: 318), it reflects on itself, and in this way "is capable of understanding its own understanding" (ibid.). Self-knowledge in Aquinas is equally an intellectual act through which a perceiver perceives his/her own mental (intentional) operations in relation to him/herself and surroundings, and s/he is able to understand these

mental operations with regard to him/herself: " the intellect perceives that it understands", and that it "understands that it understands"(ibid.). Thus, self-comprehension is an intellectual perception as no one else can feel a person's knowing or understanding something.

Self- consciousness or self-perception is seen as a judgment too as long as understanding means having the power to judge the truth or goodness of an action or event. Thus, self-understanding suggests that one has the capacity of judging one's actions as true or wrong, with the aim of affirming or correcting them in order to act on them. With self-knowledge, one can judge their own judgments and direct their rational behavior, upholding or correcting their choices whenever it is necessary to do so (Sanguineti, ibid: 318). "Self-consciousness is the basis of free will" (ibid.319). To be self-conscious means to be self-guided, notwithstanding the possibility of making wrong judgments. Thomas Aquinas writes:

> Those who free judge are those who can move themselves according to their judgment. But in order to be able to judge its very act of judging, a faculty needs to judge itself, it must know its judgment. This is the characteristic of intelligence. Irrational animals are free in a certain sense only regarding their movements and acts, but they have no judgment (ibid.: 319).

According to Aquinas, self-perception of one's own acts enables the perception of one's own life and existence as such, which implicitly means the perception of one's self, that is, of the concrete person who is thinking and living. Sanguineti (2013) notes that Aristotle makes a related point in his ethics: "whenever we perceive, we are conscious that we perceive, and whenever we think, we are conscious that we think, and to be conscious that we are perceiving or thinking is to be conscious that we exist" (Aristotle, *The Nichomachean Ethics*, 1962, 1170a32-3). Aquinas includes in self-apprehension the internal experience of having a mind: one perceives that he has a soul, that he lives, and that he exists, because he perceives that he senses, understands, and carries on other vital activities of this sort" (Aquinas as quoted in Sanguineti, 2013 :321).

David Braine (as quoted in Sanguineti 2013, 321) suggests that the soul that is internally experienced, is reached through a phenomenological experience. It was because of phenomenological experience that the ancient authors especially in the Bible speak of spirit, soul, and heart and mind (Sanguineti 2013, 321). Hence, we have biblical expressions like this "my soul proclaims the greatness of our God" (Luke 1: 46-7), "my soul is sorrowful to the point of death" (Matthew 26:38). Sanguineti (2013: 322) notes that the phenomenological conception of the soul is not dualistic and it is compatible with the Aristotelian notion of the soul as form of the body in the sense that it cannot exercise its power in the absence of the body. It is through the body that it gets access to the external world.

I thus argue that it is phenomenological experience or self-consciousness that is immaterial and causally efficacious. We see its causal efficacy in causing human action. It causes humans to perform intelligent, rational and conscious actions. It is the soul as form that defines the kind of life a human being should live. Aristotle would say humans have a function, and that function is to live a virtuous life. If Aristotle argues that humans have a function to live a virtuous life, does it mean then that Aristotle is a functionalist? Let us now discuss whether Aristotle is a functionalist or not. This will be helpful in discussing his notion of mental causation in relation to Kim (as realizer functionalist) and Davidson (as role functionalist).

8. Aristotle as a Functionalist

I will base the argument on the S. Marc Cohen's article. Cohen (1992), in the article, *"Hylomorphism and Functionalism*[i]", analyzes Myles Burnyeat's critique of a functionalist interpretation of Aristotle's philosophy of mind. Cohen argues that a functionalist interpretation of Aristotle's hylomorphism holds. However, I will argue that Aristotle should be seen as a 'soft' functionalist.

Cohen (1992: 2) firstly defines functionalism as a theory that views mental states in terms of their causal relations between mental states and other mental states, sensory inputs and behavioural outputs. For instance, my desire (mental state) to go on holiday may cause me to want (mental state) to buy new clothes. Or, I mistakenly stick my fingers in a hot water

(sensory input) and cry out 'ouch' (behavioural output). Functionalism holds that the same mental state is realizable in several different physical states or process. Mental states cannot, therefore, be reduced to physical states. They are the functional states of their physical realizers (ibid.: 2), as already explained in chapter 2.

Cohen (ibid.) notes that Aristotle had no interest in talking about mental states but refers only to the soul (*psyche*). His theory of hylomorphism states that the soul is related to the body in the same manner that form is related to matter. Aristotle's idea of the soul is biological; the body is a living thing in virtue of the soul. Moreover, the soul for Aristotle is a substance not in Descartes' sense of being separable from the body. However, it is a substance in the sense of form.

Form in ordinary language means shape but in Aristotle's philosophy, it is a functional organization, as explained in the beginning of this chapter. For instance, in the case of a statue, the bronze is the matter while the shape, arrangement, or capacity of the matter is the form. Human activities are explained in relation to the activities of the soul, as explained in previous sections. Cohen (ibid.: 3) believes that there are elements of functionalism in Aristotle's account of the soul in that mental faculties and states are materially embodied except for the active intellect. And these mental faculties are defined in terms of their functions rather than in terms of their material components.

For Burnyeat the theory of homonymy – a theory that states a body that is not actually alive is a body in name only, so an eye that cannot see is not really an eye but only in name – as making a case for Aristotle's being anti-functionalist. Burnyeat argues that for Aristotle, it is impossible to pick out a body that is not actually alive (a body only in name) because according to Aristotle, matter must always have a form. This form is the soul or substance in virtue of which the matter is alive. There is no body without a form. Therefore, the homonymy principle is incompatible with the hylomorphic theory of Aristotle. There is never a matter waiting to receive a form because every matter is compounded with a form; is essentially alive. The form is tied to the matter. So there is no question of transporting the form to another matter, which

argues then against the functionalist who claims that any given mental state can be realized in different physical or material structures (ibid.:10).

This cannot be correct, as of course a dead body consists of matter and has the form of lump of flesh. This is an example of substantive change. It seems as if Cohen agrees. According to Cohen (ibid.: 12), Aristotle uses the homonym principle to show the importance of function (form) in the definition of a living creature. A living creature or organic system that fails to perform its function that is based on its form is no longer a living creature but creature without its normal function. It is now a creature only in name without functional states or work. The functional capacity supervenes on the material body; the body that is composed of the four elements (earth, air, water and fire) realizes it. The form or functional capacity determines the kind of physical structure a body has. When the form is no longer there the physical structure that instantiates the functional capacity will not be there, and we are left with a compositional matter that will receive a new form.

Cohen (ibid.: 13)) concludes that Burnyeat has not killed off a functionalist interpretation of Aristotle philosophy. A functionalist interpretation holds that living organisms are alive or functional in virtue of their soul or form. Thus, the properties and behavior of an organism are explained in terms of the functional properties of its material components (ibid.). However, Cohen (ibid.) is not shy to point out that the functionalist interpreters are yet to explain the causal role of the soul (efficient cause) in Aristotle psychology. And for Cohen, this is a hard nut to crack; how is the soul (*psuche*) an efficient cause in Aristotle's psychology?

However, I argue that Aristotle should be considered a 'soft' functionalist because he sees the soul as the efficient cause (formal and final cause) that causes a living body to perceive, think and desire, and as the self-consciousness that causes human action as I argued in sections 6.2 -6.5 and 7 above. For Aristotle, the soul is a set of capacities or functions for perception, thought, and desire even for nutrition. I think that Aristotle should not be seen as a functionalist in the strict meaning of the term. To take Aristotle as a contemporary functionalist will defeat the meaning of the term soul (*psyche*) in Aristotle's psychology, because contemporary functionalism has no room for consciousness or awareness which Aristotle's

concept of perception or sensation entails, and it accounts for phenomenal consciousness (self-consciousness).

Caston (2006: 325) notes that Aristotle cannot be a functionalist because he holds psychological states and the soul itself as efficient causes while the functionalists, in contrast locate the real causal power in the underlying material states. For the functionalists, the mental causation works from "from the bottom up" (upward causation) but for Aristotle, it is "from the top down" (downward causation). As we saw in chapter three, for Kim downward causation implies that the causal work is done by the physical properties. The causal power of the mental is dependent on the causal power of the physical. I argue that downward causation by the mental *qua* mental is supported by Aristotle's concept of mental causation. Hence, Aristotle will save the causal efficacy of mental causation from Kim's criticism.

Thus, Aristotle believes that the soul *qua* soul is causally efficacious; its causal power is not reducible to bodily properties (DA 1.4., 407B34-408a5), although its causal power allows its matter to be functionally arranged in a particular way. Aristotle argues that the soul alone is responsible for the oneness of a living thing: on their own the material elements that constitute the living things will disperse to their natural places (DA 11.4, 416a6-9). However, for Aristotle, the mental sates, such as desire and *phantasia,* are the proper efficient causes of action and speech, while the underlying material states have a merely instrumental role (MA, 6 700b17-20; DA 111.10 433b13-27) (Caston, 206: 326). This confirms that Aristotle prefers downward causation. Hence, I propose that Aristotle be considered as a 'soft' functionalist because he defined the soul as form in functional terms.

While Kim and Davidson emphasizes the role of mental states (desire, belief and intention) in causing human action, Aristotle's emphasis is on the role of self-consciousness that entails the whole human person as a physical, emotional, intelligent and conscious (self-aware) being. To this effect, Aristotle's theory of human agency is bound to be different from that Davidson and Kim. For Aristotle, self-consciousness is the cause of human action. It implies that the agent is completely involved in his action, and that the goal he set out for himself causes his action in order to realize that goal. His action is for him and about him in relation to the society

in which he lives. The function of man, according Aristotle ethics is to be virtuous in order to live a happy life. It is the phenomenon of self-consciousness that causes man to be virtuous.

It is on this ground that I argue that Aristotle is a soft functionalist, not a role functionalist like Davidson, nor a realizer functionalist like Kim. He is what I call a 'soft functionalist" as he believes in the function of the soul, whose causal power does not depend anyhow on the matter, though the soul as a form cannot be without the matter. For Aristotle, the soul or mind or intellect is immaterial, but causally efficacious in the physical world. Since the soul is immaterial, it is not subject to the Humean idea of causation that conditions Kim and Davidson's causal theories.

Before I go on to consider the reasons why I think that Aristotle offers a better theory of mental causation than Kim and Davidson, let me now consider some criticisms against his theory of hylomorphism in the next section.

9. Criticisms of Aristotle's Hylomorphism

Christopher Shields (2016) in the article, "*A Fundamental problem about Hylomorphism"*, notes that there is a significant problem for Aristotle's theory of hylomorphism with regard to the human body. He claims this problem was noted first by Ackrill (1972: 3). Aristotle's hylomorphic analysis of change and generation suggests that when a lump of bronze receives the form of Hermes, we have a statue of Hermes. And when this statue of Hermes is melted and given a form of Domitian, then we have a statue of Domitian because the statue of Hermes has been replaced. Thus, the bronze is contingently enformed with either the shape of Hermes or Domitian (Shields, 2016: 1). Nevertheless, the one lump of bronze is persistent throughout the process of change as it is essentially bronze or metal but not essentially Hermes or Domitian shaped. Thus, the persistent matter of change is only contingently enformed with the shapes it acquires during the process of change (ibid.).

Thus, the matter underlying a change has only contingent form during the process of generation. This seemingly becomes a problem for Aristotle in relation to soul-body relations as

he argues that the human body is essentially enformed by the soul whose body it is because, a body, the matter of a human being, cannot lose its form, its soul and still be in existence like bronze. This problem is created by Aristotle's stance that a body that has lost its soul is not a real body "except homonymously" (DA 412b10-24). By homonymy, Aristotle means that the body in this case is like an artificial (manmade) eye in a painting, that looks like an eye but it does not see; it is not real body of human being with its soul. For Aristotle, only an ensouled body is potentially alive, whereas a lump of bronze is potentially a statue of any shape anytime (Shields, ibid.: 2).

The appeal to homonymy by Aristotle shows that every human body is essentially (not contingently) ensouled, and the human body becomes dead once it loses its soul; it becomes like a statue. However, the problem is that the hylomorphic account of change holds that bits of matter are contingently enformed; the bronze is not a bronze because of either of the shape of Hermes or Domitian. But bronze is the bronze it is because of its material constituents namely an alloy of copper or tin that persists through any change (ibid.). Here is the bone of contention, if, without their souls, human bodies are not bodies that are potentially alive, then the hylomorphic treatment cannot be applied to human bodies. Shields (ibid.) notes that Ackrill holds that according to hylomorphism, matter is contingently enformed, and if it is so, then bodies of human beings should be contingently enformed, not necessarily or essentially enformed. On the other hand, the homonymy argument shows that human bodies are necessarily enformed or actually alive and they become dead once they lose their souls. Thus, human bodies are both contingently (hylomorphism) and necessarily (homonymy) enformed.

This is a seeming contradiction on Aristotle's part. The contradiction is due to the appeal to homonymy. Shields (ibid.) argues that the contradiction can be eradicated by abandoning the appeal to homonymy (losing the soul means existing homonymously). Then the human body, just like any other matter, will only be contingently enformed as hylomorphism holds. This will be difficult for Aristotle as that will be tantamount to abandoning his philosophy of mind. Human beings will not be able to perform life functions characteristic of them such as eating, perceiving and thinking (ibid.).

So, Shields (ibid.) continues and proposes a solution to resolve this problem, which is the combination of hylomorphism and homonymy. This will result into two (kinds of) human bodies, one contingently ensouled (contingently alive) and another necessarily (actually alive) ensouled. Or in Aristotle's terms, this is what he refers to as non-proximate and proximate matter. Non-proximate matter undergirds the actual matter, the proximate matter. For instance, we have the bricks and mortar as the matter of the house, but also the bricks have their own matter, perhaps the clay. Then the non-proximate matter of the house will be the clay of the bricks while the bricks will be the proximate matter of the house.

Shields notices that this may not be clear with respect to a living being whose proximate matter is already complexly structured. Thus, beneath the complex matter will be non-proximate matter, which is only contingently enformed (ibid.). The proximate matter will be identified as the organic matter (the fully formed and living matter, of a living human being) while the non-proximate is non-organic. The non-organic matter would be similar to the bronze that persists through hylomorphic generation; through the acquisition of various forms. Conclusively speaking, a living human being has two bodies then – one organic and one non-organic. Shields claims that this parallels the way the 'flesh' is spoken of. We speak of living flesh healing itself when damaged. Thus, the organic body is a living body of which we are aware when we speak of the body, but there is a non-organic body that goes all the way to the flesh (ibid.: 4). Shields concludes that despite the seeming contradiction, and that his solution to resolving it might be at variance with common sense, the hylomorphic project still looks plausible. Nevertheless, the problem generated by Ackrill still poses a problem to our understanding and evaluation of Aristotle's hylomorphism in philosophy of mind.

I argue that the application of hylomorphism to living things and non-living things should not be the same, however it should be applied based on the nature of that being. Non-living things have their efficient cause outside them while living things have their own efficient cause inside them. This is suggested by the fact that Aristotle makes it clear that in the case of living organism, the efficient, formal, and final causes fall together, while this is not the case for non-living objects.

The body of a human being is ensouled; it is with its form, the soul. When it loses its soul it becomes a corpse, but not a human being, the ensouled body does not persist during the process of change. Shields wants to bring in the notion of non-proximate matter to represent the matter that persists like bronze during the change. The question is "is this non-proximate matter, part of the ensouled body" (ibid.)? If it is not, how do we differentiate the necessarily ensouled body (proximate matter) and contingently ensouled body (non-proximate matter) in a living human body when Aristotle holds that the living things have their efficient cause (the formal cause and final cause) as their soul internally? Is it possible to have two souls operating differently in one living thing; one in the proximate body and the other in the non-proximate body; or is one soul operational only in the proximate body (the necessarily ensouled body), not in the proximate body (contingently ensouled body)?

For an artifact like a house, the efficient cause that gives it the form of a house is external to it. So it is easy to talk about the proximate matter and non-proximate matter in this case. The point being made is that in living things with the soul internally structured, it cannot have two kinds of matter or body serving as one material cause. To have two kinds of matter as a material cause will be having one soul working on two different bodies in one living thing. So it becomes hard to tell that this particular thing has no soul and that one has got it. The solution proposed by Shields will only create confusion about hylomorphism.

I go further than that, because I claim that it should not be called a solution to start off with because there is no contradiction in the first place. As long as a living thing is different from a non-living thing, there is a need for a differentiation in how Aristotle's causal theory is applied to explain existence, and change of these individual substances. It could be argued that Aristotle is quite aware that the case of living things like animals and humans is quite different from the artefacts like bronze. In this case, he is not explaining a substantial change in the artefacts but to explain the development of living things.

Aristotle uses hylomorphism to explain change and I think he is successful in doing that. Aristotle's hylomorphic project helps us to understand generation and change in both non-living and living things. In non-living things, the cause of change is an external agent that gives

the matter a form or shape. In living things, the cause of change is internally glued to the matter of the body, which causes any change, be it qualitative, quantitative or locomotive change, but once this efficient cause (the soul or form) is lost, a new substance is generated, in the case of a human being, the dead body or corpse, which is a non-living thing. It has naturally taken a form or shape of a corpse. But in the case of a non-living thing like bronze, it loses the shape of Hermes when an external agent acts on it.

I think that application of the hylomorphic theory should not be the same for non-living things and living things because their efficient causes do not explain their change in the same way. The efficient cause for living things, like a human being, acts from inside the human while that of non-living things acts from outside. The change in living things are caused from inside and the change in non-living things are imposed on them from outside by an agent external to them. The problem that Ackrill contends, becomes real only if we use one application of hylomorphism for living things and non-living things.

Secondly, Christopher Shields (2016) in the article, "*Controversies surrounding Aristotle's theory of Perception"*, looks at the controversies around Aristotle's concept of perception. I argue that there is no controversy about perception in Aristotle's philosophy of mind; a controversy will arise only when we try to read into Aristotle's metaphysics what Aristotle's intention was not originally. The issue is Aristotle's "claim that in sense perception the relevant sensory faculty becomes like the object it perceives" (ibid.: 1). "What can perceive is potentially such as the object of sense is actually" (DA 418a3-4). It does mean that a sense organ in one way or another becomes like its object when it perceives. There is difficulty in understanding exactly what this likeness is supposed to mean. Shields argues that a proper evaluation of Aristotle's conception of perception needs to be made, specifically on what he means by 'likeness' (Shields, 2016: 1).

Shields notes that there are those who understand 'likeness' literally, which implies the sense organs become literally the colours they perceive. Thus, when the eye perceives red, the eye jelly, the matter of the inner eye, itself becomes literally red. "The eye simply comes to exemplify the colours present in the objects its fields" (ibid.). The most noteworthy proponent

of this approach is Sorabji (1974). This approach understands the literal likeness as a necessary condition for assimilation of form.

Another approach to understanding Aristotle's' position is non-literalist interpretation. A proponent of this approach is Brentano (1867). This alternative approach, which may be called the intentionalist interpretation, states that "the sense organs become like their objects without actually coming to exemplify the sensible qualities perceived" (Shields, 2016: 2), and instead the sense organs symbolize them in a certain way. Here, the likeness involved in the sense perception is akin to the likeness of a house and its blueprint, whereby the blueprint does not exemplify the property of a house, but rather encodes it. Things can be like to each other with respect to a particular property without exemplifying that property. Aristotle's concept of 'likeness' may be understood in that form as he speaks of the form in the craftsman's soul, which he can impart to some matter made to symbolize the form in his soul. Hence, form reception for Aristotle in sense perception may be a case of the sense organ symbolizing the perceptible form, which is not a literal property transference (ibid.: 3).

Shields (ibid.) enumerates the advantages of this approach over the literal interpretation. The claim that organs of perceptions, in sense perception, exemplify the sensible qualities they perceive is taken care of or addressed. With this approach, there is room for the phenomenon of selective attention. This is so as a perceiver will not be under any obligation to encode any property before her. Rather she becomes intentionally like only some subset of properties in her immediate surroundings (ibid.). However, this approach ignores the fact that Aristotle does not distinguish between the two approaches, and does not prefer one to the other. On this intentionalist interpretation, if form-reception and isomorphism are to be understood intentionally, is this kind of form reception adequate for perception, and if this kind of form-reception is not sufficient for perception, what else is required for an adequate perception (ibid.)?

I argue that it is not possible to notice the true colour of the perceptible form in the perceiving sense organ. So, I argue against the literal interpretation. I ask, "Who is supposed to observe the colour of the perceptible form in the perceiving organ, is it the perceiving person"?

Is it possible for a perceiving person to observe the perceptible colour while perceiving? Either she is perceiving, or she is observing the perceptible colour; she cannot be perceiving and observing at this time. If one who is perceiving a colour is asked, "What colour are you perceiving"?, she will leave the perceiving she is actually doing to focus on answering the question, thinking about the words to use to describe her experience of perception. In this case, she will only describe the mental image, the percipient object in the memory, which is not the actual perceptible form. Thus, it will be difficult to determine the right or correct perceptible form in her form perception.

We might say let us call a third person to look into her eyes to observe the colour of the perceptible object, but if this were possible, then the subjective experience that goes with any sense perception will be seriously undermined. In this case, we will not be working with the experience of the first person, but the third observer. I think due to this difficulty in obtaining the correct form of perception, Aristotle may not see it as necessary to distinguish between different approaches of form perception.

As a third point of critique against Aristotle, Georg Gasser (2009) in his article, *"Is Hylomorphism a Neglected Option in Philosophy of Mind*?", argues that Aristotelian hylomorphism is not the best alternative theory of mind to post-Cartesian philosophy of mind (i.e. contemporary philosophy), as it does not provide a metaphysical ground for the mind-body problem. I argue that Gasser's argument is a failure in that it is post-Cartesian philosophy that does not provide the metaphysical ground for mind-body problem.

Gasser (ibid.: 1) notes that modern philosophy of mind is built on two basic points namely (1) what he calls the dichotomy option and (2) the privileged access option. These two points are due to Descartes' distinction between *res cogitans* (thinking and unextended substance) and *res extensa* (extended and unthinking substance) as basic characterization of the mental and the physical (ibid.).

The dichotomy assumption is the fact that "modern philosophers conceive the mental and the physical as two different conceptual frameworks that are not reducible to each other" (ibid.:2), whereby physical phenomena are governed by physical laws and mental phenomena

are not under the governance of the physical law, but rather governed by the principle of rationality and coherence (if we follow Davidson for example). An example of this dichotomy is Davidson's anomalous monism of the mental that holds that the mental is irreducible to the physical domain. In terms of the privileged access option, only the subject has direct access to her mental states. A person is a subject of her experience, and her experience is not known to the public (ibid.).

Gasser notes that dualism holds that mental states are distinct from physical states, and they cannot be reduced to physical states. Reductive physicalism holds that mental states are reducible to physical states. However, Kim argues that complete reduction of the mental to the physical is impossible. It is some parts of the mental such as intentional and cognitive properties that are reducible while the qualitative properties such as qualia are not reducible (ibid.: 4). (See also chapter 3.) Gasser (ibid.) notes that both of these theories are based on the mental and physical divide and the consequence is a distinction between inner /outer, subjective/objective and privileged/public, due to Cartesian dualism. Thus, we have these two main divisions namely dualism and physicalism in the philosophy of mind.

He (ibid.: 5) notes that Aristotelian hylomorphism is thought to be an alternative view for understanding the mind in neither a dualistic nor a physicalist manner; it is seen (by some) as a solution to the mind-body problem created by Cartesian dualism. He gives a brief account of Aristotle's account of soul. He notes that Aristotle defines the soul as the form of the body or principle of life in living things (DA 412 a15-29). Aristotle's concept of 'soul' embraces all living things in general; it is not particularly referred to any specific mental state. He notes that for Aristotle, 'soul' is not an entity attached to the body, but is its form or nature. It is the principle of life in things that is responsible for the thing's self-developing, self-maturing and self-moving (ibid.).

On 'matter', he (ibid.) notes that for Aristotle, matter and soul are inseparable in living organisms. The concept of soul is intertwined with the concept of matter. There is no form without matter or matter without form. "Aristotelian matter is not prior to specific things but 'last' in the sense that it is the closest to the form" (ibid.). The form is the first actualization of

the matter that endows the matter with certain capacities. Thus, a living being for Aristotle is a unity of matter and form.

Gasser, in enumerating the advantages of Aristotle's hylomorphism over dualism and physicalism, notes that the separation of mind and body is not possible in Aristotle's philosophy as the living being is the primary subject of investigation, and not its mental constituents and properties. The soul is responsible for the organism's unity and its self-maintenance and growth. And the soul in the human being is responsible for the faculty of reasoning. Matter does not have its own existence; it exists alone only by logical abstraction from the notion of form, which is an internal principle of organization and changes. Thus, Aristotle's idea of matter is different from the modern concept of matter (ibid.: 8).

Gasser (ibid.: 9) holds that Aristotle's idea of an organism's body as a unity of matter and form means the organism's body is seen as a structured sum of its material constituents and this makes the notion of form as internal principle of organization and development of the body superfluous (ibid.). Consequently, "the body and its functions become the subject matter of natural science describable from a third perspective" (ibid.). The mental capacities from a first perspective are ignored and omitted. I think that Gasser fails to understand what Aristotle means by form, the essence, of a substance. The form (soul) or formal or final cause or efficient cause in the case of living being is never public like the matter (material cause). Without the form, there will not be a living organism, and no structured sum of material constituents. It unites and organizes all the different material constituents into a particular living organism. Natural science can study the material cause but not the formal or final cause of the living being. That is why Aristotle suggests two approaches to the studying of any living being namely philosophical analysis and physical analysis (e.g. DA 403a 28-30; Phy 194a21-26).

Gasser (ibid.) argues that in the hylomorphic theory, the capacities or functions rather than individual mental capacities are highlighted; there is no singular treatment of the mental. There is no real separation of the faculty of reason from other faculties that humans share with animals. Thus, the three different souls are interlocked; reason implies the activities of the nutritive and sensitive souls, and the sensitive soul entails the nutritive soul. "… we could not

reason as we do without our sensory apparatus or sense as we do without our locomotion system" (ibid.). He argues that this is not foreign to modern science. For in modern science, there is progressive development of capacities or organs from less complex ones to more complex capacities, and cognitive science shows that there is an intimate connection between bodily organs and mental states. Even modern developmental psychology teaches that the full capacities of an organism usually progresses from less complex ones into complex vital organs (ibid.).

In post-Cartesian philosophy, mental phenomena occupy a prominent place and the contents of the mind are taken as mental or conscious activity not as a faculty. Gasser (ibid.: 10) cites as an example, Davidson's causal theory of action, where he claims reasons for action ought to be viewed as causes of action if reasons are to play an informative role in our explanation of action. But, according to Gasser (ibid.), in the hylomorphic theory, the notion of mental causation is neglected, as human actions are described neither as something mental nor physical but just as the exercise of human capacities that at once are both psychological and physical. Hylomorphism tells us what an individual can do with the capacities that he is endowed with. For Gasser (ibid.: 11), this is like finding out the organic structure that enables the exercise of those capacities. For him, it is the job of "natural science to investigate these structures and detect the material elements that make the organism's behaviour possible" (ibid.). Hence, in the hylomorphic project, "human action in particular, and the behaviour of an organism in general, are conceived as multi-structured phenomena that can be assessed from different perspectives" (ibid.).

But, I counter that Aristotle argued for mental causation from a different perspective. He did not base his argument on the Humean concept of causation that conditioned most standard theories of action, which cannot account for the phenomenal consciousness as causally efficacious in the physical world. For causal theories based on the Humean theory of causality, phenomenal consciousness is epiphenomenal. But for Aristotle it is causally efficacious in the physical world. I thus do not agree with Gasser that Aristotle neglected mental causation.

Gasser does mention that the divide in the philosophy of mind is due to Descartes's bifurcation of reality into material and immaterial phenomena that gave rise to the mind-body problem. Descartes, a dualist, holds that the mind is separated from the body. And as a response to his theory we have physicalism that over emphasizes the physical in that it sees the mental as physical. Aristotle's hylomorphism unites the two realms, for it holds every substance is made of matter and form. Subsequently, a human being is made of body (matter) and form (soul). In terms of hylomorphism, there is no problem of interaction between the mental and the physical. I argue that hylomorphism values the mental more highly than post-Cartesian philosophy does. The mental, as a unity of form and matter, since it is embodied in the appropriate body, can think or imagine all physical things as well as non-physical things. Thus, hylomorphism helps us to understand why the passive intellect actualized by the actual intellect (mind) is able to think all things. So in this sense, mind or reason (thought) becomes a capacity of the soul. It is able because it is a unity of matter and form.

Post-Cartesian philosophy that separates the mental from the physical, needs to prove how the mind that is separated from the body is able to think about or imagine physical things. As the mind is separated from the body or the physical, it may not be able to think about the physical, as there is nothing uniting the two. Thus, the power of the mental is limited in the Post-Cartesian philosophy. Though the intellect in the hylomorphic theory can think about and imagine anything, still as an active intellect**,** it is separable as it is unmixed with any matter**.** Aristotle shows this in his argument about the passive intellect and active intellect.

Gasser however furthermore argues that there is a dichotomy between living things and non-living things in the philosophy of Aristotle in the same sense that there is a dichotomy between the mental and physical or immaterial or material in contemporary philosophy. I argue that the dichotomy in Aristotle's philosophy is between different areas of existence or nature, i.e., between non-living things and living things. The dichotomy is not within the same things characterized as living things or non- living things. In a living thing, for Aristotle, there is no dualist dichotomy of division between the mental and the physical or the immaterial and the material in living things, rather there is a unity of matter and form. The form, the soul, does the

same work; it is the principle of life that organizes all living things into the functional beings they are. The dichotomy in post-Cartesian philosophy is however within an organism.

Finally, Gasser (ibid.: 15) argues that Aristotle largely ignores consciousness which plays a very important role in modern and contemporary philosophy. He holds that hylomorphism succeeds in maintaining the unity of the human person (and of other sentient animals) only because it ignores a metaphysical analysis of consciousness. Descartes emphasizes the first person perspective – the subjective experience that someone is conscious and able to experience her surroundings while the growth, photosynthesis, digestion and other vegetative capacities or functions are entirely describable from a third person perspective just like any other physical processes (ibid.).

I tend to disagree with Gasser. Aristotle talks about consciousness – even if he never uses the term, as explained above – when he explains conscious activities such as perception, desire and thinking; without these faculties, one is not conscious of anything. We talk of consciousness because of those faculties. When one perceives the form of a perceptible object, for instance, the colour of a blue car, one becomes aware of the colour of the car and knows that it is blue, not green. The experience of the blue colour of the car becomes one's subjective experience. It is one's conscious activity. And when one talks about one's experience of the blue colour of the car, one will be talking about her experience, which is from the first person perspective.

Descartes may discuss the metaphysical analysis of consciousness and the first perspective aspect of the mental experience, nonetheless, he does not solve the problem of interaction between the body and the mind. However, Aristotle's notion of consciousness does not create any divide between consciousness and body, as it rests on his hylomorphic analysis whereby form as the soul represents anything mental at any level of reality, and matter the physical. In other words, with Aristotle's hylomorphism, consciousness as depicted as the form can be discussed and analyzed at any metaphysical level one wishes as every substance even the smallest one is made of matter and form. Again this reminds us of Chalmers' type f' monism.

10. Conclusion

I have argued that Aristotle's hylomorphism and theory of causality provide a possible solution to the mind-body problem because it gives a good account of mental causation. His philosophy of mind could be read as the middle link between dualism and physicalism, and a means of unifying the two theories through his matter and form theory (hylomorphism). Because as matter and form unite to form a substance, physics and philosophy together will help us to have a comprehensive knowledge of the universe. More so, it addresses the shortcomings of dualism and physicalism.

I have shown that through the activities of perception, thought and desire, Aristotle demonstrated that mental causation is possible; mental phenomena are not epiphenomenal but causally efficacious as the soul, the principle of life, is the efficient cause and final cause of any human being. In his theory of mind, there is no principle of supervenience and there is no case of over-determination as hylomorphism does not consider the physical (matter) a more fundamental element than the mental (form), nor does it consider the physical causally efficacious while the mental is non-causally efficacious. The mental does not supervene on the physical and it does not compete for causal power or causal relevance with the physical. Rather, in the hylomorphic theory, the soul is the efficient cause that initiates and completes any movement or change. For instance, desire is a mental faculty that initiates human action and it does it together with the practical reason and imagination.

And Aristotle, as non-reductivist and soft functionalist, is able to account for phenomenal consciousness as self-consciousness that enables him to give a theory of action that is human oriented. Aristotle saves mental causation from the criticism of Jaegwon Kim with his theory of hylomorphism and notion of causality, and he does it more elegantly than Davidson as he does not need a theory of supervenience – please see the final chapter for a more thorough discussion.

Chapter 5: Aristotle vs. Kim and Davidson

1. Introduction

In chapter one, I defined the playing ground for my thesis by defining some key terms such as mental realism, which is the view that mental properties have independent causal power contrary to Kim's notion that the causal power of the mental resides in their physical properties. I indicated that the causal efficacy of mental phenomena under investigation is that of phenomenal consciousness. For me, phenomenal consciousness can, *qua* mental, cause both physical and mental events. In chapter two, I showed that one of the main problems of the physicalistic theory is the inability to account for mental causation, as it has no place for phenomenal consciousness. Moreover, it was argued that even functionalism does not account for mental causation.

In chapter three, I examined the notion of mental causation in the context of Donald Davidson and Jaegwon Kim's discussions. Davidson is a role functionalist while Kim may be categorized as a realizer functionalist. Davidson and Kim as functionalists in their respective ways failed to account for phenomenal consciousness in human action, even if Davidson makes a compelling argument for human freedom via his anomalous monism. They deny human agency in their theories. I hold that any true and adequate solution to the mind-body problem should be able to account for the reality of phenomenal consciousness as causal *qua* mental that carries the agent along. I showed too that Donald Davidson does not rescue mental causation and non-reductivism because his theory of anomalous monism through supervenience is physically oriented, hence there is after all no independent mental causation. Kim argues that if it is in virtue of physical properties that the mental properties are causally efficacious, then the mental properties are rendered epiphenomenal. The mental has no independent causal power for him.

I argued in chapter four that it is Aristotle's hylomorphism and theory of causes that can save mental causation and any non-reductivist theory of mind from Kim's criticism of non-reductive attempts, primarily in this thesis, of Davidson's anomalous monism, to argue in

favour of the causal efficacy of the mental. I based my argument on the fact that Aristotle holds that the soul has causal power *qua* an immaterial form. I argued too that Aristotle's theory of hylomorphism plus his theory of causes accord an agent a place in their own actions. This is because Aristotle's metaphysics can account for phenomenal consciousness, which I argued being the classical notion of self-consciousness. It is the self-consciousness of the agent that causes them into action.

In this chapter, I will argue that Aristotle provides a better theory of mental causation than both Kim and Davidson through his theory of mind that is based on his theory of hylomorphism and theory of causes.

I will carry out my work in this chapter by comparing Aristotle's account of mental causation and Kim's account with the aim of showing that Aristotle offers a better theory of mental causation than Kim, and then I will compare Aristotle's account of mental causation and Davidson's to buttress my position.

2. Aristotle vs. Kim

Aristotle fares better than Kim in terms of explaining the problem of mental causation, as Aristotle's hylomorphism and theory of causality allow him to explain mental events such as perception and thought and desire as being causally efficacious. As we saw earlier, perception occurs when the bodily organ receives the form of the perceptible object, while thought occurs when the passive intellect receives the intelligible form.

Aristotle's theory of causality makes room for different kinds of causation for different things in nature – with this I mean both living and noon-living things. His theory of hylomorphism entails matter and form; the material cause and formal cause, potentiality and actuality. The four causes can be applied both to inanimate and animate things. While the efficient cause of inanimate objects is external to the object, in the case of living organism, it falls together with the formal and final causes, and is inherent to the living organism.

According to Aristotle's account of mental causation, it is desire, together with practical reason and imagination embodied in self-consciousness that causes human action (see section 4.6.2). And practical reason is inherently concerned with deliberation and conduct with regard to what is good and bad. Consequently, human beings as conscious subjects are rational beings with human freedom for moral actions and responsibility (see sections 4.8, 4.9).

Kim's view of mental causation is physically based as he argued that mental phenomena (desire, belief and intention) can cause mental and physical events only if they are reduced to the physical. (See chapter three.). His kind of causation does not differentiate between the causation for human action and causation for non-human action. It could be said that he has one notion of physical causation for both inanimate and animate things. The implication is that inanimate objects, such as stones, will have moral obligations and responsibility like human beings. If mental phenomena have no independent causal power but are dependent on physical properties, it does imply that 'bits of matter' are responsible for the law and order we human beings have in our society. However, Aristotle, with his theory of form and matter makes room for different kinds of causation; there is physical causation for physical objects and mental causation for humans.

Matter or micro particles in the body is part of who we are but they are not what define us. It is the soul in our case, the rationality or the intellect that defines us as human beings. Aristotle does not undermine the matter or physical world or mechanistic world, rather he is saying though matter plays an important role in the world, matter alone cannot give a comprehensive understanding of our world. We need the form too to have a better understanding our world, and in the case of humans, the form is the soul. Kim only focuses on the physical world, so he cannot explain the something more (phenomenal consciousness), because he rejects the possibility of the causal efficacy of the mental *qua* mental in the physical world. It is interesting that he acknowledges that subjective experience cannot be functionalized, but hopes that future science will show us a way in which to do so in time. Aristotle, on the other hand, can explain subjective experience as self-consciousness because he has a workable theory for mental causation.

Kim's concept of mental causation argues that mental properties have to be reduced to physical ones for them to have causal power over the mental and physical. For him everything is physical and everything is governed and explained by physical laws. Hence, he advocates the causal closure of the physical world. However, Kim seems to acknowledge the reality of the mental phenomena when he argues that all mental phenomena, such as intentions and cognitive properties, can be functionalized – except qualia. That qualia resists functionalization, suggests at least the possibility for the reality of the mental, and with it the possibility of having a kind of effect in the physical and mental world although Kim's current approach cannot account for it.

However, Aristotle's account of mental causation holds that through the soul as the inner principle of cause of change or action, we may justify or explain philosophically through argument what human beings may hold intuitively with regard to the causal efficacy of the mental. The mental is not exclusively attached or attributed only to either an invisible entity called the soul or the physical body, but rather mental states encapsulated in the soul, are found inherent in the human person composed of living flesh and blood. This explanation is based on Aristotle's theory of hylomorphism. Every substance is composed of matter and form, body being the matter and form being the soul in the human person. Through the hylomorphic theory, he avoided the problem faced by the physicalists and functionalists like Kim and Davidson in accounting for phenomenal consciousness. I have argued thus, based on an argument for showing that the classical notion and Aristotle's notion of self-consciousness is the same as phenomenal consciousness, and an argument that both the passive and active intellect play a role in creating self-consciousness.(See chapter 4.). The physicalists and functionalists, on the other hand, are yet to account for the sense of freedom and choice human beings think they have as they cannot adequately account for phenomenal consciousness.

Kim is aware of the limitation of his functional reduction for he says, "the best, or the most satisfying, outcome would have been the vindication of mental causation along the lines of non-reductive physicalism; that would have allowed us to retain mentality as something that

is causally efficacious and yet autonomous *vis-à-vis* the physical domain" (Kim, 2005:159). This outcome may be actualized in Aristotle's philosophy of mind through his theory of hylomorphism and causality. As I have shown in Aristotle's philosophy of mind, there is no reduction or supervenience of the mental on the physical needed, as rather we have a unity of matter and form or mind and body. And the soul is the efficient cause, formal cause and final cause that causes human beings to perceive, think and desire things and cause humans into action.

I have shown why I think that Aristotle's account of mental causation saves the causal efficacy of the mental *qua* mental, while Kim's account cannot do that. In order to strengthen my argument that it is Aristotle's hylomorphism that can save non-reductive physicalism from Kim's criticism, I will show in the next section why I think Aristotle's approach to mental causation is also better than Donald Davidson's. Davidson has argued for mental causation but I have found his argument not adequate to save mental causation against Kim's criticism (see chapter 3). Hence, I argue that it is perhaps only Aristotle who can save non-reductive physicalism from Kim. Thus, I will argue in the next section that Aristotle's metaphysics offers a better response to Kim's criticism of non-reductive physicalism with regard to mental causation than Davidson's account does.

3. Aristotle vs. Davidson

To argue for mental causation, Davidson argues that mental events are physical events and their interactions fall under physical laws, which means monism. However, due to the rationality and coherence of mental properties, mental properties do not fall under physical laws. Hence, mental types or properties are anomalous. This is his theory of anomalous monism. Anomalous monism implies a token identity relationship between mental events and physical events; mental events are physical events but mental types are not physical types (see section 2.7). Anomalous monism is anchored on four principles. It will be good to mention the principles again. The first three principles are:

(1) The principle of causal interaction: some mental events causally interact with physical events.

(2) The principle of the nomological character of causality; where there is causality, there must be a law. Events that are related as cause and effects fall under deterministic law.

(3) The anomalous nature of the mental: there are no strict deterministic laws with which mental events can be predicted or explained (there are no strict laws that govern mental phenomena). This implies the irreducibility of the mental to the physical phenomena and the autonomy of the mental.

The third principle appears to conflict with the first two as it denies their positions. Davidson reconciles the principles by arguing that each mental event has both a mental and a physical description and in that sense each mental event is a physical event. Thus one event for Davidson can be described or characterized both in mental terms and physical terms. (See section 2.7.) Davidson, in order to explain better the relationship between mental events and physical events, Davidson brought in the fourth principle, which is the principle of supervenience. The principle holds that each mental event supervenes or depends on a physical event (see section 2.7).

Davidson's main aim in his anomalous monism is to show that mental properties are autonomous and there is mental causation in the physical world. That is, he wants to show that our reason can cause us to perform some actions.

Davidson's anomalous monism arguments implies that:

> If we are (or can be) rational and coherent it follows that we should believe that we are (or can be) free to respond to requirements of logic and evidence. This seems to capture a deep and fundamental intuition about who we are and the nature of mind: rationality implies autonomy, and autonomy implies freedom to respond to the normative principles of logic

> and evidence, objective moral standard, and so forth … Davidson represents this freedom to respond in terms of anomalousness … Davidson's core motivation [is] to provide a realistic and workable account of autonomy (Scharf, 2010: 344).

However, Kim argues that anomalous monism and Davidson's supervenience render mental properties epiphenomenal. Kim's main argument is that mental events should be causally efficacious based on mental properties not on physical properties but anomalous monism does not show that but rather only that mental causation is dependent on physical properties. Kim, as a physicalist, believes that everything in the universe is physical or reducible to the physical. Mental entities cannot have causal influence in the physical world (see sections 3.5; 3.6).

Davidson, in his theory of anomalous monism, acknowledges the reality and autonomy of the mental. The autonomy of the mental due to the principle of rationality and coherence is argued not to be governed by the deterministic laws found in physics. But, I have argued (see section 2.8.1) that Davidson fails to adequately account for mental causation, as his argument for mental causal efficacy is still established on physicalism due to his principle of supervenience. Both Kim and Davidson accord primacy to physical properties in mental causation, even if they come to different conclusions concerning the causal power of the mental.

Aristotle and Davidson have the same aim in their different theories of mental causation for both believe in the causal efficacy of the mental. However, Aristotle's argument for mental causation is not physicalist oriented like Davidson. Davidson with his two principles of anomalous monism and supervenience accord primacy to physical properties just like Kim, but in Davidson's case he argues for mental properties having causal power due to their subvenient base that is physical in nature and also due to him being a kind of 'predicate dualist' in the sense that he argues there two ways in which to describe the mental. The latter view, while making a strong argument for non-reductive physicalism, did not convince Kim. I find it

compelling but not as well situated or effective as Aristotle's metaphysics to finally rescue non-reductivism from Kim's criticism.

However, Aristotle argues that mental events are causally efficacious in virtue of their mental properties through his hylomorphism and causal theory. Thus, Aristotle, a non-reductivist, can save non-reductive physicalism from Kim's criticism. For Aristotle, every substance is a composite of matter and form. Matter is an indeterminate principle that becomes individuated once it receives an appropriate form. Matter is in potentiality to receive a form (actuality). In applying hylomorphism to the mind-body problem, soul is to body what form is to matter or what actuality is to potentiality. And in explaining change or motion in substance, Aristotle applies his theory of causality, which entails four causes, namely the material cause, formal cause, efficient cause and final cause. He holds that the form is the formal cause, final cause and efficient cause in the case of living organisms. Advertently, in humans what we have, are two causes namely the material cause and formal cause. In natural beings like humans, the soul is the form and thus is the efficient, formal and final cause. It is the soul that actualizes the body. And the soul is irreducible to the body as form is irreducible to matter. He defines the soul as the first actuality of the living body as explained above (see section 4.2).

The soul causes every organism that is a composite of body and soul to be alive; a living thing by performing the necessary action suitable for his environment. It is the principle of life. That a plant is alive and functioning is because of its nutritive soul that helps it to produce, feed and grow. A dog is alive and functioning because of its sensitive soul that incorporates the nutritive soul, and this sensitive (perceptive) soul helps it to navigate through its surroundings looking for food for survival and running away from dangers that threaten its survival. Human beings have the other two souls (nutritive and sensitive souls) as well as the rational soul. This rational soul causes a human being to make moral choices and decisions in life. The soul and body of the organism form a unity in the organism, they are inseparable; one cannot exist without the other. Thus, there is no reducibility of mental to physical or physical to mental, but rather their relationship is explained in terms of potentiality and actuality. And the

supervenience that helps Davidson in his theory of anomalous monism is not required. There is no primacy of one property over the other; body and soul need each other to form an individual substance**.**

Hylomorphism and the theory of the four causes charter a specified theory of action for human beings, which is quite different from that of non-living things. The theory of action is based on human nature that is composed of body and soul, thus mental causation in Aristotle's philosophy of mind is quite different from Hume's concept of causation that conditions the principles Davidson uses in his mental causation argument (Kim too, as pointed in sections 3.5; 3.6)). The supervenience theory gives us a complete physicalistic explanation of human action. In terms of hylomorphism however, the soul is the efficient, formal and final cause in plants, animals and humans. It is an inherent cause in the living being; it is the internal principle of change or remaining unchanged. Human beings are beings whose acts are controlled by thought and choices, and this is due to the fact that they possess a soul as the principle of life. Thus, it is the mental phenomenon of desire that causes humans to act (see section 4.2).

While Kim and Davidson lay emphasis on the role of mental states (desire, belief and intention) in causing human action, Aristotle's emphasis is on the role of self-consciousness that entail the whole human person as a physical, emotional, intelligent and conscious being (see chapter 4). To this effect, Aristotle's theory of human agency is bound to be different from that Davidson and Kim. For Aristotle, self-consciousness is the cause of human action. This implies that the agent is completely involved in his action, and that the goal he set out for himself causes his action in order realize that goal. His action is for him and about him in relation to the society in which he lives. The function of man, according Aristotle's ethics is to be virtuous in order to live a happy life. It is the phenomenon of self-consciousness that causes man to be virtuous.

It is on this grounds that I argue that Aristotle is a soft functionalist, not a role functionalist like Davidson nor a realizer functionalist like Kim. He is what I call a 'soft functionalist' as he believes in the function of the soul, whose causal power does not depend anyhow on the matter, while the form functionalises matter in the sense of organising it,

though the soul as a form cannot be without the matter. For Aristotle, the soul or mind or intellect is immaterial but causally efficacious in the physical world. Since the soul is immaterial, it is not subject to the Humean idea of causation that conditions Kim and Davidson's causal theories.

4. Conclusion

Aristotle's hylomorphism and theory of causality provides a possible solution to the mind-body problem because it gives a good account of mental causation. Through the activities of perception, thought and desire, Aristotle demonstrated that mental causation is possible; mental phenomena are not epiphenomenal but causally efficacious as the soul, the principle of life, is the efficient cause and final cause of any human being. And Aristotle's concept of mental causation is better than both Kim and Davidson's accounts as both are unable to account for phenomenal consciousness; the mental phenomenon that causes human into action. But Aristotle is able to account for phenomenal consciousness as self-consciousness. And this self-consciousness is responsible for every human action that is intelligent, reasonable and conscious.

Mel Thompson (2012: 6) writes:

> I have a feeling that much time and effort in the philosophy of mind could have been saved had his [Aristotle's] principles been attended to more carefully… [that] the soul or mind is distinct from particular material parts of the body, that it is holistic, and that it is inseparable from the body – has implications for much modern discussion about materialism and functionalism … What Aristotle is surely saying … is that the way you describe a soul or the essence of an animal – human or otherwise – depends on, but is not reducible to, what one might say about parts of that animal body, including (most importantly) its brain … If he were alive today Aristotle would have been be quick to point out that

> detailed descriptions of neural activity are simply descriptions of parts of the body and cannot show its essence.

It will not be arrogant after the arguments put forward in this chapter to say that Aristotle's hylomorphism and notion of four causes have a lot to contribute positively to finding a contemporary solution to the mind-body problem because it saves mental causation from Kim's criticisms that denies causal efficacy and autonomy to the mental properties. In terms of Aristotle's hylomorphism, the mental is autonomous and causally efficacious in its own terms.

Conclusion

The 'mind-body problem' is a metaphysical problem that has plagued the minds of many philosophers, scientists and neuroscientists. It points to fundamental questions concerning the nature, dignity and worth of a human person.

Human beings are beings with both physical and mental properties. The mind, with its mental properties such as self-awareness, and consciousness, is located in the brain, which is a physical organ in the body. The mind is thus embodied. On the other hand, mental properties of the mind, such as subjective experience and intentionality, seem to suggest that the mind is a non-spatial, immaterial substance. This characteristic of the mind, together with its embodiedness, creates a tension, which underlies all mind-body debates. The mind-body problem thus relates to the relationship between the immaterial mind and the physical body. But, how does what is mental interact with physical substance? The mind-body problem is the problem of mental causation (the problem of interaction) and the problem of the nature of the mind (its metaphysics). In this thesis, I have focused more on the former than the latter problem – i.e. how can something strictly mental have an effect on a material or physical substance.

The mind-body problem debate is about the nature of a human person; it is a debate about what it means to be a rational and conscious human person. Thus, any theory of mind should be able to account for rationality and consciousness in a human person. The solution to the mind-body problem will guarantee the autonomy of the mental.

I argued that any true and adequate solution to the mind-body problem should be able to account for the reality of the mind and body, and their causal interaction. I argued that most of the solutions offered from physicalist perspectives to solve the mind-body problem are inadequate to do that. I specifically argued against Jaegwon Kim's philosophy of mind in order to establish my thesis that Aristotle's theory of hylomorphism may have the solution to the mind-body problem and the problem of mental causation.

To reach my conclusions, I first mapped out the playing ground for my thesis by defining the key terms such as mental reality, immateriality of the mind, mental causation and consciousness. I stated that mental reality connotes the causal efficacy of mental phenomena in the physical world, and these causally efficacious mental phenomena are immaterial in nature, but not separated from the physical in the Cartesian dualist sense. In other words, they are immaterial but inseparably intertwined with the physical. I also stated that these immaterial mental phenomena make up phenomenal consciousness. And I stated that the mental causation in question is Kim's concept of mental causation that challenges and questions the causal efficacy of mental properties in the presence of the causal efficacy of physical properties. Kim mounted this challenge through his causal closure and causal exclusion principles.

I also examined the functionalist notions of mental causation as both Kim and Davidson – and Aristotle to some extent – are functionalists. However, while Kim is a realizer functionalist, Davidson is a role functionalist. I argued that neither physicalism, which comes in two forms namely reductive physicalism and non-reductive physicalism, nor functionalism, in any of their versions, are adequate to account for mental causation, as phenomenal consciousness is beyond their explanation. Hence, the theories of human action suggested by these views are not human oriented, but based on the functions of mental states (desire, belief and intentions) occurring in the brain.

For me to argue for the cardinal point of my thesis, which is that Aristotle's hylomorphism and notion of four causes can offer a seemingly adequate solution for mental causation, I had to argue first against Kim's criticism of Davidson's anomalous monism. Kim, through his causal exclusion (supervenient) and causal closure arguments, established that

Davidson's anomalous monism cannot solve the mind-body problem as for Davidson the causal power of mental properties is supervenient or dependent on physical properties.

I demonstrated how physicalism occupies a very important place in both Davidson's anomalous monism argument and Kim's supervenient and causal closure argument. Both Davidson and Kim base their arguments on physicalism. Kim is a supporter of the theory of reductive physicalism, which is anchored on physicalism (that everything is physical or reducible to physical). And he argues strongly against non-reductive physicalism (that some facts such as mental facts are irreducible to physics). Though Davidson's aim (a non-reductivist one) is to prove mental causation and the autonomy of the mental in the physical world, he still takes physical facts to be the base of all other scientific facts. Davidson argues for anomalous monism because he believes in the truth of determinism and in the irreducibility of the mental. Hence, he uses the principle of supervenience to reconcile mental facts and physical facts in his anomalous monism.

The need to look for a solution to the mind-body problem outside the physicalistic causally closed universe is mentioned by Kim (2005) in his book *Physicalism or Something Near Enough*. There Kim argues that the physicalistic view of the mind is not adequate at present. Kim admits too that it is hard to functionalise qualia. I believe that if the universe were indeed physically closed as Kim argued, a physicalistic view of qualia or functionalized qualia would be possible, but as it is not physically closed, then we cannot functionalize qualia.

I saw this as a call to look for a solution to the mind-body problem, and one way that became apparent was to distinguish causation in terms of human agency from physical causation that is Humean oriented. This is related to the call made by Davidson, and other such as Yablo and Hutto, for us to rethink mental causation in terms of human agency. I showed that this call or invitation to rethink causation along human agency was in line with Aristotle's theory of causes as enshrined in his theory of hylomorphism. As Aristotle is able to account for phenomenal consciousness as self-knowledge which is regarded in modern terms as self-consciousness, he succeeds in accounting for a theory of action that is human based whereby the agent is involved in his action as the owner of his action, not as an outsider.

I argued that Aristotle's hylomorphism and theory of causality provide a possible solution to the mind-body problem because it gives an adequate mental realist account of mental causation. Aristotle's philosophy of mind could be read as the middle link between dualism and physicalism, and a means of unifying the two theories through his theory of hylomorphism. Because as matter and form unite to form a substance, a physical analysis and a philosophical analysis together will help us to have a comprehensive knowledge of any substance. And the two areas of inquiry may be of great assistance to neuroscience in their research on the correlation between the mind and the brain. More so, it addresses the shortcomings of both dualism and physicalism.

I argued that Aristotle, through his philosophical analysis of the activities of perception, thought and desire, demonstrated that mental causation is possible; mental phenomena are not epiphenomenal but causally efficacious as the soul, the principle of life, is the efficient, formal and final cause of living organisms. I argued that Aristotle's concept of mental causation is more compelling and adequate than both Kim and Davidson's accounts as the latter are built on an understanding of classical physics that is not 'metaphysical' enough. Aristotle's hylomorphism gives us the needed metaphysical ground as it states that every substance is a composite of matter and form; potentiality and actuality.

Thus, I argued that Aristotle's hylomorphism and notion of four causes have a lot to contribute positively to finding a contemporary solution to the mind-body problem because it saves mental causation from Kim's criticisms that mental properties are epiphenomenal. In terms of Aristotle's hylomorphism, the mental is autonomous and causally efficacious in its own terms. And, there is no problem of mind-body interaction as matter and form cannot exist without each other.

By appealing to Aristotle's hylomorphic theory, I have shown that reductive, exclusively neuro-science-based solutions to the mind-body problem cannot adequately show or explain what it means to be a rational and conscious human being. I have shown too that Aristotle's biological functionalism, which I have explained as being a 'soft functionalism', and which is based on his hylomorphic metaphysics, has serious potential to change contemporary thinking

about the mind-body problem, especially with regard to mental causation, and I have done this in this work in particular in relation to his psychology, or philosophy of mind.

My work invites contemporary philosophers to look back beyond Descartes (back into ancient times) to find a possible solution to the mind-body problem. The uniqueness of my research lies in the fact that I have used Aristotle's psychology to challenge the philosophy of mind of a well renowned contemporary philosopher, Jaegwon Kim, and thereby seems to succeed in doing what Donald Davidson could not fully succeed in accomplishing with his anomalous monism.

Future research could focus on determining if there would have been a modern mind-body problem if Aristotle's hylomorphism and notion of causes had been accepted because his hylomorphism professes unity of matter and form. If the answer is yes, what form would the mind-body problem have taken?

Bibliography

Ackril, J.L. (1981). *Aristotle the Philosopher*. Oxford: University Press.

Ainsworth, T. (2016). "Form versus Matter". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL =https://plato Stanford.edu. /archives/spr. 2016/entries/form/matter.

Anscombe, G.E.M. (1975). "Causality and Determination", in Sosa Ernest (Ed.), Causation and Conditionals. London: Oxford University Press: Pp.63-82.

Antony, L. (2007). "Everybody has got it: A defense of non-reductive materialism" in B.P. McLaughlin and J. Cohen (Eds.) *Contemporary Debates in Philosophy of Mind*. United Kingdom: Blackwell Publishing Ltd pp.143-159.

Anthony, M. V. (2003). "Davidson's Argument for Monism", *Synthese 135*, pp.1-12.

Aristotle. (1952). "Physics". *Great Books of the Western World*. Edited by R.M. Hutchins and M.J. Adler. Volume 8. Aristotle 1.Chicago: Encyclopedia Britannica Inc.pp.259-359

---------- (1952). "Metaphysics". *Great Books of the Western World*. Edited by R.M. Hutchins and M.J. Adler. Volume 8. Aristotle 1. Chicago: Encyclopedia Britannica Inc.pp.490-630

---------- (1952). "On The Soul". *Great Books of the Western World*. Edited by R.M. Hutchins and M.J. Adler. Volume 8. Aristotle 1. Chicago: Encyclopedia Britannica Inc.pp.631-670

---------- (1952). "On Sleep and Sleeplessness". *Great Books of the Western World*. Edited by R.M. Hutchins and M.J. Adler. Volume 8. Aristotle 1. Chicago: Encyclopedia Britannica Inc.pp.696-701

Armstrong, D. M. (1968) "A materialist theory of mind".

---------- (2000). "The Nature of Mind" in Brian Cooney (ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Arnon, V. (2017). "On Hilary Putnam's "The Nature of Mental States" http://sweetprince.net/words/essayson-hilary-putnam.

Atkins, J. (2014). "Psycho functionalism; the Varieties of the Functionalist Philosophy of Mind https:/Platoscinema.wordpress.com/tag/psycho functionalism.

Bayne, T. (2010). *The Unity of Consciousness*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199215386.001.0001

Bayne, T and Chalmers, D. (2003). "What is the Unity of Consciousness?" A. Cleremans (Ed.). *The Unity of Consciousness: Binding, Integration, Dissociation,* Oxford.

Bennett, K. (2007). "Mental Causation". *The Philosophy of Compass* 2/2, pp.316-337. 10.1111/j.1747-9991.2007.00063.x Journal Compilation: Blackwell Publishing Ltd

Block, N. (1990). "Inverted Earth". *Philosophical Perspectives*, Vol.4: pp.53-79.

---------- (1995). "Functionalism". *A Companion to the Philosophy of Mind*. S. Guttenplan (Ed.). (1995). Cambridge, pp.323-333.

---------- (1995). "On a Confusion about the Function of Consciousness" in *Behavioral and Brain Sciences*, 18, pp.227-47.

---------- (forthcoming) "Functional Reduction" in a Festschrift for Jaegwon Kim edited by David Sosa.

Bodnar, I. (2016). "Aristotle's Natural Philosophy". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL=https://plato.stanford.edu/archives/win2016/entrie/aristotlenatphil

Brentano, F. (1991). "The Distinction between Mental and Physical Phenomena". *Metaphysics*. Vide Hoy &Oaklander.

Brook, A. and Raymont, P. (2017), "The Unity of Consciousness". *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/consciousness-unity/>.

Brugger, E. C. (2008). "Aquinas on the Immateriality of Intellect. A Non-Materialist Reply to Materialist Objections". The National Catholic Bioethics Centre.

Cahn, S.M. (1977). *Classics of Western Philosophy, 2nd Edition*. Indianapolis: Hacket Publishing Company.

Calef, S. (2017). "Dualism and Mind". *The Internet Encyclopedia*, ISSN 2161-002,https://www.iep.utn.edu

Carruthers, P. (2016). "Higher-Order Theories of Consciousness" *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL=https://plato.stanford.edu/archives/ fall2016/entries/consciousness –higher/

Caston, V. (2002). "Aristotle on Consciousness", Mind, vol.114. USA: Oxford University Press

-------- (2006). "Aristotle's Psychology" in M. Louise Gill (et.al) (Eds.) *Blackwell Companion to Philosophy. A Companion to Ancient Philosophy*. USA: Blackwell Publishing Comapany

Chalmers, D. (1995). "Facing up to the Problem of Consciousness". *Journal of Consciousness Studies*.2. (3): pp.200 – 219.

---------- (1996). *The Conscious Mind, In Search of a Fundamental Theory*, New York: Oxford University Press.

---------- (Ed.). (2002). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

---------- (2003). "What is the Unity of Consciousness" in A. Cleremans (Ed.). *The Unity of Consciousness: Binding, Integration, Dissociation,* Oxford.

---------- (2010). *Character of Consciousness*. USA: Oxford University Press.

Child, W. (1992). "Anomalism, Uncodifiability and Psychophysical Relations". *The Philosophical Review,* 102(2): pp. 215 – 245.

Churchland, P. (1989). *A Neuro-computational Perspective*. Cambridge: MIT Press.

---------- (1991). "Reduction, qualia and the direct introspection of the brain", in *Metaphysics*. Vide Hoy and Oaklander.

Churchland, P.S. (1997). "The Horns woggle Problem" in Shear J. (Ed.). *Explaining Consciousness: The Hard Problem*. MIT Press.

Churchland, P. (2000), "Eliminative Materialism" in Brian Cooney (Ed.). *The Place of Mind.* USA: Wadswork Thomas Learning.

Coetzee, P.H and Ruttkamp-Bloem, E. (1995). *Metaphysical systems and problems*. Pretoria: University of South Africa.

Cohen, S. M. (1987). "The Credibility of Aristotle's Philosophy of Mind" in Mohan Matthen (Ed.). *Aristotle Today*. Academic Printing and Publishing pp. 103 – 121.

---------- (1992). "Hylomorphism and Functionalism" in Martha Nussbaum and Amelie Rorty (Eds.). *Essays on Aristotle's De Anima*. Clarendon Press pp.57-73.

Cole, D (2015). "The Chinese Room Argument". *The Stanford Encyclopaedia of Philosophy*, (winter 2015 Edition), Edward N. Zalta (Ed.). URL=https:/plato.stanford.edu/archives/win2015/entries/Chinese room.

Collingwood, R.G. (2001). *The Idea of Nature.* USA*:* Oxford University Press.

Cooney, B (2000). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Craig, E. (2005). *The Shorter Routledge Encyclopedia of Philosophy*, 2nd Edition. UK: Routledge.

Dale, J. (2009). *The Philosophy of Mind: The Metaphysics of Consciousness*. New York: Continuum International Publishing Group.

Davidson, D. (1970). "Mental Events", in Foster and Swanson (Eds.). *Experience and Theory*. London: Duckworth.

---------- (1980). *Essays on Actions and Events*, Oxford: Clarendon Press

---------- (1987). "Problems in the Explanation of Action" in *Metaphysics and Morality*. Oxford: Blackwell.34-49.

---------- (1993). "Thinking Causes" in Heil and Mele (Eds.). *Mental Causation*. Oxford: Oxford University Press pp. 3-17.

-------- (2001). *Essays on Actions and Events*, Oxford: Clarendon Press.

--------- (2002). "Mental Events" in David. J. Chalmers (Ed.). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

-------- (2003). "Thinking Causes", In Neil Campbell (Ed.). *Mental Causation and the Metaphysics of Mind*. 3- 134-148.

Dennett, D. C. (2000). "Intentional Systems" in Brian Cooney (Ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Demeter, T. (2009). "Two Kinds of Mental Realism" (JGen Philos,sci. 40: 59-71. Doi 10.1007/s108370099090-4

Descartes, R. (1968). *Discourse on Method and the Meditations*. Great Britain: Richard Clay Ltd.

-------- (1974). "*Ouevres de Descartes"* in Charles Adam and Paul Tannery (Eds.). 11 vols. Paris: Vrin.

---------- (1984). *The Philosophical Writings of Descartes*, 3 vols., trans. John Cottingham, Robert Stoothoff, Dugald Murdoch and Anthony Kenny. Cambridge: Cambridge University Press.

Dilley, F.B. (2004). "Taking Consciousness Seriously: A defense of Cartesian dualism". *International Journal for Philosophy of Religion 55: 135-153.*

Dretske, F. (1989). "Reasons and Causes". *Philosophical Perspectives, 3* (1989), 1-15.

Eardley, P.S and Still, C.N. (2010). *Aquinas, A Guide for the Perplexed*. India: Replika Press Pvt Ltd.

Echavariria, R. R. (2010). *Realism in Mind.* London: University of Canterbury

Falcon, A. (2019). "Aristotle on Causality". *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (Ed.). URL =https: //plato.Stanford.edu. /archives/spr. 2019/entries/Aristotle- Causality.

Fodor, J. A. (1989). "Makin Mind Matter More". *Philosophical Topics* ,17(1989), 59-80).

Foorrest, P (2016), "The Identity of Indiscernibles". *The Stanford Encyclopaedia of Philosophy,* (Winter 2016 Edition), Edward N. Zalta (Ed.). URL = https://plato.stanford.edu/archives/win2016/entries/identity-indiscernible/

Fitz, N. and Gumm, D. (2010). "Anomalism Monism", *Aporia*. 20. no2.

Fodor, (1981). "The Mind- Body Problem", Scientific American, Inc. 1981

Gaarder, J. (1996). *Sophie's World: A Novel about the History of Philosophy*. New York: Orion Books Ltd.

Gasser, G. (2009). "Is Hylomorphism a Neglected Option in Philosophy of Mind?" https://uibk.ac.at//philtheol/gasser/publ//

Gennaro, R.J. (2017). "Consciousness". *The Internet Encyclopedia*, ISSN 2161-002,https://www.iep.utn.edu

Gerson, L. (2005). *Aristotle and other Platonists*. Ithaca, N.Y.: Cornell University Press

Gertler, B. (2008). "In Defense of Mind-Body Dualism" in J. Feinberg and R. Shater-Landau (Eds.). *In Reasoning and Responsibility. Readings in Some Basic Problems of Philosophy*. California: Thomson wads worth.

Gibb, S. (2015). 'Defending Dualism". *Proceedings of the Aristotelian Society*, Vol. CXV, part 2.doi: 10.1111/j.1467-9264.2015.00388.X.

Gilson, E. (2009). *From Aristotle to Darwin and Back Again: A Journey in Final Causality, Species and Evolution.* USA: Ignatius Press.

Goff, P. (2017). *Consciousness and Fundamental Reality*. New York: Oxford University Press

Gray, J.W. (2011). "Ethical Realism"

Hammond, W. A. (1902). *Aristotle's Psychology: A Treatise on the Principle of Life.* UK: Swan Sonnenschein and Company

Hollbach, B. (1770). '*System of Nature'.*

Honderich, T. (1982). "The Argument for Anomalous Monism" *in Analysis.* 42. 59-64.

---------- (1983). "Anomalous Monism: Reply to Smith" *in Analysis*. 43. 147-49.

---------- (1984). "Smith the Champion of Mauve" *in Analysis*. 44. 86-89.

---------- (1995). *The Oxford Companion to Philosophy*, USA: Oxford University Press.

Hsieh, D.M. (2003). "Mental Causation through Constitution". diana@dianahsieh.com

---------- (2003). "Desire, Reason, and Action". diana@dianahsieh.com

Huemer, W. (2019). "Franz Brentano", *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/spr2019/entries/brentanO/.

Hutto, D. (1999). "A Cause for Concern: Reasons, Causes, and Explanations". *Philosophy and Phenomenological Research*, Vol. LIX. No 2. June 1999.

Huxley, T. (2002), "On the Hypothesis that Animals are Automata and its History" in David J. Chalmers (Ed.). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Jacquette, D. (2009). *The Philosophy of Mind: The Metaphysics of Consciousness.* London: Continuum International Publishing Company.

Jackson, F. (1982). "Epiphenomenal Qualia". *Philosophical Quarterly*, 32,127, pp.127-136.

Jacob, P. (2002). "Some Problems for Reductive Physicalism". *Philosophical and Phenomenological Research*, Vol. LXV, No.3

Kant, I. (1964).*Groundwork of the Metaphysic of Morals*. Translated by Paton H.J. New York: Harper and Row. Originally published by Kant in 1785.

Kim, J. (1984). "Epiphenomenal and Supervenient causation." *Midwest Studies in Philosophy* 9(1984), 267.

---------- (1989). "The Myth of Non-reductive Materialism." *Proceedings of the American Philosophical Associations,* 63 (1989), 39-41).

---------- (1993). "Can Supervenience and 'Non-Strict Laws' Save Anomalous Monism?", in Heil and Mele (Eds.), 19-26.

---------- (1993). *Supervenience and Mind*, Cambridge: Cambridge UP.

---------- (1999) "Making Sense of Emergence", Philosophical Studies 95, 3-36.

--------- (2001a). *Mind in a Physical World: An Essay on the Mind- Body Problem and Mental Causation*. Cambridge: The MIT Press.

--------- (2001b). "Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalists" in Carl Gillet, & Barry M. Loewer (Eds.). *Physicalism and its Discontents*. Cambridge University Press pp. 271-283.

---------- (2002). "Responses". *Philosophy and Phenomenological Research*, Vol. LXV, No.3.

------- (2003). "Concepts of Supervenience" in Neil Campbell (Ed.). *Mental Causation and the Metaphysics of Mind*. (2003), pp.217 - 239.

---------- (2005). *Physicalism or Something Near Enough.* Oxford: Princeton University Press

---------- (2007). "Causation and Mental Causation" in McLaughlin and Cohen (Eds.), 2007, 227-243.

---------- (2007). "The Causal Efficacy of Consciousness" in Velmans, Max and Susan Schneider (Eds.). *The Blackwell Companion to Consciousness*. Wiley

--------- (2012). *Essays in the Metaphysics of Mind*. Oxford: University Press

Kind, A. (2017). "Qualia". *The Internet Encyclopaedia of Philosophy*, ISSN 2161. http://www.iep.utm.edu

---------- (Ed.) (2018). Philosophy of Mind in the Twentieth and Twenty-First Centuries

The History of the Philosophy of Mind, Volume 6. London: Routledge

Kripke, S.A. (2002), "Naming and Necessity" in David. J. Chalmers (Ed.). *Philosophy of Mind: Classical and Contemporary Readings.* New York: Oxford University Press.

Kuusela, A. (2010). *Non-reductive Physicalism, Irreducibility of the Mental and the Problem of Mental Causation: A Study of Donald Davidson's and George Henrik Von Wright's Positions in the Philosophy of Mind.* ISBN 978-952-10-6608-5 (PDF). Helsinki.

Lawson-Tancred, H. (1986). *Aristotle De Anima (On the Soul).* England: Clays Ltd, St. Ives plc.

Levine, J. (1983) "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly,* 64, 354-361.

---------- (2017). "Functionalism". *The Stanford Encyclopaedia of Philosophy* (winter 2017 Edition), Edward N. Zalta (Ed.). URL=https:/plato.stanford.edu/archives/win2017/entries/functionalism.

Lewis, D. (1983). "Mad pain and Martian Pain' and 'Postscript", *Philosophical Papers* Vol.1, Oxford: Oxford University Press

Loar, B. (2000). "Phenomenal States" in Brian Cooney (Ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Lowe, E.J. (1993). "The Causal Autonomy of the Mental", *Mind*, Vol.102: 408.

Macdonald, C. and G. (1986). "Mental Causes and Explanation of Action". *Philosophical Quarterly*, 36(143).pp. 145 – 158.

Matthews, E. (2005). *Mind: Key Concepts in Philosophy*. Great Britain: MPG Books Ltd.

Maslin, K. (2007). *Introduction to the Philosophy of Mind.* Cambridge: Polity Press.

McGinn, C. (1982). *The Character of mind*. USA: Oxford University Press

---------- (1989). "Can We Solve the Mind-Body Problem?" *Mind* 98: 349-66.

McLaughlin, B. (1993). "On Davidson's Response to the Charge of Epiphenomenalism" in Heil and Mele (Eds.). *Mental Causation*. Oxford: Oxford University Press pp. 27-40.

Melchert, N. (1999). *The Great Conversation: A Historical Introduction to Philosophy*, *3rd Ed*. London: Mayfield Publishing Company.

Melden, A. I. (1961). *Free Action.* London: Routledge and Kegan Paul.

Merricks, T. (1995). "Critical Notice of Jaegwon Kim's Supervenience and Mind", Philosophical *Books*. 36.158-161.

Miller, A. (2019). "Realism", *The Stanford Encyclopedia of Philosophy* (Winter. 2019 Edition), Edward N. Zalta (ed.), URL== https://plato.stanford.edu/archives/win2019/entries/

Miller, F. D. (1999). Aristotle's Philosophy of Soul. The Free Library. 1999 Philosophy Education Society, Inc. 7. Oct.2019. https://www.thefreelibrary.com//aristotle%27S+philosophy+of+soul.-a064426384

Nagel, T. (1974). "What it is like to be bat?" *Philosophical Reviews,* Vol.83: 435-50. doi: 10.2307/2183914. Jstor 2183914.

-------- (2002). "What Is It Like to Be a Bat" in David J. Chalmers (Ed.). *Philosophy of Mind; Classical and Contemporary Readings.* New York: Oxford University Press.

Nannini, S. (1999). "Physicalism and the Anomalism of the Mental". In de Caro M. (Eds.), Interpretation and Causes. Syntheses Library (Studies in Epistemology, Logic, Methodology and philosophy of Science.) vol.285. Springer Dordrecht. Doi-https://dor.org/10.1007/978-94-05-92277-1

Nath, S. (2013). "Ryle as a Critique of Descartes" Mind-Body Dualism". *International Journal of Scientific and Research Publication,* Volume 3.Issues 7.

---------- (2014) "Type-Token Dichotomy in the Identity Theory of Mind"

*Journal of Business Management and Social Sciences Research***,** Vol. 3, No. 4,

O' Connor, T and Wong, H.Y. (2015). "Emergent Properties", *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL== https://plato.stanford.edu/archives/sum2015/entries/

Offray De La Meettrie, J. (1748). *Man a Machine* (*L'Homme aachine)*

Omoregbe, J. (1996). *Metaphysics without Tears: A Systematic and Historical Study.* Lagos: JERP Press.

---------- (2001). *Philosophy of Mind: An Introduction to Philosophical Psychology*. Lagos: JERP Press.

Pasnau, R. (2012). “Mind and Hylomorphism”. *The Oxford Handbook of Medieval Philosophy.*

Place, U.T. (2000). “Is Consciousness a Brain Process” in Brian Cooney (Ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Priest, S. (1991). *Theories of the Mind*. England: Clays Ltd

Putnam, H. (2002). “The Nature of Mental States” in David. J. Chalmers (Ed.). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.

Quine, W.V.O. (1995). “Naturalism; Or Living within One’s Means”. *Dialectica, 49*. 251-261

Radner, D. (1985). “Is There a Problem of Cartesian Interaction?” *Journal of the History of Philosophy 23*

Rainone, A. (1999). “Thirty five Years after “Actions, Reasons’ Causes”, What Has Become of Davidson’s Causal Theory of Action”. In de Caro M. (Eds.), Interpretation and Causes. Syntheses Library (Studies in Epistemology, Logic, Methodology and philosophy of Science.) vol.285. Springer Dordrecht. Doi-https://dor.org/10.1007/978-94-05-92277-1

Robb, D and Heil, J. (2018). “Mental Causation”. The Stanford Encyclopaedia of Philosophy (Summer 2019 Edition), Edward N. Zalta (Ed.) URL=< https://plato.stanford.edu/archives/sum2019/entries/mental-causation

Robinson, H. (2017). "Dualism". *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2017/entries/dualism/>.

Rockwell, T. (2004). “Physicalism, Non-reductive, Dictionary of Philosophy of Mind”. https://sites.google.com//site/minddict/physicalism-non-reductive

Rosenthal, D. M. (1998). “Two Concepts of Consciousness.” In *Philosophical Studies* 49: 329-59.

Rost, J. M. (1989). *The Mind of Aristotle: A Study in Philosophical Growth*. Canada: University of Toronto.

Russell, B. (1927). *The Analysis of Matter*. London: Kegan Paul

Ryle, G. (1949). *The Concept of Mind*. Hutchinson’s University Library.

Sanguineti, J. J. (2013). “The Ontological Account of Self-Consciousness in Aristotle and Aquinas”. The Review of Metaphysics, vol.67. p.311-344

Scharf, D. (2010). “Why Jaegwon Kim’s physicalism is not Near Enough, An Implicit Argument for a non-Cartesian Interactionism”. NeuroQuantology, 3: 337-353.

Schlosser, M. (2019) "Agency", *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/agency/>.

Searle, J.R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: University Press.

---------- (2000). “Do Minds Compute” in Brian Cooney (Ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Sihvola, J. (2007). “The Problem of Consciousness in Aristotle’s Psychology, in (Sara Heinamaa (et.al.) (Eds.), *Consciousness: From Perception to Reflection in the History of Philosophy.* Dordrecht: Springer.

Skirry, J. (2008). *Descartes: A Guide for the Perplexed*. New York: Continuum International Publishing Group.

Shields, C. (2016). “Aristotle’s Psychology”. *The Stanford Encyclopedia of philosophy (*Winter 2016b Edition), Edward N. Zalta (Ed.). URL =https://plato.Stanford.edu./archives/win 2016/entries/Aristotle Psychology.

---------- (2016). “Aristotle”. *The Stanford Encyclopedia of philosophy* (Winter 2016b Edition), Edward N. Zalta (Ed.). URL =https://plato Stanford.edu./archives/win 2016/entries/Aristotle.

Smart, J.J.C. (2000). “Sensations and Brain Processes” in Brian Cooney (Ed.). *The Place of Mind*. USA: Wadsworth, Thomas Learning.

Sosa, E. (1984). “Mind-Body Interaction and Supervenient Causation” *in Midwest Studies in Philosophy,* IX, 271-81.

---------- (1993). “Davidson’s Thinking Causes” in Heil and Mele (Eds.). *Mental Causation*. Oxford

Sorabji, R. (1974). “Body and Soul in Aristotle”. *Philosophy*, 49 (1974), 63-89.

---------- (1990). *Aristotle Transformed.* Ithaca, N. Y.: Cornell University Press.

Sreekumar, N. (2015). “Aristotle’s theory of Causation and the Ideas of Potentiality and Actuality, Aspect of Western Philosophy, IIT Madras.

Stoljar, D. (2016). “Physicalism”. *The Stanford Encyclopaedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (Ed.). <https:plato.stanford.edu//archives//win2016.

Strawson, P. F. (1959). *Individuals*. London: Methuen.

---------- (1985). “Causation and Explanation” in Vermazen and Hintikka (Eds.). *Essays on Davidson: Actions and Events*. Oxford: Clarendon Press. 115-35.

---------- 1992: *Analysis and Metaphysics*. Oxford: Oxford University Press.

Stumpf, S.E. (1966). *Socrates to Sartre: A History of Philosophy*. USA: McGraw- Hill.

Sturm, T. (2012). “Consciousness regained? Philosophical arguments for and against reductive physicalism”. < http//creativecommons.org/license/by-nc-nd/3.0/.

Tanney, J. (2009). “Rethinking Ryle. A Critical Discussion of the Concept of Mind” in G. Ryle (Ed.). *The Concept of Mind 60th Anniversary Edition*. London: Rout ledge.

Thompson, M. (2012). *Understand Philosophy of Mind: Explore Who You Are*. Great Britain: CPI Cox and Wymen Reading.

Turing, A. (1950). “Computing Machinery and Intelligence”. *Mind*, 50: 433-460.

Tye, M. (1995). Ten Problems of Consciousness. Cambridge, MA: MIT Press.

Van Gulick, R. (2017). "Consciousness", *The Stanford Encyclopedia of Philosophy (*Winter 2017 Edition*)*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/consciousness/>.

Van Riel, R. and Van Gulick, R. (2019), "Scientific Reduction", *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2019/entries/scientific-reduction/>.

Vella, J.A. (2008). *Aristotle: A Guide for the Perplexed*. Great Britain: Continuum.

Velleman, J.D. (1992). “What Happens When Someone Acts.” *Mind*, voi.101. p.401.

Voss, S. (1993). “Simplicity and the Seat of the Soul” in Stephen Voss (Ed.). *Essays on the Philosophy and Science of Rene –Descartes.* New York: Oxford University Press.

Walter, S. (2007). “Determinables, Determinates and Causal Relevance”. *Canadian Journal of Philosophy* Vol.37, Number 2, June 2007, pp.217-244.

Wetzel, L. (2018). "Types and Tokens", *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/types-tokens/>.

Yablo, S. (2002). “Mental Causation” in David. J. Chalmers (Ed.). *Philosophy of Mind: Classical and Contemporary Readings.* Oxford: Oxford University Press.

Yoo, J. (2018). “Mental Causation”. *The Internet Encyclopedia*, ISSN 2161-002,https://www.iep.utn.edu