Richard Swinburne's arguments for substance dualism.

MA by Research in Theology and Religion

David Horner

September 2018

Richard Swinburne's arguments for substance dualism.

Submitted by David Horner to the University of Exeter

as a dissertation for the degree of

MA by Research in Theology and Religion

in September 2018

Acknowledgements.

I would like to thank my supervisors, Dr Jonathan Hill and Dr Joel Krueger for their support and encouragement in the writing of this dissertation and for their patience in trying to keep me on the straight and narrow. I want to acknowledge the many conversations, on this and other topics, I have had with my friend and philosopher, Dr Chris Boyne, who sadly died in June of this year. I thank all my other chums at The Bull, Ditchling, for listening to my metaphysical ramblings. And finally, I thank my wife, Linda, for once more putting up with this kind of thing.

Abstract

This dissertation is a contribution to debates in the philosophy of mind and of personal identity. It presents a critical account of arguments for substance dualism to be found in Richard Swinburne's *Mind, Brain, and Free Will* (2013). Swinburne's principal claim is that persons are essentially pure mental substances whose sameness over time is constituted by a unique 'thisness'. A human being consists of two parts: a contingent part, the body (physical substance), and an essential part, the mind or soul (pure mental substance) which is characterised by 'thisness'. It is, on this account, logically and metaphysically possible that a person can be disembodied. The dissertation analyses Swinburne's relationship to other major theories in the philosophy of mind, especially his critical rejection of physicalism and materialism. Swinburne mounts a defence of substance dualism by building upon some key fundamental ideas and principles. The first area of discussion is Swinburne's novel contention that any satisfactory account of the mental and physical lives of human beings must meet the requirements of a 'metacriterion' which supports his division of the world into physical and mental substances, properties, events and time. Swinburne underpins the metacriterion by proposing a canonical vocabulary based on a theory of informative designators. The main line of attack here is on the inadequacy of Swinburne's theory of designation as a convincing theory of how language works and is used. Secondly, the metacriterion is complemented by a theory of privileged access of subjects to their mental events which is not available to others. Criticism of this doctrine is derived from the work of Austin, Ryle and Wittgenstein. Thirdly, Swinburne deploys the principles of credulity and testimony to defend the causal interaction of mental and physical substances. He claims the principles are fundamental, *a priori*, and epistemic. The argument of the dissertation is that they are none of these things. My conclusion is that Swinburne's principal arguments for substance dualism and personal identity are unsound.

Contents

List of Tables.

Abbreviations.

CCP	Causal Closure of the Physical
DP	Direct Perception
DRT	Direct Reference Theory
EA	Epistemic Assumption
FC	Fundamental Criterion
PAP	Principle of Alternative Possibilities.
PC	Principle of Credulity
PLA	Private Language Argument
PT	Principle of Testimony
RBU	Relations-Between-Universals
SPL	Substances-Powers-Liabilities
TE	Thought Experiment
UP	Unobservability Principle

Chapter One: Introduction

1 Rationale.

Richard Swinburne's *Mind, Brain, and Free Will* (2013) is a defence of substance dualism and free will. However, in this dissertation I concentrate primarily on his defence of substance dualism. Swinburne reasserts the boldly Cartesian claim that human beings are fundamentally dual consisting of a soul (the 'essential part') and a body (the 'inessential part'). Persons are souls interacting with bodies. It is the soul (or mind) that constitutes the 'thisness' of a human being and is the guarantor of personal identity, human free will and moral agency. He makes these claims against what he takes to be a prevailing philosophical orthodoxy that conceives persons as 'merely complicated machines' (Swinburne, 2013, p.1). In arguing for 'substance dualism' he believes that he is arguing in support of a belief held by '...most ordinary people, at least in Western countries...that humans consist of two parts – a soul and a body'. His intent is to provide a set of concepts and arguments that show that '...the traditional majority Western view... is the correct one' (2013, p.3). This is a strong and important claim and one which requires careful examination.

The case he sets out builds upon previous work especially *The Evolution of the Soul* (1997). However, he maintains that the case presented in *Mind, Brain, and Free Will* is 'deeper' and 'stronger' and provides different and novel arguments with a fuller discussion of the underlying philosophical arguments about mind and body. If Swinburne's conclusion is indeed correct, that 'humans are pure mental substances' (2013, p.163), this must have profound implications for our everyday understanding of what it is to be a person and for related philosophical and theological debates about souls, minds, and bodies including the possibility of 'survival' after death. For example, the possibility of our continued, personal existence in a disembodied form must be a necessary condition for some such survival.

In an early review Holdsworth (2014, p.635) suggests at least four reasons why *Mind, Brain, and Free Will* deserves to be a focus of interest. First, he suggests that the work embodies a 'counter cultural anthropology' in the way that it opposes contemporary forms of materialism and physicalism and combats current trends which fail to distinguish adequately physics and metaphysics. Secondly, it addresses some key problems '…about starting-points and navigation-points in philosophy'. Swinburne lays out a rich ontological and epistemological landscape as the groundwork for his arguments for substance dualism. Thirdly, and perhaps most importantly, he suggests that there may be important practical implications if Swinburne's arguments for dualism are correct. For example, his arguments suggest that, whilst neurological research is valuable, it must have limited prospects. On Swinburne's account mental properties do not supervene upon physical properties nor are deducible from them therefore purely neurological investigations of 'mental life' must always remain limited. There will remain, in principle, a gap, however sophisticated our 'brain imaging' techniques may become, between such imaging and what we are intentionally picturing and experiencing. 'Asking the subject' must remain a pre-eminent fact in experimental method. Fourthly, Holdsworth suggests that Swinburne's 'philosophical anthropology' implies a renewed emphasis on conflict and 'inner tensions' in the human psyche:

> Without the potential conflict of contemporaneous desires and beliefs in the same individual, the individual is a wanton (*sic*). Only given the prospect of such conflict does the question of what to believe about the relative worth of competing desires have import. (Holdsworth, 2014, p.641)

2 Aims and Objectives.

The aim of the dissertation is to analyse and evaluate Swinburne's arguments and establish whether the conclusions he draws are both true and valid. Does he really make a convincing case for the view that human beings are essentially non-physical? The principal aim is to analyse and evaluate the case Swinburne presents for 'substance dualism' in Chapters One to Six of *Mind, Brain, and*

Free Will (2013). This may be ambitious enough without extending the discussion into, albeit important, implications of dualism for free will and moral agency (Chapters Seven and Eight respectively).

The objectives of the study include:

(a) To define core concepts of Swinburne's ontology such as 'substance', 'event', 'property', and epistemology (the credulity and testimony principles);

(b) To situate Swinburne's contribution within contemporary literature on the mind/body problem especially his critique of physicalism;

(c) To analyze the trajectory of Swinburne's thinking about substance dualism especially the introduction of a metacriterion and the theory of informative [rigid] designators;

(d) To evaluate the cogency of the privileged access argument in supporting the dualism of mental and physical substances;

(e) To evaluate Swinburne's use of the credulity and testimony principles, for mind/body interaction and in critiques of competing non-interactionist theories;

(f) To analyze the role of thought experiments in understanding personal identity from a dualist perspective;

(g) To evaluate the arguments for the logical possibility and metaphysical possibility of disembodiment and the mind or soul as pure mental substance.

3 An Overview of *Mind, Brain, and Free Will*.

In Chapter One of *Mind, Brain, and Free Will* (2013), 'Ontology', Swinburne lays down metaphysical commitments which provide the groundwork for subsequent arguments. His claim is that '...everything turns on questions about the identity of properties and substances' (Swinburne, 2013, p.39). Fundamental to this project is establishing a 'metacriterion' which will enable us to possess a

(canonical) vocabulary to equip us to ‘tell the whole history of the world’ and to stipulate identity criteria for the basic constituents of the world.

First, Swinburne sets about this in Aristotelian fashion in terms of substances, properties and events. Substances are ‘things’ as diverse as electrons and souls. Substances can consist of other substances. Substances are what they are because of their properties (characteristics of substances) and events (temporal descriptions of substances and properties). Properties can be essential or inessential to the identity of a substance in which they are instantiated. He claims that if substances, properties and events are included in a description of the world, or slices of the world, then everything can be accounted for. The failure of other theories of mind/body relationships is that they fail to meet this metacriterion notably failing to identify the distinction between physical and mental substances, properties and events,

Secondly, events are the instantiation of a property in a substance at some time. A complete history, on this account, may be told by a description of events. Swinburne outlines the principles of a canonical vocabulary adequate for the specification of identity criteria for substances, properties and events. To establish such criteria Swinburne adapts Kripke’s (1980) theory of designation. A ‘rigid designator’ always refers to the same object in every possible world in which it exists. Swinburne includes the condition that for a rigid designator (which he re-names as an ‘informative designator’) we must know the conditions of its application. If we know this then we cannot be wrong!

Swinburne distinguishes between informative [rigid] designators and uninformative [rigid] designators. The former is a term which picks out properties of a substance believed to be essential to it. The latter picks out the accidental properties or predicates that may only contingently describe a substance. The theory specifies logical identity conditions of substances, properties and events. We can then distinguish essential properties of objects from superficial ones. For example, physical properties picked out by informative designators include ‘being square’; ‘having a positive electric charge’, and so forth. Mental properties (of perception) picked out by informative designators include ‘seeing a desk’, ‘hearing a telephone ring’ and so on. The world consists of substances and events; a complete history is a

history of events. Only such an account of identity criteria, Swinburne claims, can provide a complete history of the world although this does not limit the many ways the world may be divided up (Swinburne, 2013, p.9).

What appears to follow:

(a) Having identity criteria for substances, properties, and events means that we can evaluate claims of identity, for example, claims by materialists or physicalists that mental events are just physical events or realized in physical events;
(b) And whether I am the same thing (i.e. substance) as my body. Swinburne believes we can make the move from the logical possibility of the mind existing without a body (i.e. that this is 'conceivable) to its metaphysical possibility (i.e. that it is not only conceivable but actual). This argument pivots on the claim that there is a non-physical element which constitutes 'me' which is identifiable by the informative [rigid] designator 'I'.

Chapter Two, 'Epistemology', sets out the epistemological commitments which show events that occurred or occur according to the metacriterion.

First, two fundamental, *a priori*, epistemic principles underpin what it is to have a justified belief about what is or is not the case. These are:

(a) 'The Principle of Credulity'. What seems to be the case probably is the case in the absence of counter evidence ('special considerations'). For example, Swinburne claims that we are justified, on this principle, of thinking that 'mental events cause brain events' because that is just how it seems to us.
(b) 'The Principle of Testimony'. If what people tell us, i.e. testify to, is ultimately caused by what is apparently perceived then we are entitled to believe it is the case. Such testimony about an event arises from our 'privileged access' to our experiences.

Swinburne's belief is that 'some' philosophers of mind simply don't establish such key epistemic principles and yet they are crucial, for example, to

arguments about whether intentions cause or do not cause brain events – from a 'scientific' (*sic)* point of view. These principles may be 'defeated' if there is evidence (a 'defeater') that what is perceived or testified to is not the result of what the subject perceived or testified to.

Secondly, in addition to these principles Swinburne asserts that there is an allied epistemological assumption (EA) that lies behind the holding of a justified belief in a scientific theory. The EA consists of three elements: (a) a justified belief that a theory's predictions are true; (b) that the only sources of evidence for the occurrence of a predicted event are: experience, memory, and testimony; (c) justification is subverted by evidence contradicting any of these sources. He claims that what goes for justified belief in scientific theories is equally appropriate for any theory concerning 'metaphysically contingent matters' (Swinburne, 2013, p.64). These principles play a role in Swinburne's criticisms of findings in neuroscience which seem to undermine free will.

In Chapter Three, 'Property and Event Dualism', Swinburne develops the idea that there are two kinds of events and properties in the world: mental and physical. Mental events and physical events have different identity conditions and are ontologically and logically distinct. Two moves seem to be in play here.

First, the contrast is between our unique, direct, privileged access to our own mental properties and events and the inferential access to our mental properties and events by others. Swinburne defines a mental property as '...one to whose instantiation in it a substance necessarily has privileged access on all occasions of its instantiation' (Swinburne, 2013, p.67). These are 'pure mental events'. Pure mental events encompass propositional events: beliefs, thoughts, intentions, desires, and non-propositional events: sensations.

Secondly, the identity conditions for such pure mental substances, properties and events are derived from the appropriate use of their informative [rigid] designators: '...two properties are different iff their informative designators are not logically equivalent' (Swinburne, 2013, p.68). It follows from the appropriate use of informative designators that mental and physical properties, for example, cannot be logically equivalent. Nothing about a brain's physical constitution seems to entail conscious experience. This is a key argument against the identity theory of a mind/body relationship.

In Chapter Four, ‘Interactive Dualism’, having established the fundamental, ontological distinction between physical and mental substances and so on, Swinburne must address the classical Cartesian problem of how it is possible that immaterial substances can have causal efficacy with regard to physical substances. His positive case is that we just do generally experience our mental events (say, having an intention) leading to physical actions. He refers to this as ‘the normal view’. His negative case is the refutation of counter-evidence that tend to contradict the normal view. He invokes the epistemological principles of credulity and testimony, set out in Chapter Two, to refute such counter-evidence.

Our general understanding of large scale physical principles seems to provide the first obstacle to the acceptance of interactionism. Fundamental postulates of physics, such as the conservation of energy, leave no space for interactive, substance dualism. For example, the transfer of energy underlying causation is simply not compatible with interactions between physical and non-physical substances. The principle of the conservation of energy implies energy cannot simply arise out of nothing or the non-physical. This is the causal closure of the physical (CCP). The interactive dualist account implies that for mental substances (minds) to influence physical substances (brains) then transfers of energy must be involved. Given the non-physical and non-extended nature of mental substances any such energy must seem to appear *ex nihilo*. We have no evidence that this is the case. Causation just is wholly physical. However, Swinburne responds by suggesting that it is not even clear that CCP holds for physical matter (citing Quantum Theory).

A second obstacle arises from some contemporary neurological and psychological evidence. Libet’s psychological experiments, for example, seem to show that the brain engages in activity prior to the reporting of intentions to act. This appears to undermine claims that mental states cause brain states and that it is the brain which determines our actions whether we are aware of this or not. Wegner cites a range of phenomena that seem to show that we are frequently mistaken or deluded about the extent to which our conscious mental states and actions are the results of free will.

Swinburne finds, on grounds of the credulity and the testimony principles, that neither of these two lines of argument are compelling. The CCP argument just can't overthrow what is entailed by the normal view. There is an important paradox that to gather the required evidence based on subjects' testimony and experiences the very theories which are being postulated must be false. The assumptions of the counter-theories are self-refuting. The normal view must be correct and 'very probably' our intentions often do cause brain events.

In Chapter Five, 'Agent Causation', Swinburne argues for a model of causation that is both favourable to the dualist position and looks forward to his later arguments for free will. Building on the arguments for interactive dualism he makes the case for a distinction between non-intentional causation and intentional or agent causation. However, both types of causation are subsumed under one umbrella in the sense that all causation, he claims, is causation by substances rather than events. This prepares the ground for agent causation as free will.

First, he challenges the view (seemingly held by 'some' philosophers and scientists) that the causes of events are (always?) other events. He then introduces two general accounts of causation: (a) The relations-between-universals (RBU) interpretation in which laws of nature express relationships between universals grounded in necessary regularities from the succession of events; and (b) the substance-powers-liabilities (SPL) interpretation accepts fundamental laws as regularities, not regularities of events, but '...regularities in the causal powers and liabilities of substances' (Swinburne, 2013, p.129). What follows is that powers of substances are not powers to produce effects but '...to exercise a certain kind and amount of causal influence'. Swinburne illustrates this by a discussion of gravitational influences of planets according to Newton's laws.

Secondly, he draws a distinction between non-intentional causation and intentional causation. Non-intentional events are embedded in chains of causation in which one event follows another (logically distinct) event by some law of nature. For example, the ignition of dynamite causes an explosion, the ignition event being caused by an electrical spark and so on. On the SPL

interpretation this should be reconceptualised as powers and liabilities of substances rather than as events.

Thirdly, what follows is that intentional ‘events’ are not really events at all. Agent causation (as distinct from non-intentional causation) means that: ‘...A person having an intention (acting) is simply the person exercising causal power’ (Swinburne, 2013, p.2). And a person is defined, Swinburne goes on to argue (in Chapter Six, ‘Substance Dualism’), as a pure mental substance which intentionally, exercises causal powers. This discussion is key to his further understanding of the nature of free will.

Substances with specific capacities (powers) describe the world as it is – rather than (abstract) laws of nature. All causation is causation by substances rather than events. This characterisation is again essential to meeting the metacriterion requirement to be able to tell a complete history of the world. Non-intentional and intentional causation are thereby rendered under a ‘simpler’ account which explains why effects can be understood in terms of different forms of causation: intentional and non-intentional causes. For example, I throw a rock which crashing through a window is accounted for within the same SPL paradigm of causation and the dual substance theory.

Swinburne, in Chapter Six, ‘Substance Dualism’, argues that personal identity rests ultimately on our being ‘a pure mental substance’ (a mind or soul). It is our ‘thisness’, that makes someone the same person over time but also the same person whether embodied or disembodied. He reaches this conclusion by arguments for the synchronic and diachronic unity of personal consciousness.

First, the synchronic unity of consciousness is constituted by the unity of co-experience of a mental substance. He argues, on grounds of metaphysical assumptions and neurological evidence, that it is not possible to have a ‘full description of the world’ in which substances are individuated only by physical properties:

> For it is an evident datum of human experience that conscious mental events of different kinds (visual sensations, auditory sensations, etc) are co-experienced, that is belong to the same (mental) substance, both at one time and overtime. (Swinburne, 2013, pp.142-143)

Secondly, Swinburne draws a distinction between a ‘complex view’ and a ‘simple view’ of personal identity. The complex view is the belief that sameness and continuity over time can be grounded in psycho-physical characteristics: memory, character, physicality (brain matter) and so on. He advocates the simple view that, each person is a ‘particular’ characterised by ‘thisness’ which is not dependent on, or reducible to, specific physical or mental properties. This is the diachronic unity of the person.

Swinburne believes that the logical possibility of the simple view is supported by the failure of the complex view. This failure is illustrated by the outcomes of various thought experiments, for example, considering the consequences of ‘split brain’ transplants. He contrasts the relative weakness of the complex view in interpreting the results of such experiments with the relative strength of the simple account. He deploys his basic ontological and epistemological claims to arrive at these conclusions. He believes he establishes the logical possibility of pure mental substances existing without a body but not *vice versa*. This is like the Cartesian argument for the conceivability of the self as a thinking substance independent of the body.

Thirdly, Swinburne’s further move is to argue that this logical possibility of the existence of pure mental substances can be converted into their metaphysical possibility. Such a possibility means that such states of affairs can occur given certain circumstances. For Swinburne the circumstance in this case is that ‘I’ is an informative designator. I can name, indubitably, a non-physical element which (always) is me; I cannot be mistaken in identifying myself as a self, an experiencing subject. It follows that I can exist without other psycho-physical characteristics such as memory or continuity of brain matter. Human beings are pure mental substances, having souls as their essential parts and bodies as inessential parts. Physical properties belong to human beings in virtue of belonging to their bodies; ‘pure mental properties’ belong to them by belonging to their souls; souls are non-physical, pure mental substances. It further follows that we might very well continue without our current body or indeed any body although Swinburne acknowledges that this is not generally our normal state.

4 My Approach.

First, my approach in this dissertation is primarily one of interrogation and analysis of Swinburne's core concepts, principles and arguments. It is based on a close reading of the first six chapters of *Mind, Brain, and Free Will* (2013) which culminate in the justification of substance dualism. It is this justification with which we will be ultimately concerned. I try to follow the arguments where they lead. I identify and analyse Swinburne's conceptual and argumentative strategies in criticizing opposing physicalist theories. He importantly attempts to demonstrate the soundness of his case by showing that alternative theories are simply defective. This is exemplified in his defence of 'interactive dualism'. I discuss the extent to which such strategies are successful.

Secondly, much of this discussion is contentious; a fact that Swinburne himself acknowledges. The text is complex and densely argued; some critics (McGuire and Hause, 2014, p.670), for example, have described the style as '…often painfully detailed'. Blackford (2014, p.112) is also critical of Swinburne's style for 'lack of examples' and having '…no lighter moments'! I explore the logical geography of his terminology especially given his tendency to stipulative definition. Holdsworth (2014, p.642) notes that much of Swinburne's argumentative work is achieved by 'linguistic stipulations' justified by the claim that '…any philosopher is entitled to define technical terms as they wish' (Swinburne, 2013, p,71). And so they are. For example, according to Swinburne, minds and bodies are defined as 'logical substances', pure mental substances in the former case, and physical substances in the latter. But stipulative definitions have risks since they usually run against the grain of publicly accepted, standard usage whether technical or otherwise. We have every reason to question them for that reason. Stipulative definitions need reasonable and valid grounds for acceptance. Swinburne's definitions are, therefore, an important locus of scrutiny and investigation in what follows. This is especially necessary in Swinburne's assertion of the need for a metacriterion requiring that '…our vocabulary should enable us to tell the whole history of the world' (Swinburne, 2013, p10).

Thirdly, I have structured the dissertation around three core elements that I believe are keystones for his justification of substance dualism. The first element is a metacriterion (my Chapter Three). This sets out a fundamental requirement which any proposed account of mind/body relationships must meet. The requirement is the need for a vocabulary which will encapsulate identity criteria for mental and physical substances, properties, and events. Only in this way will we be able to tell 'a complete history of the world' which will fully represent the '…the data of the mental and physical life of human beings' (Swinburne, 2013, p.10). This 'canonical vocabulary', he argues, is based on a specific theory of designation which we examine and critique at some length and forms perhaps the biggest departure from his previous work.

The second element is privileged access (my Chapter Four). This is a familiar theme in Cartesian discussions of mind/body relations. It is the importance of the special relationship we each have to the contents of our own consciousness. We 'know' our own thoughts and feelings directly in a way other people cannot. We can only infer the contents of others' consciousness. I invoke some standard criticisms of these claims because they are, for Swinburne, part of the argument for supporting the view that the world consists of two kinds of events: physical and mental. Not only are there two kinds of events but mental events may persist without physical events.

The third element is 'credulity' which can also be bracketed with the related principle of 'testimony' (my Chapter Five). These two principles are, he claims, fundamental, *a priori* and epistemic. Having established the key distinction between mental and physical these principles play a central role in Swinburne's moves to overcome the classical Cartesian problem of just how incorporeal mental substances, such as minds, can interact with all too corporeal physical substances, such as bodies. The principles are deployed to undermine various types of counter-evidence which appear to threaten interactive dualism. Swinburne argues the paradox, using the credulity and testimony principles, that to collect the evidence which is supposed to support deterministic physicalist theories those theories must themselves be false. Evidence from subjects can only provide evidential grounds if the testimony is the result of undetermined judgements. However, I go on to show why, and how, the two key

principles invoked by Swinburne are neither fundamental nor *a priori* nor epistemic.

Nevertheless, Swinburne believes that the arguments advanced for substance dualism remain compelling. In my Chapter Six, 'Personal Identity', I examine his belief that mental substances (minds) can exist in a disembodied state. Further that our potential to exist in such a state does not rest on the continuity of any mental properties, memory, character and so forth. The metaphysical possibility of disembodiment rests on the existence of pure mental substances (souls or minds) which can be picked out by the informative designator 'I'. I unpack the modal argument he deploys here and show that the assumption that 'I' can be treated as an informative [rigid] designator is simply false. What Swinburne does in the crucial last phase of his argument for substance dualism is to attempt, invalidly, to derive existence from verbal or conceptual definitions. Not only is the idea of the metaphysical possibility of disembodiment untenable but its logical possibility seems contradictory.

In the next chapter I set the scene by placing Swinburne's substance dualism in the wider context of contemporary approaches to the mind/body problem. I draw attention to those theories to which Swinburne considers and critiques.

Chapter Two: The Mind/Body Problem.

1 Introduction.

In his advocacy of substance dualism Richard Swinburne presents his solution to the longstanding mind/body problem. In doing so he rejects all physicalist accounts. He writes that '…subtly different forms of physicalism all hold the doctrine that the only events are physical events, some of which may also in a sense (although not my sense) be 'mental events'' (Swinburne, 2013, p.93). Fundamental to dualism is the representation of the mind/body problem as concerned with mental to material correlations (Heil, 2013, p.16). This naturally arises from a core assumption that mental states and events are, and (logically) must be, distinct from physical states and events. And this schism between the mental and the material has posed the question of the nature of the correlations between minds and brains. What are the operative causal mechanisms? Historically, there have been a range of responses to these questions within dualism itself. These include epiphenomenalism, parallelism (Leibniz's pre-established harmony), and occasionalism (relying on the intervention of God) (Heil, 2013, p.68). Swinburne's brand of dualism, as we noted in the previous chapter, is a form of interactive dualism which gives causal efficacy to mental and material substances. This chapter situates Richard Swinburne's refutation of non-dualist theories within the landscape of contemporary trends in the philosophy of mind. In what follows I use David J. Chalmers' taxonomy (2002, pp.247-272) as a convenient map of the relevant landscape.

1.1 A Taxonomy of Metaphysical Views of Consciousness.

Chalmers' taxonomy of major trends in metaphysics of consciousness consists of six major broad types of theory (see Table One below). These six classes are themselves broadly divided into two major groups. First, there are reductive,

materialist theories claiming that consciousness is fundamentally a physical process which require no expansion of (physical) ontology. These are: Type-A theories such as Eliminativism; Type-B Identity theories; and Type-C theory such as Double-Aspect theory. Secondly, there are non-reductive theories of consciousness which claim that consciousness is something irreducible in nature and requires the enlargement of (physical) ontology. Chalmers' taxonomy is by no means definitive but gives us a broad perspective in which to place Swinburne's defence of substance dualism (Type-D theories).

Swinburne's prime objects of criticism are those theories which broadly adhere to some variety of materialism or physicalism that is Type-A, Type-B, and Type-C theories in Chalmers' terminology. (I follow Chalmers in treating materialism and physicalism as synonymous terms). This is part of Swinburne's general strategy to establish substance dualism by the critical destruction of alternatives and demonstrating 'the errors of physicalism' (2013, p.93). Swinburne takes it as axiomatic that it is obvious that properties and states of mind cannot logically be reduced to properties of brains or any physical (material) entity. Physicalism or materialism, on the contrary, maintain that consciousness (phenomenal truths) may ultimately be explained within the framework of physical principles.

Given the range and extent of the critical literature in philosophy of mind I have mainly confined myself to those theories and theorists specifically cited by Swinburne in his text. The chapter, therefore, concentrates on the varieties of physicalism criticized in *Mind, Brain, and Free Will* (2013). These include Identity Theory, Functionalism, Instrumentalism and Eliminative Materialism (Types A – C). Section Four moves briefly from Swinburne's critique of physicalism to Type-D Dualism and Chalmers' theory of Property Dualism. Here Swinburne presents a more sympathetic response recognizing a kinship with his own view of the irreducibility of mental events to physical events. In contrast he rejects epiphenomenalism (Type-E Dualism). However, we delay a fuller consideration of Swinburne's treatment of epiphenomenalism to my Chapter Five, 'Credulity and Testimony'.

Table One: Chalmers' Taxonomy of Metaphysical Views of Consciousness (Chalmers, 2002, pp.247-272).

Type	Characterization	Description	Examples
A	Physicalism	No 'hard problem' problem of consciousness. No ontological gap between physical and phenomenal truths.	Eliminativism Analytic Functionalism Logical Behaviourism
B	Physicalism	The gap is epistemological not ontological. Phenomenal states are identified with certain physical or functional states.	Identity Theory
C	Physicalism	'Deep' epistemological gap between the phenomenal and the physical. But this is bridgeable, in principle, with future knowledge.	Solubility of the epistemic gap (Nagel, 1986)
D	Dualism	Ontological and epistemological gap. Denies causal closure of the physical (CCP). Causal role for phenomenal properties.	Interactionism Substance Dualism Property Dualism
E	Dualism	Accepts CCP. Phenomenal properties have no causal efficacy.	Epiphenomenalism
F	Monism	Phenomenal properties exist at the fundamental level of physical reality.	Neutral monism Panprotopsychism

2 Physicalism

2.1 General Characteristics.

Swinburne defines Physicalism, at the most general level, in the following way:

> The subtly different forms of physicalism all hold the doctrine that the only events are physical events, some of which may be also in a sense (although not my sense) 'mental events'. (Swinburne, 2013, p.93)

This definition is broadly consistent with Chalmers' view that physicalists (or materialists) ultimately reduce consciousness to a physical process requiring no expansion of physical ontology. Chalmers (2002, p.248) specifies the mind/body problem as the 'hard problem' of explaining just how it is that purely physical processes give rise to consciousness: '…explaining on the basis of natural principles how and why it is that physical processes are associated with states of experience.' More prosaically we might say it is the problem of the fact that people just seem to have insides as well as outsides. The physicalist response to the epistemic gap between physical and phenomenal truths is generally a reductive one in which any solution to the 'hard problem' of consciousness is one in which consciousness is, ultimately, to be interpreted as a physical process. This is what Swinburne rejects in his advocacy of substance dualism, that a full account of the world must encompass phenomenal truths as well as physical.

Broadly physicalism maintains that every substance is a material substance and every property a material property and therefore every event must in some sense be a physical event. This removes awkward questions about just how immaterial minds or souls are connected to, or are correlated with, or are causally related to brains and bodies. However, if minds are not (non-material) entities at all but only 'matter suitably organized' then exactly what are they and how does 'matter suitably organized' give rise to self-consciousness and the sense of 'what it feels like to be me'? It is through this open door that Swinburne enters.

2.2 The Errors of Physicalism.

In this section we will explore Swinburne's general criticisms of physicalism: 'the 'errors of physicalism' (Swinburne, 2013, pp.93-99). The shape of Swinburne's argument is first to show that there is an ontological gap between physical and mental substances and properties by the metacriterion argument. This ontology then entails that there is a fundamental epistemic gap between physical truths and phenomenal truths shown by the privileged access argument. A complete account of the world is only possible by an account of both physical and phenomenal truths. This must mean that physicalism is false. Any explanation of the relationship between mental and physical substances, properties and events will itself involve reference to consciousness itself.

2.2.1 The metacriterion argument.

For Swinburne (2013, p.10) the root of the errors of physicalism lie in the failure to adopt an appropriate 'metacriterion' by which we can adequately represent the '… data of the mental life and physical life of human beings'. The data of this putative science of mental and physical life is constituted by 'events'. What is required, according to Swinburne, is a set of 'canonical descriptions' of events such that anyone who knew the canonical descriptions of the relevant events would be able, from a subset of those events, deduce 'everything that had happened'. The canonical description of an event is constituted by a description in terms of substances, properties and times. The description entails that any competent language user would know which substance, property, time or event was involved. The descriptors here Swinburne refers to as 'informative designators'. These provide conditions for identifying mental and physical events.

Swinburne's argument is that canonical descriptions for mental and physical events are, and must be, different because the informative designators of substances, properties and events involved pick out different things.

> …I am understanding by the 'canonical description' of an event one in terms of the informative designators of the substance(s), property and time involved in the event. The conjunction of those informative designators will constitute an informative designator of the event. We should count any two canonical descriptions as picking out the same event if and only if any possible world which contained such

> an event picked out by the one would contain the other, and conversely. For only then is the occurrence of one event nothing in the history of the world 'over and above' the occurrence of the other event. (Swinburne, 2013, p.27)

Swinburne therefore rejects physicalist theories in general by their failure to meet appropriate identity conditions… 'Logically equivalent informative designators refer to the same thing, and logically non-equivalent informative designators refer to different things' (Swinburne, 2013, p.20). I will examine this argument in greater detail in the following Chapter.

This metacriterion provides for Swinburne the key method for establishing identity and difference. It establishes the grounds for determining when two referring expressions designate, or pick out, the same substance, property and event (in every possible world where they exist). Swinburne's contention is that the failure of philosophers of mind to adopt such a requirement leads them into error. Physicalists, for example, tend to collapse the ontological difference between mental and physical events. Swinburne's metacriterion is the essential key to establishing the identity conditions for substances, properties and events. In Chalmers' terms they lead to a profound epistemic gap between physical and phenomenal truths and the corollary claim that causal closure of the physical (CCP) must be false.

2.2.2 The privileged access argument.

Swinburne here develops what Chalmers (2002) calls 'the knowledge argument' against physicalism in general. This asserts that:

> P1. There are truths about consciousness (phenomenal truths) that are not deducible from physical truths.
>
> P2. If there are truths about consciousness that are not deducible from physical truths then, physicalism is false.
>
> C. Physicalism is false.

This line of argument is supported by the concept of 'privileged access'. 'Pure mental events', in Swinburne's sense, are events to which the subject (substance) has privileged access whereas physical events are events where the subject's access is not privileged. Pure mental events are private events

which are unobservable and known only to the subject in the sense that only I can know what it is to have *my* pain. There can, logically, be no public access. Others can observe my physical states and behaviours (public access) but there are truths about my experience (my mental events) only accessible to me. These contrasting properties of mental and physical events establish for Swinburne that no pure mental event can be the same as a physical event and so cannot supervene on a physical event. This conclusion runs counter to a core claim of physicalism that, in principle, if mental events are reducible to physical events or explainable in terms of physical functions then they are public events.

For Swinburne pure mental events are just not observable in the ordinary way but can be known, accessed, by immediate awareness of the person who has them. In current social cognition research this has been named the 'unobservability principle' (UP). It is the very privacy of our psychological states that 'marks' them as mental. And hence they are properties or events of a mental substance (mind or soul). However, in contemporary work on embodied and distributed cognition UP has been challenged by strategies that defend 'direct perception' (DP). The DP claim is that:

> …expressions and gestures are more than mere signals or signs of thought and feeling; in many cases, they are part of the dynamic process of thinking and feeling itself. (Krueger, 2012, p.169)

If DP theory is shown to be in any way correct this would seriously imperil one of Swinburne's fundamental principles, that of privileged access (see Chapter Four, 'Privileged Access'). But we must now turn to a discussion of how Swinburne deploys his general principles in the critique of not just physicalism or materialism in general but of various species of physicalism.

3 Varieties of Physicalism.

The previous section introduces two key arguments by which Swinburne intends to show that physicalism in general is simply false. But maybe not all

types of physicalism can be forced into this general Procrustean bed. In this section, therefore, we examine Swinburne's approach to specific types of theory picked out in *Mind, Brain, and Free Will* (2013).

3.1 Identity Theory.

Identity theory, Type-B Physicalism, proposes that phenomenal states may be identified with certain physical or functional states. Mental concepts are distinct from physical or functional concepts but empirically refer to the same thing in nature. In Chalmers' vocabulary there is no epistemic gap between physical and phenomenal truths. In the 1950s this theory was championed as a solution to the mind/body problem by philosophers including U.T. Place and J.J.C. Smart amongst others (Heil, 2013, p.69; Nagel, 2012, p.35ff.). This theory aimed to falsify dualism, and its rigorous separation of physical (stuff) and mental (consciousness), in favour of a purely physical explanation of mental phenomena. The theory accepted that mental states are genuine internal states of the persons possessing them but that mental events are simply brain events. Types of 'mental' properties just are types of physical properties. Such types of property as 'having a red sensation' or a 'desire to eat fudge' are types of brain events that are patterns of neurons firing.

U.T. Place is cited by Swinburne (2013, p.25) as the first advocate in1956 of the 'modern version' of mental/physical event identity. Pace writes that:

> … a process or event observed in one way is the same process or event observed in (or inferred from) another set of observations…if the latter event provides an explanation of the former set of observations. (Pace quoted by Swinburne, 2013, p.25)

The exemplar is lightning: an electrical discharge and the observation of light in the sky. According to Swinburne, the identity theorists' conclusion would be '…that an electrical discharge (property) in the sky (substance) is the same event as lightning in the sky because it explains why we observe what we do when we observe lightning' (Swinburne, 2013, p.25). But for Swinburne, of course, what we have here is not one property but two.

In explaining the rise of this theory Swinburne suggests its attraction lies in its apparent simplicity. If we make Swinburne's assumption that there is an ontological difference between mental and physical properties (based on his metacriterion) then there must be a difference between the occurrence of a mental event and the occurrence of a physical event. But this seems to lead to all kinds of difficulties in correlating the two in a lawlike way – 'the nomological problem'. Understanding and explaining the potentially innumerable correlations seems to generate insurmountable complexity which may be ultimately unexplainable and hence the appeal of simplicity of identity theory.

Swinburne's objections to this approach seem to be threefold. First, a methodological one, with 'the nomological problem' that just because science apparently can't explain something this does not imply that, as a matter of fact, it doesn't happen. We might just have to accept that a very complex explanation is required. If this were so then we need to alter our scientific theories to fit the data. For example, if we can't discover correlations specified by the theory then we must seek other explanations.

Secondly, and more importantly, his objections rest on the identity criteria of properties, substances and events and the necessary and sufficient conditions for something to be either a mental or a physical property or event: the metacriterion.

> Since the informative designators of any physical properties are not logically equivalent to those of any mental properties (since there are different criteria for applying the designators), no mental property is identical to a physical property. The criteria for being in pain are not the same criteria for having some brain property… (Swinburne, 2013, p.69)

Thirdly, type-type identity theory faces the problem of 'multiple realizability', that is, some types of mental event may be correlated with different types of brain events, For example, this may be the case for different creatures.

3.2 Identity and Modes of Description: Papineau.

David Papineau, in *Thinking and Consciousness* (2002), presents a further variation on identity theory. This comes under fire from Swinburne for its 'monist' account of physical events and 'conscious events'. Papineau argues that descriptions of physical and conscious events represent (merely) a conceptual difference between ways in which we think about the world. This runs counter, of course, to Swinburne's conception of the substantial difference between the physical and the mental as two entirely different features of the world. Papineau believes that there are good, empirical reasons to believe in the completeness of physics; all physical effects are fully caused by and may be explained by purely prior physical histories (monism) and this provides a fundamental obstacle to dualist accounts. For Swinburne this amounts to the claim that we might tell a complete history of the physical world in physical terms; a position that Swinburne is at pains to criticize and reject in Chapter Four, 'Interactive Dualism', of *Mind, Brain, and Free Will* (2013). He argues, paradoxically, that if the principle of the causal closure of the physical (CCP) were true it would be, in principle, impossible to provide the evidence needed to sustain a justified belief in CCP. (This will be discussed further in Chapter Five, 'Credulity and Testimony', below).

However, Papineau believes that 'physical events are caused only by physical events' and this seems to rule out dualism.

> P1. Conscious events do cause physical events, for example, my conscious thirst causes me to get a glass of wine.
>
> P2. All conscious events cause, and are caused by, physical events.
>
> P3. Physical events are caused only by physical events.
>
> C. All conscious events are also physical events.

According to Swinburne, Papineau's story of mind/body relationships contradicts the key identity criteria for mental and physical properties and mental and physical events. Papineau's claim is that: '… the same property may be picked out by different concepts; and that the difference between the conscious and the physical is really a conceptual difference' (Swinburne, 2013, p.97).

In arguing that the difference between the mental and the physical is a conceptual difference Papineau turns aside from the language of substance and property and its problems in favour of analysis in terms of 'mode of presentation'. The difference between the conscious and the physical according to Papineau is essentially a difference between concepts, the ways we pick things out, and not between properties and so events in the world. Rather than the central issues being one of identity, difference and reference - the difference between the physical and the conscious is one of 'sense'. Talk about being in 'pain' and talk about 'neurons firing' are two distinct modes of presentation of the same physical phenomenon. This we might say transforms a conventional view of the mind/body problem, as though it were a quasi-empirical question, into a question of the sense of our talk about consciousness, talk about stuff, and of their logical relationships.

In response to Papineau Swinburne again deploys his metacriterion requirement for the designation of substances, properties and events. He argues that: 'A mode of presentation' is itself a property. If there are two 'senses', then there are two predicates ('being in pain'; 'C-fibres firing') designating two modes of presentation of *one* supposedly physical property. These predicates, 'being in pain' and 'C-fibres firing', are, according to Swinburne, usually used in English as informative designators. And according to his metacriterion, if they are not logically equivalent then they designate different properties (and events): one – mental - 'being in pain' and two – physical – 'C-fibres firing'.

Even if 'being in pain' is not used as an informative designator of a property but as a designator whose application depends on the physical property underlying the surface property (supervenience) then there are still two different and not logically equivalent ways (i.e. by its surface property or by its underlying nature) by which this supposedly physical property can be picked out. The phenomenal (or surface) is equivalent to the conscious and the material is equivalent to the physical. These different concepts would pick out this one property by means of two different predicates – one of them informatively designating a conscious property possessed by a physical property informatively designated by the other predicate. Swinburne's conclusion is that the 'two concepts account' advocated by Papineau collapses into another 'two properties account' and thus reasserts

the claims of dualism'...if we are to describe the world fully we need to postulate two kinds of property, and two kinds of (logically disjoint) events' (Swinburne, 2013, p.98).

3.3 Functionalism: Putnam and Armstrong.

Another and important group of physicalist theories which Swinburne takes to task are functionalist and for good reason in that according to Heil: '…These days functionalism dominates the landscape in the philosophy of mind, in cognitive science, and in psychology' (2013, p. 87). Chalmers considers functionalism as a form of identity theory (Type-B Physicalism) in that it 'identifies' phenomenal states with functional states. Equally, in common with identity theory, it seems to provide the methodological underpinnings for a 'scientific' solution to the mind/problem. Its differences lie in its attempts to overcome some of the limitations of the theory. Functionalist explanations, after the manner of computer science, aim to explain minds at a higher level of abstraction than that of the physical hardware of the system (brains). Functionalism is, in this sense, not reductionist - a characteristic it has in common with Swinburne's approach. Similarly, unlike other physicalist theories functionalism doesn't rule out altogether the conceivability of immaterial substances (Cartesian Souls) but largely presupposes that probably all substances are in fact material substances (Heil, 2013, p.87). Mental events supervene on physical events. Of course, this is a move that Swinburne cannot accept.

Functionalism rejects the view that states of mind are fundamentally states of the brain – in favour of the assumption that states of mind are linked to, or to be identified with, 'functional roles'. The functionalist model of the mind is that of an input – output system in the sense that experiencing some mental state such as pain is simply to manifest input-output conditions characteristic of pain (which is also no doubt also a brain state in humans). I have a hangover because I drank too much red wine; the input of red wine on this model causes characteristic outputs, behaviours, holding my head, saying I have headache, taking an aspirin and so on. However, it is a part of the functionalist picture that such

behaviours do not exhaust my responses. Functionalists reject the (conceptual) behaviourist claim that all mental states can be analysed, or explained, in terms of behavioural dispositions. My headache will probably lead me to formulate various beliefs and intentions such as to remedy my headache by taking aspirin (Heil, 2013, p.9).

For Swinburne this will not do. Pure mental states, in particular, are such just because they do not depend on inputs and may not in fact issue in any outputs. According to Swinburne given the requirement of his metacriterion:

> …no pure mental event is identical to or entails any (contingent) physical event involving the same substance, and so cannot supervene on any physical event. (Swinburne, 2013, p.70)

Functionalism implies that mental events must supervene on physical events but does not privilege any physiological arrangements or material composition to produce minds. One of the features that motivate functionalism's popularity is the analogy it suggests with computing hardware, software and levels of abstraction.

In contrast to identity theory functionalism assumes that states of mind are capable of 'multiple realization'. Mental states are characterized by their roles, but these roles may be realized by different types of material or physical states. Identity theorists map types of mental states to neurological configurations. This would imply that other types of creatures who do not share our neurological make-up would not, for example, feel pain, or heat and so on. This seems empirically, at least, problematic. For functionalists being in pain or feeling heat might be realizable by very different physiologies.

Swinburne cites Hilary Putnam's paper 'The Nature of Mental States' (Putnam, 2002, pp.72-79) as an example of a functionalist approach. Putnam defends the hypothesis that 'being in pain is a functional state of the organism' (2002, p.76) against two alternative hypotheses: (a) that pain is a brain state [identity theory] and (b) that pain is a behavioural disposition [behaviourism]. Putnam argues that:

> …the identification of psychological states with functional states means that the laws of psychology can be derived from statements of the form 'such-and-such organisms have such-and-such Descriptions' together with identification statements ('being in pain

> is such-and-such functional state' etc.). The presence of the functional state…is not merely 'correlated with' but explains pain behaviour on the part of the organism. The identification serves to exclude questions … representing an altogether wrong way of looking at the matter e.g. 'What is pain if it isn't either the brain state or the functional state?' … and leads to both fruitful predictions and fruitful questions. (Putnam, 2002, p.79)

Swinburne glosses the functionalist view in his own terms as the belief that:

> …what makes any property a property of a kind which I have called a 'pure mental property' is that events with that property have a certain function in a person's life of thought and behaviour and in particular tend to have certain kinds of causes and effects (in or outside the brain). (2013, p.95)

He goes on to interpret this as meaning that having a pain is just the property which events have if they tend to be caused by, for example, physical damage. And this will cause a desire for the pain to cease.

> What, according to the functionalist, makes a desire or any other mental property the property it is, is the place of events with that property in a whole web of interrelated mental events (thoughts, sensations, intentions desires). (Swinburne, 2013, p.95)

Swinburne's objections to functionalism in general, and Putnam in particular, are that they represent (a) a form of monism and (b) a form of substance physicalism in asserting that human beings are the same substance as their physical bodies. A token mental event is then the same token event as some brain event even though the two types may not always be correlated. Such a view:

> …simply ignores the fact that there are pure mental events, picked out in English by such words as 'pain', 'afterimage', 'purpose', and 'thought' to which their possessor has privileged access; and any account of the nature of human beings has got to give these a significant role. (Swinburne, 2013, p.95)

Here we see Swinburne's criticism of functionalism at the intersection of his epistemology (the principle of privileged access) and ontology (the metacriterion requirement for the identification of substances, properties and events).

Another functionalist to receive attention from Swinburne is D.M. Armstrong. In *A Materialist Theory of the Mind* (1968) Armstrong pioneered a variation on the

functionalist theme emphasizing functional roles in terms of causes and effects. Mental states are 'apt' to produce causes and effects. Sensation, knowledge, intention and so on are analysed according to the different effects they are 'apt' to cause:

> …Suppose I have the intention to strike somebody. My mind is in a certain state, a state that I can only describe by introspection in terms of the effect it is apt for bringing about: my striking that person. (Armstrong quoted in Swinburne, 2013, p.95, fn.)

He argues that events can be understood as mental events by the fact that they are 'apt' to be caused by other events of particular types and are given to causing events of other types. To be a mental event of a certain type is ultimately to supervene on being a physical event that evinces certain normal causes and effects.

Swinburne argues that Armstrong's and other similar functionalist characterizations of mind are simply 'absurd'. He seems to rest this criticism on an interpretation of the 'normal' requirement in 'certain normal causes and effects'. The consequence of such a theory in which:

> …a brain event is the mental event it is in virtue of the normal causes and effects outside the body of similar brain events is that any brain event would constitute spots in the subject's visual field or that person's intention to hit someone if similar brain events caused or were caused by the normal extra-bodily manifestations of that mental event and not if they didn't. So whether a person had spots in their visual field or intended to hit someone would depend not what was happening in them but what happens in most other people – which is obviously (sic) false. (Swinburne, 2013, pp.95 – 96.)

Even if we suppress the 'normal' requirement, i.e. if a mental event of a particular type is any brain event with certain causes and effects outside the body, absurd results seem still to follow:

> …any event caused by a bodily disorder and causing the subject to wince was a pain, and any brain event which caused the subject to hit someone was an intention to hit them; and so on. (Swinburne, 2013, p. 96.)

He observes that people frequently wince when they ‘think’ some medical intervention may cause them pain even when this is not the case and similarly some people seem to hit others ‘unintentionally’.

Swinburne maintains that functionalist accounts, resting on certain physical causes and determined effects, are unsustainable. His principal objection, as with his objection to identity theory is that the theory does not acknowledge *his* category of pure mental events. And that such mental events may be picked out by the use in English by such expressions as ‘pain’, ‘after-image’, ‘purpose’, and so on to which a person has a unique and privileged access. He reverts to a re-statement of fundamental elements of his ontology and epistemology.

To summarize we have considered two trends in the contemporary philosophy of mind: identity theory and functionalism. These are both classified by Chalmers (2002) as Type-B Physicalism (or materialism) because their program is the identification of phenomenal states with either physical or functional states (of organisms). In both cases Swinburne has reasons for rejecting them rooted in his metacriterion requirement and the concept of privileged access. In the next section we will consider Eliminativism (Type-A Physicalism). A fundamental assumption here is that there is no relevant epistemic gap between physical and phenomenal truths. Once the so called ‘easy problems’ have been solved by neuroscience and psychology then there will be no ‘hard problem’ of consciousness remaining. Again, Swinburne argues that such an approach is deeply in error.

3.4 Eliminativism: The Churchlands.

Eliminativism, championed by Patricia and Paul Churchland, rejects both dualism and so called ‘folk psychology’ (Heil, 2013, p.150 ff). Dualist accounts of mind and folk psychology’s descriptions of mental contents are misconceived because the terms they use simply do not refer to anything. Eliminativists argue that we must abandon certain ways of talking about minds and bodies. They maintain that there just are no intentional states, beliefs, desires and no reasons for action and so on. What is to be eliminated or reduced is certain non-scientific ways of talking and accounting for our mental furniture.

The language which we use to describe and explain our mental goings-on is wholly misleading unless it can be reduced or translated into neuro-physiological categories. The eliminativist program replaces the language of folk psychology with a scientifically based understanding of how the neural activities explain our cognitive characteristics. Talk about minds is not talk about anything in the world; there are no such mental entities:

> …psychological categories and modes of explanation are either 'reducible' to more fundamental neurobiological categories, or psychological categories and modes of explanation apply to nothing: beliefs, desires, and the lot are like phlogiston or witches, non-existent. (Heil, 2013, p.150)

Swinburne gives little space to the eliminativist case apart from a long footnote (Swinburne, 2013, p.99, fn.). The Churchlands are mentioned in the context of Swinburne's defence of a person's privileged access to her 'propositional events'. He takes the Churchlands to be arguing the following:

> P1. There is no one-to-one correspondence between types of propositional events such as beliefs and brain events.
>
> P2. There is no privileged access to such propositional events.
>
> P3. Brain events are caused by other brain events and cause our behaviour.
>
> C. There can be no such things as propositional events.

Propositional events, and similarly sensations, have no role in any scientific theory. Swinburne is happy to accept the first premise but, of course, rejects the other premises and the conclusion. He has already argued, in Chapter Three, 'Property and Event Dualism', of *Mind, Brain, and Free Will* (2013) that we do, as a matter fact, have privileged access to our propositional beliefs, intentions, thoughts, desires and sensations. He goes on to justify the claim that beliefs and intentions (via brain events) do cause our behaviours (Swinburne, 2013, pp.100-124). I form an intention to raise my arm (a mental event) which causes the occurrence of my arm rising (physical event).

If the philosophy of mind is primarily concerned with the logical relations between talk about consciousness and talk about stuff (albeit 'neuro-biological'

stuff) then the Churchlands want a total revolution, a total reconstruction, of 'mind talk'. The proposal is an attempt at re-definition in which experience (beliefs, sensations, feelings and so on), which, at present, carry an essential reference to consciousness, are re-defined to refer to the physical goings on in the brain. This program is exemplified in the thrust of Paul M. Churchland's *The Engine of Reason, the Seat of the Soul* (1995) subtitled 'a philosophical journey into the soul' but is in fact an exploration of recent research in neuro-science together with artificial neural networks. But this reductive program like others of this lineage (Type-C Physicalism):

> …only begins to look sensible if we mistake it that the one word 'experience' could carry, simultaneously, both its present and a new and incompatible sense; and hence that experience, which essentially involves consciousness, could be reduced to mere bodily movements, which do not. (Flew, 1975, p.74)

Despite the fancy neuro-science footwork, in the face of our actual experience, there are no good reasons to abandon our everyday 'consciousness talk' in favour of scientific concepts about neurological stuff which mean entirely different things. For Swinburne this view conforms to 'the principle of credulity': '…that what seems to us to be so probably is so, that our apparent experiences are probably real experiences' (Swinburne, 2013, p.42). The eliminativist view runs completely counter to this principle. However, in the next section we examine Daniel Dennett's instrumentalist theory which aims to defend, extend and elaborate the Churchlands' approach.

3.5 Instrumentalism: Dennett.

Daniel Dennett in 'True Believers: The Intentional Strategy and Why it Works', extends the Churchlands' work on eliminative materialism (Dennett, 2002, pp.556-567; Heil, 2013, p.141). He develops a form of instrumentalism that is critical of dominant forms of functionalism. Dennett argues that nothing in the world corresponds to, or 'answers' to, our talk of mental states and processes in the way that both functionalist and substance dualists suggest. (The arguments are similar also to those advanced by Dennett's former tutor Gilbert Ryle and

also another Oxford luminary in the tradition of linguistic analysis, J.L. Austin). Dennett's program is that such talk should be, echoing the Churchlands, 'eliminated' from descriptions and explanations of the nature and behaviour of sentient creatures; primarily flesh and blood creatures such as ourselves. Dennett's alternative strategy is to treat minds as, in effect, social constructs arising out of the 'interpretative practices of language users'. Talk about minds depends (or ought to depend?) on its usefulness (utility) in describing and explaining complex sentient systems.

Such descriptions and explanations may be categorized in terms of a triad of 'stances': The intentional stance' (mindful behaviour); the design stance (mechanism and function); the physical stance (physical foundations); (Dennett, 2002; Heil, 2013, p.130-135). Dennett's view is that for a creature to have beliefs and desires about the world is just (merely) the ability to explain and describe that creature's behaviour from the position of 'the intentional stance'. Beliefs, desires, thought processes, and so on are not inaccessible, private, undetectable, immaterial entities. The analogy drawn here is with a child looking at a globe and seeing a line which marks the equator and mistaking this for some actual physical feature or property of the world. We then look for the specific feature or property but can't find it and so go on to deduce that we can't find it because it is a kind of 'immaterial' substance or property (Heil, 2013, p.132).

By this analogy we might say that Swinburne is like the child who mistakes a representation of the equator for some actual feature but immaterial feature. He certainly believes the mind to be a real but immaterial substance, with real properties. He mistakes talk about minds and their states, characteristics, and behaviour for actual but immaterial mental substances, properties and events. In *Mind, Brain, and Free Will* (2013) Swinburne does not explicitly cite the work of Dennett but his realist framework and privileged access argument are clearly counters to Dennett's arguments. Swinburne as we have seen talks of minds and their properties as on a par with those 'material things', the furniture of the world, such as plants, rocks, trees and so on. Mental substances are no less 'real'. He returns again and again to the proposition that in order to give a complete description of the world we must (logically) give that account in terms of physical and mental substances, properties and events.

3.6 Event Identity: Donald Davidson.

Swinburne makes short work of Donald Davidson's variant identity theory: event identity. The ground for Swinburne's refutation is a reminder of the metacriterion requirement:

> …I am understanding by the 'canonical description' of an event one in terms of the information designators of the substance(s), property, and time involved in the event. The conjunction of those informative designators will constitute an informative designator of the event. We should count any two canonical descriptions as picking out the same event if and only if any possible world which contained an event picked out by the one would contain an event picked out by the other. (Swinburne, 2013, p.27)

Davidson's event identity criterion states that two events are identical or equivalent '…if and only if they have the same causes and effects' (Davidson, 1980, p.179).

Swinburne rejects this on the grounds that:

> P1. The criterion could only apply to events which are caused and have effects;
>
> P2. There is no reason to suppose that all events have effects (for example, elsewhere Swinburne argues that 'pure mental events' may have neither causes nor effects).

Even if we accept it only applied to events with cause and effects we can conceive, even so, of instances where even here it would fail. He then argues that this would apply in cases of 'over-determining' causes. He proposes the following (convoluted) case to support this contention. Suppose an event has two different effects E2 and E3 *each of which causes* E4 and suppose that each is a partial cause of E4 (a necessary condition). Swinburne suggests that for Davidson's principle to be plausible the phrase '*each of which causes* E4' must be understood as '*each of which is a sufficient cause of* '. However, the consequence of this is that E2 and E3 could still be distinct events if E4 is over-determined that is, if either E2 or E3 would have caused E4 without the other. We would need to provide an additional clause to the original principle along the

lines ‘so long as effects are not over-determined’ – so long as there are not two sufficient causes of E4. ‘Then if E2 and E3 are both sufficient causes but not separate sufficient causes, they will be the same sufficient cause of E4 and so the same event’ (Swinburne, 2013, p.29).

The nub of this example, Swinburne claims, is that it shows Davidson’s argument is ‘viciously circular’ and ‘useless’ because Davidson’s principle provides no criterion by which to determine whether E2 and E3 are, or are not, separate sufficient causes or the same sufficient cause. And that will depend on whether or not they are the *same* event and that cannot be determined by the theory. A further criterion of event identity would be needed to show whether the events are the same or different in other words the principle fails to be an adequate criterion of identity.

Elsewhere Swinburne points out some agreement with another thesis defended by Davidson to the effect that: ‘…there cannot be strict psychophysical laws’ (Davidson, 1979, p.220). (This position is important in Swinburne’s later arguments for free will.) Swinburne agrees that Davidson’s thesis here is the same as his if we understand ‘strict’ as meaning ‘general’. However, where again he differs from Davidson is on the question of criteria for physical and mental event identity. In his paper ‘Mental Events’ Davidson puts the case for Anomalous Monism. This variation on mind/body identity theory asserts that while all events are physical not all mental events can be given physical descriptions. Davidson argues that physical-mental causal interaction is the law-like causal interaction of two physical events. Mental characteristics are in some sense dependent or supervenient, on physical characteristics but ‘…dependence or supervenience here does not entail reducibility through law or definition’ (Davidson, 1979, p.225). This clearly is at odds with Swinburne’s dualist position that events are of two very distinct types, physical and mental (in his sense) (Swinburne, 2013, p.190, fn.). The various physicalist theories we have discussed are, according to Swinburne, in error and we have sought to point out the grounds on which Swinburne believes this to be the case.

4 Non-Physicalist Theory.

In this section we consider briefly several theories which, in common with Swinburne's substance dualism, present non-physicalist solutions to the mind/body problem: Property Dualism (Type-D theory), and Epiphenomenalism (Type-E theory). They are non-reductive and require an expansion or re-conceptualization of a strictly physicalist ontology.

4.1 Property Dualism.

Property Dualism rejects the causal closure of the physical (CCP), embraces the idea of an epistemic gap between physical and phenomenal truths but argues for a causal role for phenomenal properties (Chalmers, 2002, p.261). On this view, although there is only one type of substance, or entity, this has both physical and phenomenal fundamental properties. In addition, phenomenal properties are irreducible to physical properties and play a role in affecting physical properties: '...the "downward" causation of the mental on the microphysical'. In summary, according to Heil (2013, p.185), the precepts of property dualism are that:

P1. Mental and material properties are distinct kinds of properties.

P2. A substance can possess both mental and material properties.

P3. Mental properties depend on material properties.

These are at least consistent with Swinburne's view of the irreducibility of mental (phenomenal) properties to physical ones. In concluding his survey of the 'errors of physicalism', for example, Swinburne writes that:

> There just are what I describe as 'physical events', and 'pure mental events' (including conscious events) which are different from each other, and philosophers should face up to this. (Swinburne, 2013, p.98)

Swinburne cites, with approval David Chalmers' work *The Conscious Mind* (1996). He suggests that this was responsible for influencing many philosophers

(unnamed) to recognize property, and so in his sense, event dualism. He quotes Chalmers to the effect that there are:

> …both physical and non-physical features of the world. The falsity of logical supervenience implies that experience is fundamentally different in kind from any physical feature. (Chalmers quoted in Swinburne, 2013, p.98)

Swinburne suggests (2013, p.98, fn.) that '…Chalmers seems to allow the possibility of (what I am calling) 'propositional events' having phenomenal (i.e. in my terms 'sensory') properties, and so not being fully supervenient on the physical'. Chalmers argues that it is not the (physical) biological and chemical constitution of creatures that is the foundation of consciousness. Consciousness 'arises from' functional organization which in turn emerges from various processes and physical states. A particular kind of *functional* state will give rise to a particular sort of conscious experience. In summary then Chalmers acknowledges a type of property and event dualism in a similar sense to Swinburne's own usage. Property Dualism on this account steers a course between the Scylla of pure physicalism and the Charybdis of full blown Cartesianism.

4.2 Epiphenomenalism.

Another variant of dualism, epiphenomenalism, recognizes an epistemological gap between physical and mental events. For example, Broad, in his seminal *The Mind and Its Place in Nature* (1925), defines epiphenomenalism according to four key propositions:

> P1. Certain bodily events cause certain mental events.
>
> P2. No mental event plays any part in the causation of any bodily event.
>
> P3. No mental event plays any part in the causation of any other mental event.
>
> C. All mental events are caused by bodily events and by them only.

Conscious events never cause brain events and all conscious events have physical causes such as brain events. This is incompatible with interactive dualism. Swinburne believes, like physicalism, although for different reasons, that epiphenomenalism is in error. There are good arguments to show, according to Swinburne, that there is interaction between conscious events and physical events. This position, Swinburne claims, corresponds to our everyday understanding of what's going on – 'the normal view' – when we form and execute intentions:

> P1. We just do, as a matter of fact, cause physical events to occur by forming and carrying out intentions (i.e. pure mental events).
>
> P2. We normally believe that our intentions guide our bodily movements, and this is just how it seems to us.
>
> P3. The scientific consensus is that bodily movements are caused by brain events which are preceded by the formation of intentions.
>
> C. If intentions cause movements they do so by causing brain events.

However, I postpone a fuller treatment of these reasons and arguments to Chapter Five, 'Credulity and Testimony'.

5 Conclusion.

In foregoing sections, I have picked out Richard Swinburne's reactions to a variety of trends in contemporary philosophy of mind. We have concentrated on those theories which Swinburne considers compete for the same (logical) space as substance dualism and which he decisively rejects. They have included physicalist theories: identity theory; functionalism; eliminativism; instrumentalism; and event identity; and a non-physicalist, dualist theory: epiphenomenalism. These are theories which provide alternatives to key aspects of Swinburne's substance dualist solution to the mind/body problem. They maintain, for example, a closed physical ontology and the reducibility of phenomenal truths to microphysical ones.

Swinburne's rejects such theories on the following grounds: (a) ontologically that physicalism is inadequate for a complete description of the world and; (b) epistemologically that phenomenal truths are not reducible to microphysical truths. The world is not a causally closed physical system and there is a causal nexus between phenomenal and microphysical properties. A complete description of the world must include two fundamental types of substances, properties, and events i.e. mental and physical.

I have identified three key arguments by which Swinburne supports these claims. First, he uses the concept of a metacriterion requirement to establish the kind of canonical vocabulary which will allow him to prescribe the identity conditions for mental and physical substances, properties and events. Secondly, a privileged access argument provides the basis for further defining the mental/physical distinction and the nature of phenomenal truth. Thirdly, Swinburne believes that the principles of credulity and testimony are fundamental, *a priori*, and epistemic principles in the refutation of physicalism and grounds for interactive dualism. In the following chapters we develop the consequences of these arguments for Swinburne's answer to the mind/body problem as substance dualism.

Chapter Three: A Metacriterion.

1 Introduction.

I suggested in the previous chapter that a central pillar of Swinburne's thinking in *Mind, Brain, and Free Will* (2013) is the metacriterion requirement. This criterion is briefly '…the requirement that our vocabulary should enable us to tell the whole history of the world.' He believes that a complete history of the world entails a history of physical events but also a history of mental events. To achieve this aim he believes that canonical descriptions will provide comprehensive descriptions of the world in terms substance, properties, events and times. The words in the canonical vocabulary will be such that a competent language user would be able to identify which substance (mental or physical), property and event is involved. The words which pick out substances, properties and events are, according to Swinburne, informative designators which designate the same object 'in every possible world' in which it exists. An object, for the purposes of this theory of designation, is a substance, property or event either mental or physical. Such objects are characterized by, identified by, some essential property. The non-identity of mental and physical substances, properties and events establishes the veracity of substance dualism.

In this chapter I analyse Swinburne's arguments for such a metacriterion. I examine the ways in which this rests on a particular model of how language is supposed to relate to the world; Direct Reference Theory (DRT). This he appears to adopt from the work of Saul Kripke (1980) and Hilary Putnam (1975). An important element in this story is the doctrine of 'natural kinds'. I go on to criticise both the doctrine of natural kinds and Swinburne's fundamental ideas about informative designators. The theory severely over simplifies the ways in which language operates generally and in relationship to mental concepts in particular.

2 The Metacriterion Argument.

The metacriterion, according to Swinburne, underpins identity criteria for mental and physical substances, properties and events: '…determining when we should deem that two referring expressions pick out the same property and so on' (2013, p.10). This emphasis given by Swinburne to DRT in *Mind, Brain, and Free Will* (2013) is a relatively new feature of his work compared to, say, *The Evolution of the Soul* (1987). The key move is that once we have discovered the essence of substances we can derive metaphysically necessary conclusions:

> When we know fully what we are talking about (e.g. that in talking about 'water', we are talking about 'H_2O'), mere a priori considerations can show whether some sentence is metaphysically necessary or impossible for this kind of reason. Hence there is available a definition of a sentence as metaphysically necessary (impossible or possible) iff it is logically, necessary, impossible or possible) when we substitute co-referring informative designator for uninformative designators, that is designators which pick out the same substance or property. (Swinburne, 2013, p.19)

2.1 Swinburne's philosophy of language.

Swinburne asserts that the task of the philosopher is to stipulate relevant rules for terms which can adequately describe 'a complete history of the world'. Swinburne believes that our ordinary talk about our physical and mental lives is simply inadequate to the task of developing a systematic 'scientific' approach to the mind/body problem. He argues that only a specialized vocabulary can accurately describe 'the data' of our mental and physical lives. Ordinary language is too limited, too vague, too ambiguous, to do the job. For example, the word 'property' has a range of meanings and usages; 'characteristic', or 'feature' or 'aspect':

> But 'ordinary language' has no precise rules for when two properties are the same. Is red the same property as the property of reflecting light of such-and-such a wavelength, or is being triangular (having the shape of a closed rectilinear figure with three interior angles) the same property as being trilateral (having the shape of a closed rectilinear with three sides)? Ordinary usage provides no criteria. (Swinburne, 2013, p.10)

Swinburne's metacriterion is the requirement that our vocabulary ought to be able to tell the whole history of the world. And by 'vocabulary' here Swinburne means the 'referring expressions' that will identify (or differentiate) the key constituents of the world: substances, properties, times, and events. He simply stipulates that these are the core categories which will be the locutionary building blocks for his ontology. The requirement to tell this putative history determines that two equivalent referring expressions pick out the same property, substance and times. Referring expressions are required to pick out the same property (or whatever) iff that property (or whatever) can be designated by the same informative designators. Thus, two properties are the same if their designators are logically equivalent.

> …I am understanding by the 'canonical description' of an event one in terms of the informative designators of the substance(s), property and time involved in the event. The conjunction of those informative designators will constitute an informative designator of the event. We should count any two canonical descriptions as picking out the same event if and only if any possible world which contained such an event picked out by the one would contain the other, and conversely. For only then is the occurrence of one event nothing in the history of the world 'over and above' the occurrence of the other event. (Swinburne, 2013, p.27)

Is Swinburne equivocating between strong supervenience and identity? For example, it could be the case that in every possible world some event A occurs iff event B occurs, but they are still distinct events. He seems to believe that what we have here is a causal relationship, but the relevant informative designators would pick out events A and B as logically different events. If the events referred to here are a physical event (B) and a mental event (A) Swinburne would want to deny supervenience. He interprets the claim that mental events 'supervene on' brain events as meaning that mental events are 'constituted by' or 'realized in' brain events. The stipulation of identity conditions for mental and physical events forestalls this claim. His definition of the supervenience of substances is: '…A-substances supervene on B-substances iff necessarily for every A-substance x there is a B-substance y, such that

necessarily if y exists x exists' (Swinburne, 2013, p.22). Swinburne is arguing that no mental substance necessarily supervenes on a physical substance.

'A competent language user' would know fully which properties, substances, times and events are involved by using the appropriate canonical descriptions picked out by informative designators. 'Events which involve the instantiation of the same properties in the same substances at the same time will therefore be identical' (2013, p.27). Canonical descriptions of physical and mental events cannot be the same:

> P1. If our canonical descriptions refer to the same objects, then the descriptions of data of human mental and physical life are logically equivalent.
>
> P2. This would sustain the correctness of 'identity theory', the prime physicalist account which makes the following types of claims: 'mental events are just brain events'; 'humans are just complicated physical organisms'; and so on.
>
> P3. But the canonical descriptions of the data of the mental and physical lives of human beings are profoundly different.
>
> C. Identity theory must be false.

Thus:

> …being red (being the property picked out by the informative designator 'red') and 'reflecting light of such-and-such wavelength' are not the same property – since having one does not entail having the other or conversely. If one said that these properties are the same although the designators are not logically equivalent, then knowing that some object had the property picked out by 'reflecting light of a certain wavelength' would not tell you all there is to be known about it; that is red (i.e. looks a certain way to most people) is a further piece of information about the world. (Swinburne, 2013, p.23)

Swinburne rejects criteria of identity which are *a posteriori* in favour of identity criteria which are *a priori*: the criterion of identity is synonymy. However detailed a physical description may be, a mental description cannot be deduced from it: '…For one property to supervene upon another requires the possibility of

deducing the presence of the supervening property from the presence of the conditional property' (Holdsworth, 2014, p.636). This is fundamental to Swinburne's defence of property dualism.

William Jaworski (2014, p.19) criticises this approach, especially the assumption that language users who can correctly apply terms like 'red' know the 'essence' of what they are referring to. It is, Jaworski suggests, plausible to suggest that correctly using terms like 'red' by a 'competent speaker, favourably positioned with faculties in working order' means I can pick out red things in the environment. But even if this is the case it just doesn't follow that we would thereby know the metaphysically necessary and sufficient conditions for 'being red'. He illustrates the point in the following way: if at a reception you inform me that people with yellow tickets only should be allowed through the door then certainly I can pick out the permitted guests. However, even though I can pick out such guests I am not required to know the conditions, how and why, these particular people were given yellow tickets.

But Swinburne's assumption is that words in a language have a 'correct use' and that '…normally whom or what is designated by a word is the same whoever uses it' (Swinburne, 2014b, p.53). What is more the sufficient and necessary conditions for a thing or person to be that or who they are, are similarly built in. To describe the world, he claims, we require a language that mirrors the world. Swinburne rejects the capacity of ordinary, everyday language to adequately represent 'the data of the mental and physical life of human beings'. He believes that what is needed is some special (philosophical) vocabulary which will accurately enable us 'to tell the whole history of the world'. (And presumably if we are not able to do such 'telling' then we cannot resolve the mind/body problem.) In the next section I go on to characterize and critique this approach. It is Swinburne's claim that the deployment of 'informative [rigid] designators' necessarily entails 'names' of the objects which they pick out.

There are important risks to Swinburne's approach here. He is choosing to redefine what he considers to be vague, ordinary language terms to make them more precise in a way stipulated by a proposed new usage. But doing so tends, in effect, to give another and different meaning to the words. We may not be

talking then about the same sense or reference or indeed talking about anything at all.

First, it may be, as Swinburne claims, that ordinary language is vague in the sense that it does not provide ‘precise rules for when two properties are the same’. It maybe that this is because such precise rules are not possible, or necessary, to formulate or that ordinary usage is quite capable of handling the difference between a scientific description of light and wavelength and our ordinary use of the term ‘red’. From the fact that the terms of two descriptions are not logically identical it doesn’t follow that they don’t refer to the same thing.

Jaworski points out the argument on its own does not establish Swinburne’s claim to the dualism of nonphysical and physical properties. It only implies that:

> …properties which underwrite the correct use of “red” must be different from those that underwrite the correct use of “reflecting such-and-such wavelengths of light”, and it does not follow from this alone that the former properties must be nonphysical and the latter physical. (Jaworski, 2014, p.21)

This amounts to a form of dualism about the conditions under which predicates, or terms are correctly applied. Swinburne (2014, p.54), however, rejects this interpretation because the case for nonphysical properties is established by the fact that innumerable properties are individuated by informative designators which refer to their instantiation in substances that have privileged access to such instantiations.

Secondly, Swinburne means us, for example, to understand the word ‘mind’ as meaning, or being identical to, the expression ‘pure mental substance’. In effect he is proposing a new set of meanings to mental conduct words. He is claiming that the *definiendum*, ‘mind’, means the same as the *definiens* a ‘pure mental substance’. In everyday usage we just don’t mean by mind a ‘pure mental substance’. Nor does the act of stipulative definition imply that the new term has reference; it is a recommendation or imperative for future usage. However, he believes these radical linguistic moves can be justified by the adoption of a particular denotative theory of language discussed in the next section.

2.2 Direct Reference Theories.

Swinburne's case for picking out 'minds' as 'mental substances' draws on Kripke's version of a theory of 'direct reference' set out in *Naming and Necessity* (1980) first published in 1972. Kripke's concept of 'rigid designators' (a theory of 'proper names') is fundamental to showing the non-identity of physical and mental substances, properties and events. Kripke defines 'rigid designation' as '…a term that in every possible world designates the same objects' (1980, p.217).

The notion of direct reference promoted by Kripke is critical of the work of Frege (1848-1925). In a paper Frege (1988), first published in 1892, draws a distinction between the sense (*Sinn*) of an expression and its reference (*Bedeutung*) (Baldwin, 2001, p.59; Blackburn, 2006, p.340). Frege argues that two expressions can have the same reference but present this reference in different ways. This 'mode of presentation' constitutes the sense of the expression. For example, the terms George Orwell and Eric Blair refer to the same person. But I might be unaware of this and be surprised to discover that the person I had known as Eric Blair turns out to be the person I know as George Orwell. It is the sense of expressions that determine the thought expressed by a sentence in which they occur. In contrast reference determines its truth or falsity. For Frege, proper names possess both *Sinn* and *Bedeutung*, sense and reference.

Kripke asserts that names are without sense and acquire their references by procedures such as ostension and then proceed from speaker to speaker via a 'causal chain'. Thus, Direct Reference Theories (DRT) are sometimes also referred to as Causal Theories of Reference (Grayling, 1990, pp.182-186). Each speaker receives the name with the intention to use it to refer to the same object referred to by the speaker from whom he received it. Proper names, therefore, lack descriptive content and refer directly to their objects. Kripke uses this 'rigidity' of names, in every possible world, to support a metaphysical thesis about the necessity of identity and of origin.

Swinburne adapts Kripke's terminology of 'rigid' and 'non-rigid' designators and uses the terms 'informative [rigid] designators' (performing the function of

names) and ‘informative [non-rigid] designators’ (descriptive phrases which may change their reference).

> For a rigid designator of a thing to be an informative designator it must be the case that anyone who knows what a word means (that is has the linguistic knowledge of how to use it) knows a certain set of conditions necessary and sufficient (in any possible world) for a thing to be that thing (whether or not he can state those conditions in words.) Two informative designators are logically equivalent if and only if they are associated with logically equivalent sets of necessary and sufficient conditions. (Swinburne, 2013, p.12)

An informative designator is a term which, in every possible world, *necessarily* designates (denotes) the same object. He goes beyond Kripke by extending the supposed special referential status of proper names to a range of other types of words including common nouns. ‘Red’ invariably refers to, designates, the same colour (Swinburne, 2013, p.11).

Swinburne’s version of DRT trades on the intuition that a proper name provides a more direct route to denoting a person (or object) than a description. In addition, it suggests that a name attaches to its bearer in a special way. This presumes, of course, two things. One is that the person or object exists and, secondly, the person or object has gone through some naming ritual (a metaphorical ‘baptism’). But once established a proper name must have one bearer and exactly one; identity is then necessitated. In every possible world where he exists ‘Richard Swinburne’ is ‘Richard Swinburne’ and necessarily in all possible worlds but will only possess such properties as are essential (i.e. not transiently descriptive) for being ‘Richard Swinburne’ which presumably would not include ‘Professor of Philosophy of Religion at the University of Oxford’. Proper names on this account directly refer rather than describe. Names and description have similar grammatical roles but differing logical forms; for example, Donald Trump and ‘the occupant of the Oval Office’. These clearly are not equivalent – the latter being a description which applies at some times but not others (Stroll, 2000, pp.215-217).

In Swinburne’s interpretation DRT overcomes limitations of ordinary language that prevent the systematic development of identity criteria, for example, for mental and physical properties. It provides a means of fixing a canonical

vocabulary. If this proposed canonical vocabulary operates on the model of proper names (extended to include common nouns) it implies that, if once our canonical vocabulary has been established, then the objects that it names are *necessarily* those objects. Descriptions refer only indirectly i.e. 'pick out' objects through the intermediation of descriptive phrases. On the DRT model a true identity sentence whose flanking expressions are properties are necessary and not contingent. The theory assumes that names, in order to count as 'informative designators', must (logically) denote objects that exist rather than fictional names.

Kripke claims that the theory of rigid designators leads to several significant but contentious conclusions. Importantly, from the point of view of Swinburne's adoption of Kripke's theory, is the claim that some necessary truths can be known *a posteriori* (Grayling, 1990, p.98). For example, 'Water is H_2O' is a true identity statement whose truth is discovered empirically and yet one which subsequently becomes a necessary truth in all possible worlds in which water exists.

Swinburne builds Kripke's theory into his metacriterion. Thus:

> But the requirement that we be able to tell the whole history of the world (putting someone in a position to know everything that has happened, is happening, and will happen), by listing a subset of events which entail all events, does put restrictions on how we can do this, restrictions on the possible identity criteria. The restriction…is that the properties, substances, and times constituent of events must be picked out by informative designators. Since the identity of an informative designator is constituted by the criteria for its application, such designators will pick out the same properties and so on iff they are logically equivalent. (Swinburne, 2013, p.23)

However, there is a problem with talk of 'necessity' embedded in the idea of informative designators. O'Hear (1995, p.56) points out that the concept of 'necessity' normally attaches only to how things are described (i.e. that it is *de dicto*). Necessity, at least from a logical point of view, does not attach to objects or things in themselves i.e. *de re*. Swinburne writes that: '…the truth or falsity of a metaphysically contingent sentence, however, depends not merely on what the sentence claims but on how the world is, independently of how we describe it' (Swinburne, 2013, p.19). No juggling with definitions, descriptions or verbal

usage can necessarily guarantee that some matter of fact is the case. There is no reason to suppose, that in the strong sense of 'necessity' objects picked out under certain concepts in a certain way must be constituted or behave in a particular way. However, an essentialist argument based on a theory of 'natural kinds' seems to offer a route to 'metaphysically necessary sentences'.

2.3 Naming and natural kinds.

In *Naming and Necessity* (1980) Kripke extends his theory of naming to 'natural kinds' that is to common nouns that refer to species or substances found in nature (tigers, lions, water, gold and so on). DRT includes the idea that natural kind terms are also therefore 'rigid designators'. Names pick out their referents directly (and necessarily); if we have established the 'essence' of a substance, say, 'Water is H_2O' then the 'picker' (a competent language user) can't be wrong. Once this is established the identity of water is necessarily fixed by its composition. Subsequently when we refer to 'water' we must always be referring to a substance whose chemical structure is H_2O. Other (phenomenological) properties of water, such as being colourless, odourless and tasteless, are inessential to its identity. Thus, we may conclude that there is no possible world in which something could be H_2O and not water i.e. this is metaphysically necessary. Swinburne seems to believe that the same argument can be made in relationship to artefacts (Swinburne, 2013, pp.30-31).

If, in a possible world, there is a liquid that is colourless, odourless and tasteless but has the composition XYZ then this, necessarily, is not water. Putnam in his famous 'Twin Earth' thought experiment (1995, pp.22-25) argues to this conclusion. Both Kripke and Putnam reject the Fregean view that sense determines reference. For Frege it would be the case that 'water' meant, as in ordinary discourse, 'a liquid with certain observable properties' i.e. odourless, transparent, colourless and so on. For Putnam and Kripke the meaning of 'water' is determined by what it 'is' (not a concept) and this 'is' is determined by its chemical composition. The referent term for water is then fixed when its chemical composition is established:

> There is a class of rigid designators, to which Kripke and Putnam drew our attention in the 1970s, which pick out things (and especially substances and kinds of substances) by certain of their superficial properties, but where what makes a substance or substances of that kind is the essence (of which competent language users may be ignorant) underlying those properties, a set of necessary and sufficient conditions for a substance or substance kind to be that substance or substance kind. In ignorance of the latter, we do not fully understand what we are saying about a substance when we say that it is that substance or a substance of that kind. (Swinburne, 2013, p.11)

The implication is, of course, that up until that point (say around 1750) when the word 'water' was used, people didn't 'really' know what they were talking about. But this seems an odd result of the analysis – in that before that period users got on perfectly well with the business of life referring to water without benefit of knowledge of its chemical composition. Similarly, people got on with transacting in gold, identifying it as yellow and malleable, without benefit of knowing that it ought to be identified by its atomic number, 79 (Grayling, 1990, p.183).

On the Putnam-Kripke account reference is operative not by meaning but by some historical linking of present uses of a term to the occasion when the referent of the term was fixed (Grayling, 1990, pp.183-184). Similarly, for Swinburne it is certain sufficient and necessary conditions which determine what something is – its chemical composition or atomic number. We can only identify what something is, and thereby what counts as sameness over time, by discovering its true nature or essence. Something may look like gold, such as iron pyrite, but not be gold. Phenomenological characteristics of natural kinds do not determine their true nature. This line of argument takes on its full significance in Swinburne's account of personal identity. The question 'Whether humans are the same substances as their bodies, and what makes a person at an earlier time the same person as a person at a later time?' might be answered by being able to 'pick out', by direct reference, their essential, rather than inessential, properties (See Chapter Six, 'Personal Identity').

Swinburne has already started by stipulating that the world may be categorized in terms of substances, properties and events. Kripke's theory provides the ideal means for specifying the identity criteria of substances because rigid designators appear to enable us to know which objects are the same and which

are different; objects in this case being the same substance, property, time and event. Once the ‘essence’ of natural kinds has been established then, on this theory of natural kinds, their ‘name’ is necessarily established. This is integral to Swinburne’s notion of ‘canonical vocabulary’ which, as we have seen, he believes, establishes the identity of mental substances and physical substances.

However, the problem with this ‘essentialist’ form of argument is brought out by an anecdote concerning the Count of Saxony and Marshal General of France, Maurice de Saxe (1696-1750). Saxe apparently asked anxiously ‘…But what guarantee have we that the planet we call Uranus really is Uranus’ (Flew, 1961, p.82). The reply to this must be that such a question makes sense only if we interpret it roughly as meaning ‘Which planet does everybody, and particularly the person who first named it, call Uranus?’ Although Saxe’s question is apparently absurd it is not altogether absurd to ask, as Flew suggests, ‘whether something which is universally accepted as right really is right’. There is a distinction here between questions of verbal usage and questions about a challenge to some accepted standard. In the next section we consider the argument that natural kinds are determined by their composition.

3 Objections to Natural Kinds.

3.1 Composition and natural kinds.

In the Twin Earth Argument, above, Putnam (1995, pp.22-25) asserts that something is water iff it has the chemical structure of water. In other words, suppose there is a liquid, say on a Twin Earth, which is indistinguishable from our water i.e. having similar phenomenal properties such as taste, colour, the ability to quench thirst, washing power and so on but has a different composition (XYZ). According to Putnam, this Twin Earth liquid would not be water. Natural kinds are not fixed in advance but depend on the objective laws which determine their constitution. Stroll (2000) argues that (a) Putnam is mistaken about this but (b) even if the proposition that ‘water is necessarily H_2O’

was not false the inference that ‘water must mean H_2O’ does not follow. He proposes three reasons why the composition argument is mistaken.

3.1.1 The Isotope Argument.

Water is composed not only of H_2O but also D_2O molecules (heavy Water) whose molecular weight is 20 compared with that of H_2O at 18. There exists a further isotope - Tritium, T_2O, which can equally be said to be part of the composition of water. And there are even more compositional complexities arising from oxygen combining to form other molecules of different atomic weight. The apparently simple identity claim that ‘Water is H_2O’ begins to look more complicated. The claim that there is some ‘essence’ which invariably differentiates water (in every possible world in which it exists) begins to appear to be more like an act of linguistic convention.

Putnam’s defence here might be that the term water refers to two different kinds of stuff with very different properties. For example, heavy water is toxic. ‘Water’ here is a homonym and fundamentally is not the same stuff having different chemical compositions. However, Stroll (2000, p.40) makes the further point that the essentialist theory, in common with many philosophical arguments, rests on too limited a set of select exemplars or paradigms (water, gold, tigers and so forth) which seem to suit the purpose but ultimately fail to deliver the required outcome.

3.1.2 The Functional Argument.

Stroll asserts that the meaning of a term has little to do with the composition of an item but depends more importantly on the function in serves in a language. This is illustrated by a variation in Putnam’s Twin Earth Scenario which shows that function and appearance can override composition. Suppose through some interactions between Twin Earth and Earth the inhabitants of Twin Earth discover that ‘water is H_2O’ on Earth and not XYZ. Stroll claims that it is entirely reasonable to suggest that Twin Earthers would not stop using the word ‘water’ because it is used for the same set of purposes on both planets. Furthermore, it

is implausible to suggest that Twin Earthers would not, as Swinburne (2013, p.19) claims, know what they were talking about.

There are cogent alternatives to the view that what is always decisive in naming a substance is composition or microstructure: '…it cannot be a necessary truth that water is identical with H_2O' (Stroll, 2000, p.241). The same point applies to human artefacts as much as to natural objects. It seems reasonable to say that what counts as a 'desk' is the use to which it is put not the material and properties from which it is composed. For example: (a) desks maybe made from many different types of materials; (b) compositionally, they have many different types of 'micro-structures'; (c) we subsume what 'desk' means not according to their varied composition or microstructures but by their use. More strongly we might urge, with Wittgenstein, that the word functions through its role in a particular 'language game' and not through some set of necessary and sufficient conditions enshrined in a definition (Wittgenstein, 1980, p.17).

3.1.3 The Isomer Argument.

The existence of isomers seems to provide a further counter instance (Stroll, 2013, pp.241-242). Isomers are in effect variants of substances that have the same chemical composition but that have different qualities (properties). An example of this phenomenon is ethyl alcohol and methyl ether. Both are composed of the same numbers of carbon, hydrogen and oxygen atoms but in radically different arrangements: C_2H_2OH and CH_2OCH_2 respectively.

The substances are pure substances and have different physical properties such as boiling and melting points. They are not 'inter-convertible' and there is no way of predicting these different properties from simple knowledge of their elementary chemical composition.

> …because they are pure substances we can use Kripke-Putnam vocabulary and call them natural kinds. The example shows that Putnam and Kripke are wrong in holding that natural kinds are identical with their chemical components since these are the same in all cases of isomers. (Stroll, 2000, p.242)

3.2 Meaning versus composition.

Stroll (2000) argues that the essentialist argument confuses two different questions. The relevant distinction is between asking, on the one hand, 'What does 'water' mean?' and, on the other, 'What is water?' The confusion lies in the assumption that a debate about what 'water' (or any other natural kind term) means will be settled once it is determined what water is i.e. what its composition is. But why should this be the case? Putnam asserts '…once we have discovered that water (in the actual world) is H_2O, nothing counts as a possible world in which water isn't H_2O' (Stroll, 2000, p.236). Therefore 'Water is H_2O' is an identity statement and 'is' here means 'identical with': Water is H_2O.

But Stroll suggests why this won't do. Consider:

> Water is H_2O;
>
> Ice is H_2O;
>
> Therefore: Water is Ice.

If A=B and B=C, then it follows that A=C. But the conclusion must be false because it is just not the case, empirically, that water is identical to ice. They have different properties i.e. ice is solid and opaque, water is liquid and transparent. Ice would not float in water if ice and water were identical as claimed by Putnam and Kripke. But, of course, Swinburne would want to say that opacity, transparency, solidity, liquidness are all inessential properties. This begs the question. Why should we assume that phenomenological, observable features are merely 'superficial'?

We can extend the argument if we say, 'Steam is H_2O'. The same logical problem arises. The point is the differences between ice, water and steam cannot simply be accounted for by determining chemical composition. The choice of composition as the 'essence' of a natural kind, to the neglect of observable physical differences constituted by various phenomenological properties, simply runs in the face of correct, everyday usage. The observable physical differences here are explained in terms of everyday locutions that ordinary, non-scientific, people have used historically for this purpose. But this is precisely what Swinburne sets aside as inadequate in his account of identity criteria for substances (Swinburne, 2013, pp.28). Why assume that

phenomenological, observable features are merely ‘superficial’? An essentialist response to Stroll’s argument might be not that the Kripke/Putnam theory of natural kinds is wrong but that the wrong property for the essence of water has been identified.

3.3 Learning about natural kinds.

Is it reasonable to suggest we can only ‘know’ what water means at the end of some protracted learning process when we have learnt its chemical composition? Wittgenstein, for example, offers a more plausible alternative account of how we might understand how a child comes to know what water is. This proceeds through coming to grasp the phenomenological characteristics of water - that is that it is liquid, fluid, transparent, odourless and potable etc. This is what water means and learning takes place through the use of the word in relationship to such properties and the relevant language game.

It could reasonably be argued that these are the essential properties of water. If so, then isn’t water a ‘natural kind’? And certainly, it is the case that it’s not purely arbitrary whether what’s in your glass is water or some other liquid, say gin, which shares some of the phenomenal properties of water. But the theory stipulates that there must be one overriding, primary property which is the essence of the substance and which is picked out by a rigid designator. Not only that, but the Kripke/Putnam claim is that the essential property comes with a sense of necessity. Whereas what we have sketched above is a sense that our grasp of what water is, is empirical and experiential. And what’s more no contradiction is involved in believing that some of its characteristics might have been other than they in fact are.

A child is able to ‘pick out’ and designate water by such characteristics. At some stage they might be taught, and come to know, that the chemical composition of water consists of hydrogen and oxygen although this is arrived at not by experience. On the Swinburne theory it is only then that they know what they are talking about and use ‘water’ as an informative designator. Similarly, Swinburne claims that:

> While being ‘water’ (as used in the eighteenth century) is an uninformative designator of a property, being H_2O (as used today) is – I believe – an informative designator of a property. Being H_2O

> is the property of being composed of molecules, each consisting of two atoms of hydrogen and one atom of oxygen. (Swinburne, 2013, p.13)

These claims seem odd in a number of ways.

First, we have already seen that the compositional story of natural kinds, such as water, is mistaken. If we take Swinburne's claim literally, then there is a paradox of just how English speakers were supposed to be communicating with each other if they could, according to Swinburne, only exchange uninformative designators that didn't tell (inform) them what water was and meant. How would they have understood 'Bring me a glass of water'; 'Don't spill the water!' if to know these things involves knowing that water consist of hydrogen and oxygen?

Secondly, are they 'uninformative' in some highly technical sense that they didn't designate 'in every possible world' where water exists? But even so if the designator was genuinely uninformative how would I know whether we were talking about water or not? If an uninformative designator does allow us to identify and talk about water and communicate sensibly, then what is added if philosophers then want to bring in some extra notion of an informative [rigid] designator? The move is meant to show that we can derive metaphysically necessary (possible or impossible) sentences in a purely *a priori* way (Swinburne, 2013, p.19).

Stroll observes theoreticians:

> …instead of beginning with the fact that early English speakers communicated with one another and asking how that is possible, they have developed an *a priori* theory that makes the fact of such communication inexplicable. Here we have philosophical paradox in its strongest form. (Stroll, 2000, p.244)

Thirdly, and by analogy, children are in the position of our early English speakers. But nevertheless, by observation, experience, and manipulation they learn to use the word 'water' and communicate with others. The notion that we can only 'know' what water is when we know its chemical composition, or in Swinburne's terminology when we know its essential properties, is specious.

3.4 Fundamental objections.

There is no good reason to suppose that there are such underlying properties that necessarily attach to objects and determine their identity. There is nothing, of necessity, which explains the behaviour or functions of individual natural objects or human artefacts. Stroll (2000, p.234) argues that statements about chemical composition, for example, are neither necessary nor are they actually identity statements. It is a mistake to assume that a natural kind can be encapsulated, expressed, by a simple identity statement.

O'Hear (1991, p.58) provides a number of reasons to show that it is a mistaken assumption: '…that identifying an object in a certain way entails that it *has* to behave or be constituted in a particular way'. First, we do have well-founded beliefs, as Kripke, Putnam and Swinburne suggest, that we can categorize and pick out natural objects. We have 'sortal concepts' such as whales, roses, gold, mountains and so forth by which to pick out objects and which have certain resemblances to each other.

> In ordinary talk it is often fairly vague just which properties are the essential properties of the kind to which a particular substance belongs, but clearly there are some properties which any substance shares with other substances, such that if it lost them it would cease to be the same kind of thing and cease to exist. (Swinburne, 2013, p.29)

But resemblance is not necessity. Why should these resemblances involve a necessity which we don't find generally in our scientific investigations and understandings of the world? For example, if we think about Karl Popper's famous 'all swans are white' case there was no necessity that a black swan might not be discovered. Well you might argue that colour, in this case, is not an essential property. But how are essential properties then to be established? As we saw above in the case of water and H_2O there is no (logical) necessity involved. There can be in principle no properties which can be picked out as necessarily essential. It must be the case that all properties are contingent.

O'Hear (1991, p.58) illustrates this by considering individuating artefacts:

> Individuating an object as say a car, may lead us to well-founded expectations about its origins, performance and structure, but no more necessity can be milked out of this concerning these

> expectations than is given by the evidence from past experience that things that look like this generally behave like that. The problem for anyone who wants to derive a type of necessity concerning the properties or constitution of individuals individuated under particular sortal concepts is to explain how this necessity can amount to more than saying that in the past similar looking things tended to behave like that.

This implies that it is not possible to show that contingent facts are necessary truths. So that it cannot be (logically) possible that, given some 'real' definition of a word for that thing, its essential behavioural properties must necessarily follow. It is a kind of metaphysical dream that if only one could know the essential natures of things we would discover the ultimate reasons for why they behave as they do (Flew, 1961, p.115). In the next section I want to unpick the theory of direct reference and the belief in the value of informative designators in the light of the problems and criticisms of the concept of natural kinds discussed above.

4 Informative Designators Revisited.

4.1 The problem of transience.

Grayling (1990, p.185) says two related questions ought to be asked about direct reference theories. First, just what precisely constitutes the reference of a rigid designator? For example, for Swinburne the word 'red' is an informative designator. But there are many shades and hues of red, for example, Scarlet, Cochineal, Vermilion, Rossa Corsa, Hematite, Madder, Dragon's Blood and so on (Clair, 2016, pp.134-155). Secondly, what is the status of other referring terms such as those that may be used to pick out non-natural kinds? An informative [rigid] designator is defined as a word that designates the same item in every possible world in which it exists. But what is an item? Swinburne's answer to the question casts the net very widely:

> A rigid designator is a word which 'in every possible world, designates the same object', and I shall understand 'the same

> object' (or 'thing' to mean the 'the same substance, property, time, or event'); that is, designates the same object, whatever happens to that object so long as it exists…A non-rigid designator is a word that applies to something only as long as it has some non-essential property. (Swinburne, 2013, pp.10-11)

An informative [rigid] designator is supposed to be a word which 'in every possible world, designates the same object' i.e. has the same extension. But what is this object or item? Grayling points out that this cannot be the extension of the term since the very concept of a 'possible world' is predicated on the point that extension may vary across worlds (Grayling, 1990, p.69). Suppose there is a possible world in which Aristotle is not a teacher, then the extension 'teacher' is different in that world from the world in which Aristotle teaches. And, of course, teacher is not a natural kind term like water or gold and it is not clear whether it does or does not rigidly designate (a problem with Swinburne's extended use of the theory). The result is that it is possible that there are two possible worlds in which there is a disjunction between the extension of a given term. Grayling concludes that the extension of a term cannot be what the term designates.

A way out of this problem is to claim that an informative designator designates the kind or species itself. The downside of this strategy is that it confers rigidity on non-natural kinds that is terms such as 'teacher', 'bachelor', and so on. So, the term teacher would pick out all and only those who imparted information to, demonstrated to, or instructed others in a known educational way (Grayling 1990, p.185).

Secondly, Putnam and Swinburne take the view that all such kind terms are informative [rigid] designators. But there is a price to be paid for the expansion of what counts as an item that can be rigidly designated. Primarily this runs dead against our ordinary ways of talking. It is simply not the case that when we use natural kind terms we conventionally have a *kind* thing in view. Commonly we apply the label if anything fits with minimum sufficiency some general specifications. Thus, to name a non-natural *kind* is just saying what it is. This contrasts with the procedure for natural kinds, such as animal species or elements, when we can first name a natural kind and then go on to discover what it is. The central idea here is that naming and designating avoided

description and supposedly provided direct reference but for non-natural kinds description re-appears;

> …The descriptive theory appears therefore to apply here, for non-natural kinds lack biological or atomic essences, and appear to have what Locke called 'nominal' essences only. (Grayling, 1990, p.185)

Thirdly, DRT attempts to discover only one relationship, one royal road, to explaining how names refer. There are other relationships, for example, contrast 'mentioning X versus referring to X; picking out X versus identifying X and so on. For example, I may be able to pick out persons, say males from females in a group – without being able to identify any particular individual. Swinburne might then want to say that this is informative but not rigid but what sense does that make? Swinburne's attempt to describe the world as if it were a process of naming is surely ultimately too limiting. Why should we accept the central thrust of Swinburne's argument that words are 'proper names' (*unum nomen unum nominatum*)? As Austin asks rhetorically:

> But why, if one 'identical word' is used, *must* there be 'one identical' object present which it denotes. Why should it not be the whole function of a word to denote many things? Why should not words be by nature 'general'? However, it is in any case simply false that we use the *same* name for different things: 'grey' and 'grey' are *not* the same, they are two similar symbols (tokens) just as the things denoted by 'this' and by 'that' are similar things. In this matter, the 'words' are in a position analogous to that of the objects denoted by them. (Austin, 1970, p.39)

Fourthly, consider the case of the Mandela Particle. This was a new fundamental particle discovered in 1973 by the Leeds University Cosmic Ray Group. The particle was said to be 40 to 70 times the mass of a proton. 'The particle fitted neatly with theoretical proposals that a particle of that mass should indeed exist – the intermediate vector boson which would mediate the weak interactions' (Anon., 1975, p.310). Measurements made by a research group at Durham were subsequently unable to replicate the findings of the Leed's team. Eventually the emerging consensus was that the 'discovery' of the Mandela Particle was a mistake caused by faulty equipment.

‘Mandela particle’ is clearly a proper name and not a description. It fulfils all the criteria for a rigid designator: the name gives a direct reference to an ‘object’; the name was given in an act of ‘baptism’ at its discovery; the Mandela Particle is a fundamental particle of mass 70 times that of a proton’ and (like ‘water is H_2O’), according to the theory, is a true identity statement, necessary, and *a posteriori*. But as a matter of fact, the particle was subsequently discovered not to exist. Therefore, we appear to have an informative [rigid] designator which references something which does not exist.

What’s wrong with some objects being fictional in some worlds but not others? The objection might be that there is an object that exists in many possible worlds but not in the actual world. The object could theoretically be identified across these possible worlds with this same name. The name is then a rigid designator because it refers to the same object in these (but not every!) possible world in which it exists. But the non-existence of the object in the actual world doesn’t seem to make it incompatible with the theory. If this is right, then the case of the Mandela Particle could refer to a particle in whichever world it exists because it has certain essential properties. It just so happens that our actual world is one in which the particle doesn’t exist.

But it is incompatible because, at least in Swinburne’s interpretation, an informative designator is a word which ‘in *every* possible world, designates the same object’. We must read this as it can’t be an informative designator if it only designates in *some* possible worlds. This is part of the problem with the theory in that it wants to imply actual existence in all possible worlds. If informative designators can flop in and out of existence, of being rigid and then non-rigid (flaccid?), informative or uninformative, just how do we know when we are actually using an informative designator? The element of necessity seems to evaporate. So, in the next section I want to consider the problem of the treatment of fictive entities as a major weakness of the theory (O’Hear, 1991, p.168).

4.2 The problem of fictive designators.

Fictive names refer to things that do not exist; named fictional characters cannot exist in this or any other world because they are, by definition, fictional. Grammatically, fictive names are proper names. It seems perfectly acceptable to believe that ‘Hamlet’ designates a character in the play of the same name! However, here the principle of informative designation breaks down because we can’t press designation into the game of distinguishing things that exist from those which don’t. We might accept that they help to define the criteria for defining a concept, say of ‘mental substance’, but clarifying a concept is very far from showing that there is such a thing. We have to have warrants for believing something exists independently of simply knowing a name. In the cases of Swinburne’s key concepts, mental substance, property, and event and so on there is by no means any consensus that they name actual existing entities.

Fictive names create a dilemma. The first horn of the dilemma is that the theory implies that a designator cannot refer unless there is a particular substance or that it ‘labels’, ‘tags’ or ‘designates’ (all synonymous). The second horn of the dilemma is that we use proper names that do actually *refer* to fictive characters in drama and literature; non-existent substance. It makes perfect sense to talk about, and use proper names, in referring to characters and kinds of things which don’t exist (unicorns). If we accept the first sentence of the dilemma then we must deny that we can informatively designate non-existent objects or individuals. If we accept the second horn of the dilemma then we must reject the theory.

The first horn of this dilemma may be challenged by suggesting that the theory doesn’t require the actual existence of objects and individuals to talk about them. This is because, if we accept the notion of informative designators, we seem to be accepting that we can, and do, talk about objects and individuals in non-actual but possible worlds. Then the ‘particular individual’ mentioned in the first sentence of the dilemma needn’t necessarily be an actually existing individual. It could be an object or individual that might exist in a non-actual possible world. Thus, the theory doesn’t seem to require the actual existence of such individuals.

However, it seems clear that this is not how the notion of informative designation is to be interpreted. Swinburne is clear the object rigidly designated must exist:

> A rigid designator is a word which in 'every possible world, designates the same object', and I shall understand 'the same object' (or thing') to mean 'the same substance property, time or event'; that is, designates the same object, whatever happens to that object *so long as it exists* [italics mine]. (By 'every possible world' is meant 'whatever else might be the case'.) (Swinburne, 2013, p.10)

What can we make of the names of fictional characters? We cannot assume that an informative designator picks out something that necessarily exists in the way that Kripke and Swinburne seem to suggest. The theory suggests: (a) One cannot refer to that which does not exist; and (b) whatever is referred to must exist. But this is just false;

> …For it is a plain fact that we do use language to refer to non-existent (including fictive) objects by name and to make true (or sometimes false) statements about such objects. (Stroll, 2000, p.226)

Just as it is clearly true to say that Othello was married to Desdemona, but that Hamlet was not married to Ophelia. Surely to distinguish what is a fictive item from what is an actual item in the world cannot be determined by designation but again must (logically) depend on conditions of verification or falsification and not linguistically generated identity conditions.

Kripke has attempted to address the problem of fictive proper names and (non-existent) natural kinds on a number of occasions - for example in the Addenda to *Naming and Necessity* (1980, pp.156-157). However, none of the explanations Kripke admits seem plausible. For example, one response to problem of fictive species, such as unicorns, could be that we might discover in the past that there existed some animal that had unicorn like characteristics. But given the mythic status of unicorns and the lack of information about their internal structure '…there is no actual or possible species of which we could say that it would have been a species of unicorn' (1980, p.157). Similarly, in response to the problem regarding fictional proper names it might be claimed that we might discover that there was a detective with adventures like those of Sherlock Holmes. But this would not show that Doyle was writing about this

person. The possibility of coincidental likenesses of real to fictional characters is deeply implausible. Kripke writes '…I thus no longer write, as I once did, that "Holmes does not exist, but in other states of affairs, he would have existed"' (1980, p.158). Fictional entities and empty names remain a real problem for DRT. It is a topic that Swinburne does not address.

Swinburne acknowledges that:

> …you can tell the history of the world in many different ways (using different categories of kinds of substances, properties, and times, and so of events), subject to the condition that the system of categories would enable you to list a subset of events which entail all events. (Swinburne, 2013, p.9)

But no such list of categories can be generated by a theory of informative [rigid] designators. This must be as true for mental substances, properties and events as it is for 'natural kinds' and artefacts. A complete history of the world must surely include all of fictive events as well as actual events and all the possible interpretations of such events. For whatever can be said about non-fictional objects can be said about fictional objects; and whatever can be said in our daily talk outside fiction can be said within or about fiction. For example, we can state counter-factuals about fictional characters. We might propose or imagine that David Copperfield was in fact female. Such talk includes every possible use of ordinary discourse: jokes, lies, true and false statements, direct and indirect references, asides and the application of names to characters, places and things (Stroll, 2000, p.227).

5 Conclusion.

I have argued in this chapter that Swinburne's theory of informative designation suffers from perhaps insurmountable logical problems. The conclusion of Stroll, O'Hear, and others is that the theory is deeply flawed. The idea of 'natural kinds' and the claim that the essence of an object (broadly interpreted) can be identified on the basis composition have been shown to be false. There are no essences in the sense that Swinburne needs to support the project of a canonical vocabulary which would '…tell the whole history of the world'. We

have seen that the theory of informative [rigid] designation presents a picture of the way language works which is greatly oversimplified if not erroneous. The case is not made for a philosophical vocabulary which, Swinburne believes, is needed to correct the vagueness of ordinary language. As J.L. Austin writes:

> …our common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth making, in the lifetimes of many generations; these surely are likely to be more sound, since they have stood up to the long test of the survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than you or I are likely to think up in our arm-chairs of an afternoon – the most favoured alternative method. (Austin, 1970, p.182)

Chapter Four: Privileged Access.

1 Introduction.

In the previous chapter we examined the way in which Swinburne uses a particular form of reference theory to underpin a two-substance ontology. By 'tidying-up' the vagueness and ambiguity of language and with the determination of a 'canonical vocabulary' we might then be able to '…tell the whole history of the world'. The metacriterion, as far as Swinburne is concerned, specifies the nature and identity conditions of substances, properties, and events both mental and physical. In this chapter I examine Swinburne's concept of 'privileged access' and its role in answering the question 'What is the essential difference between mental and physical?' According to Swinburne the difference is that subjects alone have direct access to their mental properties and events. What characterizes my sensations, beliefs, feelings, intentions and desires is that they are 'pure mental events' to which I have unique, privileged access. The corollary is, of course, that such access is logically denied to public or third person scrutiny. The aim of this chapter is to evaluate Swinburne's privileged access argument and its implications. We begin with a number of alternative definitions of privileged access and then come to Swinburne's preferred definition.

2 What is 'privileged access'?

2.1 Definition.

Swinburne acknowledges that there may be other, and perhaps equally valid, ways of understanding the contrast between mental and physical. In particular he cites two: (a) 'the propositional/non-propositional' (or 'intentional/non-

intentional') and (b) 'the non-physical science/physical science' distinction. In the first case '…a mental property is one which involves an attitude towards some (apparently) conceivable event under a certain way of thinking about it. A physical property is then any property whose instantiation does not involve such an attitude' (Swinburne, 2013, p.67 fn.). Swinburne's objection to this propositional/non-propositional contrast is that pains and sensations of patterns of colour, for example, would not count as mental events because they are what they are independently of the subject's attitude to them. But equally these mental events are '…obvious trouble-makers for "mind-brain" identity'.

In the second case, 'the non-physical science/physical science' distinction, he understands this as the claim that the instantiation of a physical property is to be explained by 'extended physics' whereas the instantiation of a mental property is one which cannot be explained in this way. Swinburne finds this basis for the mental/physical distinction unsatisfactory because it is 'hopelessly vague'. It is unclear as to what would constitute 'extended physics' that is to say a science which incorporates present-day physics but is extended to much else as well but is still 'physics'. His doubts about these two alternative conceptions underpin his preference for characterizing the distinction by the presence or absence of privileged access (Swinburne, 2013, p.67 fn.).

Swinburne identifies the distinction between physical and mental properties in the following way:

> I define a mental property as one to whose instantiation in it a substance necessarily has privileged access on all occasions of its instantiation, and a physical property as one to whose instantiation in it a substance necessarily has no privileged access on any occasion of its instantiation. (Swinburne, 2013, pp.67-68)

The very mark of a mental property is that a subject always has privileged access to that property whenever it is instantiated. My experience of something which is red is something which I have unique access to. I have this access to this red image, which others don't have, precisely because it is instantiated in me. Others can only infer that I am 'seeing red'.

Flew defines it more straightforwardly as '…the particular relationship we have to contents of our own consciousness, but that none of us has to the contents of anybody else's' (1979, p.287). I alone can be conscious of my mental events.

Others need to infer what is in my mind from my physical and linguistic behaviour; I can know these things directly and with certainty (other things being equal, for example, I am not under the influence of drugs). The private realm of mental properties and events is directly known to the person whose properties and events they are. However, to believe this seems to give rise to the particular problem of 'other minds' – for how do we 'really' know that other minds have similar experiences? But this must only be the case if in this context 'to know' is interpreted as 'to know directly'.

2.2 The argument.

Swinburne elaborates:

> Someone has privileged access to whether a property P is instantiated in him in the sense that whatever ways others have of finding out, it is logically possible that he can use, but he has a further way (by experiencing it) which it is not logically possible that others can use. A pure mental property may then be defined as one whose instantiation in a substance does not entail the instantiation of any metaphysically contingent physical property in that substance. (Swinburne, 2013, p.68)

Now if we are meant, as perhaps I think we are, to understand 'access' as meaning that only I can *experience* my thoughts and only you can *experience* your thoughts then this seems relatively unexceptional. I might know because, say you are a close friend, what you are thinking but this must only be 'indirectly'. So, I can equally well know that Tom is angry without being able to directly introspect his feelings. J.L. Austin (1970, p.116) concludes his influential paper on 'Other Minds' with the gist of his argument in the following way:

> *(1) Of course* I *don't* introspect Tom's feelings (we should be in a pretty pickle if I did).
>
> *(2) Of course* I *do* sometimes know Tom is angry. Hence
>
> *(3)* to suppose that the question 'How do I know that Tom is angry?' is meant to mean 'How do I introspect Tom's feelings?' (because, as we know, that's the sort of thing that knowing is or ought to be), is simply barking our way up the wrong gum tree. (Austin, 1970, p.116)

However, I suggest that for Swinburne there is just more to it than this. First, it is not simply that he wants to draw attention to the fact that *my* introspections are *my* introspections and *your* introspections are *your* introspections and that I can't introspect your introspections. His claim is that this justifies a certain way, *his* way, of separating mental and physical. Thus 'my remembering an appointment' is interpreted on these grounds, as a mental event which is the instantiating of a property in a mental substance. But how do we get to this from the unexceptional claim that you cannot experience my experiences? Yes, it would be foolish to deny the case that we cannot experience others' experience, but Swinburne's further claims do not follow straightforwardly from these simple observations about introspection. We are, it seems, on his account required to accept that the physical/mental distinction implies already the existence of mental substances and so on. Again, we might be prepared to accept the idea of substance in relation to physical objects, but it is far from clear, at least on these grounds, that we are required to accept the existence of 'mental substances' on the basis of a distinction between public and private access to our mental phenomena.

Consider Swinburne's answer to the question 'What is a "mental substance"?' This seems unpersuasive if not doggedly circular:

> I define a mental substance as one for which the possession of some mental property is essential. For a mental substance to exist, when not having a conscious event, it must have some beliefs or desires, or at least a disposition to have sensations or thoughts to form intentions. (Swinburne, 2013, p.141)

The problem, as always with such exotic entities as mental properties and mental substances, is just how we might first identify them, and then subsequently re-identify them, given the absence of any physical substratum. But surely when we talk of beliefs, desires, sensations, thoughts and intentions these are characteristics (predicates) which we normally ascribe to flesh and blood persons such as ourselves?

Secondly, in the previous chapter we discussed some of the problematic aspects of Swinburne's attachment to the theory of 'informative designators'. The theory is again in play here. When Swinburne mentions a property or event we are asked to assume that he is naming a property or event picked out by the informative designator used in mentioning it. For example, when he mentions

the property of ‘having a toothache’ he is, as a consequence, naming the property picked out by the informative designator ‘having a toothache’. Isn’t there something odd here? I might say ‘there has just taken place in me a mental event ‘remembering an appointment’ picked out by the informative designator ‘remembering an appointment’’ but we might say this means nothing more than ‘I’ve just remembered an appointment’? (Wittgenstein, 1978, PI I 306).

Swinburne’s use of privileged access as a justification for the mental/physical distinction is at least problematic. Perhaps there is a distinction to be made here between ‘privileged access’ in a relatively ‘soft’ sense – as interpreted by Austin - and a much ‘harder’ sense used by Swinburne. In the softer sense reference to ‘privileged access’ does not entail the radical disjunction between mental and physical or private and public. The danger in this ‘hard sense’ of access is that it conjures up an inappropriate model of an inner observer ‘picking out’ and naming inner, mental objects versus a public world of observers similarly picking out and naming objects. This implies that privileged access is about special modes of acquainting ourselves with inner mental objects and events. But, as Heil argues (2013, p.84),

> Now if ‘directly observing sensation’ just amounts to *having* that sensation, there is no puzzle at all in the idea that you can directly observe your sensations. This is just to say that only you can have *your* sensations. And this is no more mysterious than the thought that only your refrigerator can undergo *its* defrosting.

Now this does not necessarily refute Swinburne’s account, but it does, at the very least, present an alternative picture of what self-awareness might mean. However, as we have seen it is possible to account for the asymmetry between public and private access without necessary recourse to dualism. In the next section I argue that Swinburne’s model of inner perception (and naming) of mental events leads to a *reductio ad absurdum* which further calls into question Swinburne’s sense of privileged access.

3 Perception of Mental Events.

3.1 Two types of mental event.

Swinburne distinguishes two kinds of conscious mental events; one type involves physical properties and a second type which has no physical properties. First, 'impure mental properties' entail the instantiation of a physical property that is to say: '…I can only see a desk or hear a telephone ring if I have the physical properties of being causally affected by desk or a telephone' (Swinburne, 2013, p.69). Secondly, pure mental properties are entirely removed from such physical dependency. Examples given by Swinburne include the properties informatively designated by say 'having a headache' or 'having ten spots in my visual field' or 'hearing a tune in one's mind'. He claims that it is just *obvious* that these are pure mental properties and moreover it is *obviously* logically possible that someone can 'have a headache' or 'spots in their visual field' without this being entailed by any brain or any other physical event. (But who or what 'has a headache', or 'has ten spots in their visual field' or is 'hearing a tune in their mind' – if no brain or physical events are entailed?)

What 'pure' and 'impure' mental properties and events have in common, according to this model, is a similar process of privileged perception and designation of properties and events whether physical or mental. But is this plausible?

Firstly, and uncontentiously, suppose I see an over-ripe banana that looks brownish yellow, squashy, elongated and so on. The way it appears to me depends on my visual experience (of a physical substance let's concede). It is uncontentious that perception involves sense experiences. I see the banana.

Secondly, Swinburne's argument is that these experiences are mental events albeit impure. In Swinburne's language I would be having an impure mental event involving the instantiation of certain properties. Self-awareness resembles awareness of ('external') physical objects and events but 'turned inwards' such as in the case of consciousness of 'having a headache'. I observe myself 'having a headache'. In other words, the distinctive element is that such

experiences are mental events '…which we are aware in having them and which have just the properties which we are thus aware of them having' (Baldwin, 2001, p.46). It is the notion of 'turning inwards' or 'having inner perceptions' that is problematic.

Thirdly, Ryle (1976, p.203) argues, this analysis must lead to absurdity. In effect it results in a regress of experiences or perceptions. According to the theory the subject's experiences are in themselves mental events '…of which we are aware in having them and which have just the properties which we are thus aware of them having' (Baldwin, 2001, p.46). For example, having a visual experience, on this account, implies an awareness of the experience itself. Now if we are observing (and informatively designating) this itself is a form of awareness, a special 'inner perception'. But because all perceptions must imply experiences, our inner perception of our own visual experience (observation) must lead to a further non-visual experience and so on. And this leads to the regress.

The regress undermines the claim that privileged access must involve a model of 'inner' perception. It implies that there is no symmetry between perceptions or observations of mental objects and of 'outer' physical objects. Thus, my mental process of remembering, for example, is not an object or event that I inwardly perceive. It just is that my having it *is* my sensing it (Heil, 2013, p.243).

3.2 'Hearing a tune in the mind'.

Let's consider Swinburne's example of 'hearing a tune in one's mind' (2013, p.69) he asserts these are pure mental properties instantiated in a mental substance. It is then just *obvious* that it is *obviously* logically possible that someone can hear a tune in their mind without this being entailed by any brain or other physical event. It appears from the claim about privileged access, in the sense that I cannot introspect your experiences or you mine, we arrive at a picture of two separate and independent worlds (or one world which is physical and spatial and another 'world' which is non-spatial – the mind.

But what does it mean ‘to hear’ a tune in your mind. Ryle remarks that this cannot in any sense be a literal ‘hearing’ (ontologically equivalent, say, to ‘seeing a banana’). We are dealing here with a metaphor.

> When people employ the idiom ‘in the mind’, they are usually expressing over sophisticatedly what we ordinarily express by the less misleading metaphorical use of ‘in the head’. The phrase ‘in the mind’ can and should always be dispensed with. Its use habituates its employers to the view that minds are queer ‘places’ the occupants of which are special status phantasms. (Ryle, 1976, p.40)

It is important to be clear that Ryle is not claiming that ‘minds’ do not exist. The argument is about the existential status of minds. As already cited in the quotation above, his argument is that we are talking about two senses of ‘exists’. And what his analysis illustrates is that those different senses are rooted in language, in the logical grammar, we use to talk about mental and physical things. He argues that to talk of a person’s ‘mind’ is to talk of a person’s abilities, liabilities and inclinations to do and undergo certain sorts of things as part and parcel of the everyday world. In this sense, he like Swinburne, is arguing, that whereas my capacity to do mental arithmetic is quite a different kind of thing to my physical characteristics. For Ryle the antithesis between a public, physical world and the private, mental world is bogus because it just does rest on the misreading of the logical grammar of our everyday mental concepts:

> The sorts of things I can find out about myself are the same as the sorts of things that I can find out about other people and the methods of finding them out are pretty much the same. (Ryle, 1976, p.149)

And the evidence for this must be that we have established this through the logical grammar of our everyday utterances about minds and mental characteristics. Such utterances are ways of saying things about the abilities, affections, and dispositions and other characteristics of persons. The problem is that the architecture of Swinburne’s theory paradoxically, conflates mental and physical characteristics.

This is exemplified in the following case (Swinburne, 2013, p.70). Both Swinburne (when he is in pain) or a filing cabinet (which weighs ten kilos) have the property of 'being in pain' or the property of 'weighing ten kilos'. He has privileged access as to whether (or not) the property of 'being in pain' is instantiated in him and no-one has privileged access to whether the property of 'weighing ten kilos' is instantiated in him, it follows that he has privileged access to whether the property of 'being in pain' or 'weighing ten kilos' is instantiated in him. He goes on to say that a filing cabinet does not have privileged access to whether the property of 'being in pain' or 'weighing ten kilos' is instantiated in it for the luminous reason that a filing cabinet is not conscious and therefore cannot have privileged access to anything at all.

But the comparison is in this case false. Pure mental events are private events which are publicly unobservable according to him. I can know that he is in pain even if it is a truism to say that I cannot feel his pain. But I also know what it is to feel pain. How privileged is 'privileged' access? There is nothing in the notion that my experiences, being *my* experiences entails their incorrigibility. There is no logical contradiction in asserting that someone might fail to recognize his frame of mind for what it is. For example, they mistakenly suppose themselves to know things which are actually false; deceive themselves about their own motives; fail to notice a clock has stopped ticking; they do not know they are dreaming when they are dreaming, and sometimes they are not sure they are not dreaming. But it is equally the case that, as a matter of fact, we do deploy mental-conduct concepts regularly and effectively in describing and assessing others. If the privileged access account were true in its strong form, then this would not be possible.

Ryle (1976) asserts that this must be wrong because sensations and so on are neither observable nor unobservable because they are just not the kinds of things that can be observable or unobservable. To believe so is to accept the model Swinburne proposes. But you might say this must be false because I can observe (for example) the pain of a headache from which I'm suffering. Someone else could observe me clasping my hand to my brow or complaining of the pain – that is witness the 'pain behaviour' I am exhibiting. But what they cannot do is observe the pain directly in the way that I can. However, when I am in pain would I really be saying something like 'Oh I observe I am in pain'? Isn't

it that I just say, ‘I’m in pain’, ‘my arm is hurting’, and so on? I’m simply in pain – I don’t observe it in the way I might be said to observe a squashy brown and yellow banana. I express it. There isn’t an object ‘painful’ in the way there is an object ‘brown and yellow squashy banana’. In the next section we cast further doubts on the idea that privileged access means access to some inner, private self whose characteristics are impenetrable to the public world.

4 Private Self and Private Language.

4.1 Wittgenstein’s Private Language Argument (PLA).

Swinburne (2013, p.87) recognizes that the everyday words we use to discuss and describe the productions of private access are in fact public not private words; ‘mental words are words of a public language’. However, a corollary of the categorical divide he advocates is the sense of an isolated private self transparent to itself but utterly opaque to others. This is a core theme in Swinburne’s account of personal identity which pivots on the concept of a private self as a mental substance capable of being disembodied (See Chapter Six, ‘Personal Identity’).

If you say ‘I’m feeling queasy’ I would know what you mean and what you must be feeling in order to say such a thing. Given a public language then it seems perfectly clear that we can (directly) understand what someone means when they say, ‘I am in pain’. It would be odd to say that this was then something that was ‘inferred’. I don’t infer your queasiness or pain. Conversely the problem for a Cartesian model, as Swinburne, himself, observes, is: ‘How can a subject use words with a public meaning to describe what is, in an important respect, private to him or to herself? And how can others understand that subject’s description of events?

Wittgenstein famously rebuts this picture of the private self in the *Philosophical Investigations* (1978) first translated and published in 1953. His approach is to pose the question: ‘How could there be wholly private criteria to describe private sensations?’ If the privileged access story were true, then it ought to be possible

for our private self to construct a wholly private language. He tries to imagine what would be entailed in the construction of such a 'private language' that is a language which is to be used by only one person to report the on goings of their mental life. The meanings of the words of the language would be the product of the linguistic practices adopted by that person alone.

Wittgenstein supposes the person to be having some current sensation ('S'). Suppose they then have another sensation and want to know whether this is also an 'S'. In the normal circumstances of a public language such recognition would be part of the grammar of our known sensation words. In the private language scenario, no such guarantee is available to the lone inventor of a private language. No independent, public criteria are available to determine correct usage. (The counter argument is, as we discuss below, that I could check it via my memory.) Our private language user cannot know whether they have really identified the same sensation again. In the circumstances of a private language Wittgenstein argues we could never know whether or not we had succeeded in endowing our words with a private reference or any reference at all. The only possible criterion could be what the person thinks is right – but that is no criterion at all. They might think this second sensation is an 'S' but could be wrong about this but would never know they were wrong or – if right would not know that either. There simply would be no criteria of correctness to guide description or misdescription.

Wittgenstein's twist here is to deny that this is an epistemological problem. It is intrinsic to the logic of a private language. There are simply no rules or procedures attaching terms to intrinsically private entities: '…whatever is going to seem right to me is right. And that only means we can't talk about right'. Private language is not simply practically impossible but logically so. It is for this reason that the private language argument, if correct, is a powerful tool in de-reifying the conception of the mind as an occult arena populated by 'mental objects' and 'mental events'. Simon Blackburn believes this is a strong challenge to '…the Cartesian way of thinking of the mind as an inner theatre, whose show is known in a privileged and unique way by its possessor' (2005, p.293).

In summary:

P1. The meaning of words is a function of the rules of their use.

P2. In a private language rules of usage are rules that only I can follow.

P3. A rule that only I can follow is a rule that cannot sustain the contrast between obeying the rule and violating the rule.

P4. If there is no way to get it wrong, then there is no way to get it right.

C1. A private rule, therefore, involves a contradiction.

C2. Therefore, there can be, in principle, no private languages.

4.2 Swinburne's Reply.

Swinburne acknowledges the potential threat of the private language argument to the concept of privileged access. However, he believes that Wittgenstein is wrong in claiming that there can be no misdescription in a purely private language.

> A subject could misdescribe a sensation as 'E', meaning by that that it was qualitatively identical to the previous sensation so described, even if the subject had no way of discovering this was so or not. (Swinburne, 2013, p.88)

What are Swinburne's grounds for this claim? He challenges Wittgenstein's key assertion that in the context of a private language that there is no sense of right or wrong (P4): A rule that only I can follow is a rule that cannot sustain the contrast between obeying the rule and violating the rule. Swinburne sees this as a problem of unreliable memory since Wittgenstein remarks:

> Always get rid of the idea of a private object in this way: assume that it constantly changes, but that you do not notice the change because your memory constantly deceives you. (Wittgenstein, 1978, Pt. II, xi)

But this only means that the subject might be unable to know with some degree of reliability whether they had described their sensation, by the rules of their private language, correctly or incorrectly. This would imply that they may or may not have applied the 'correct' private language word to the sensation or not. In

consequence right and wrong descriptions are still notionally applicable even though they couldn't be certain.

In addition, it can be argued that the emphasis on the unreliability of memory is just too strong. Such unreliability would be equally applicable and undermining to public language. By the same logic would it not also be the case that I can't know whether or not any word that I use in everyday speech is used correctly or not? I might consistently misremember what correct usage is according to the public rules of the language.

Swinburne (2013, p.57) argues from a claim that 'apparent memories' may be assumed to be more or less reliable evidence for 'what happened'. He claims that: (a) apparent memory is convincing evidence of events having happened and strongly convinced evidence of apparent memory just is strong evidence of what has happened; (b) any apparent memory can be checked against another apparent memory. Swinburne claims, for example:

> An apparent memory that sensations of type E only occur after sensations of type F and an apparent memory of the subject that they had a recent sensation of type F would be evidence that the subject had described their present sensation as 'E' correctly, and an apparent memory of the subject that they had not had any recent sensations of type 'F' would be evidence that the subject had misdescribed his or her present sensation. (Swinburne, 2013, p.88)

This seems to miss the point of Wittgenstein's argument. Memory is of no help here. On what basis can I check my private apparent memory against another private apparent memory without ending up in some regress of referring to other private apparent memories. It is quite a different case if there a reliable check, an independent, public check, as evidence on the truth or otherwise of the apparent memory.

> What is definitionally essential to a logically public language is that the criteria determining the applicability of its expressions must be, in principle if not always in practice, publicly available. Correspondingly, it would be essential to a logically private language that it must be possible to explain the meaning of its terms by reference only to the subjective experience of the user. (Flew, 1994, p.314)

What Wittgenstein shows is that it is simply not possible in the case of a private language to meet, to use Flew's words, the 'definitionally essential' condition of explaining '… the meaning of its terms by reference only to the subjective experience of the user.'

There is simply no way in the private language case of reliably talking about or describing our experiences – memories included. Similarly, the same problem arises with a truly private mental object – sensation of type 'F'. How does it help to posit some other purely private object 'F' that occurs following 'E'? We have evidence that we have misdescribed 'F' because of the non-occurrence of 'E'. But the same problem infects both 'E' and 'F' if they are both to be identified in a purely private language. There needs to be some independent checkable rule that sensations of type E follow on from sensations of type F. Swinburne wants, I think, to suggest that 'apparent memory' can in a sense be both judge and jury. But it is precisely the absence of an independent check which is at the core of the private language argument. Wittgenstein's point is not one about memory, its reliability or otherwise, but about the logic of public language. In the following section we examine some convergences and divergences between Swinburne's and Wittgenstein theories. These further illuminate our central topic of privileged access.

5 A Reconciling Project?

5.1 A science of the mental.

In spite of his rejection of the private language argument Swinburne shares some aspects of Wittgenstein's approach to language. In particular he accepts the fundamentally public nature of language and language learning. He acknowledges that, despite his claims about apparent memory (the reliance on apparent memory of past mental events), the lack of a public test renders a subject at least 'uncertain' as to whether or not they have described their experience correctly…'And crucially, others could have no knowledge of the mental event the subject is having' … ruling out the possibility of 'a science of

the mental' (2013, p.88). It is not altogether clear what Swinburne means by 'a science of the mental' although he clearly believes that *Mind, Brain, and Free Will* is a contribution to such a putative science. He plainly doesn't identify 'mental science' with psychology, neuroscience, cognitive sciences and so on. And language is crucial to this.

> Since the informative designators of any physical properties are not logically equivalent to those of any mental properties (since there are different criteria for applying the designators), no mental property is identical to a physical property. The criteria for being in pain are not the same criteria for having some brain property, e.g. 'having one's 'c' fibres fire') or behaving in a certain way in response to a bodily stimulus (e.g. crying out when a needle is stuck into you). (Swinburne, 2013, p.69)

If this is the case, and in addition no mental properties supervene on physical ones, a 'science of the mental' also must rest on the private/public distinction. Such a science must be defined by the distinction between 'how the subject feels' and 'brain and behavioural events'. These ideas return to his fundamental concept of what is meant by 'privileged access' and it is epistemological:

> Someone has privileged access to whether property P is instantiated in him in the sense that whatever ways others have of finding this out, it is logically possible that he can use, but he has a further way (by expressing it) which is not logically possible that others can use. (Swinburne, 2013, p.68)

The criteria for 'being in pain' are 'how the subject feels' and are a matter for a putative science of the mental based on the privileged access of the subject. In contrast the criteria for brain and behavioural events are a matter of public access and are the object of, for example, neuroscience.

Swinburne believes that 'a science of the mental' must rest upon a model of a public language in order to understand the mental life of human beings.

> ….as Wittgenstein points out, we learn the meanings of words for describing our mental events by the public circumstances which normally cause them and the public behaviour which – we believe – they normally cause. (Swinburne, 2013, p.88)

For example:

> We come to understand what believing, intending, or desiring (the content of) some sentence to be true consists in by being taught the necessary connection of these propositional events with each other and the public behaviour to which they typically give rise. (Swinburne, 2013, p.89)

Swinburne should not be understood in using the term 'propositional event' to mean that utterances stand for propositional events. The term 'propositional event' is used precisely to mean 'utterances' themselves. And this is related to Wittgenstein's approach in the sense that their meaning is constituted by their relationship to each other and to public behaviour.

But Swinburne and Wittgenstein again part company in relation to Swinburne's adoption of a theory of informative designators. Wittgenstein emphasizes the idea of usage as a key to the meaning and development of language: meaning is what is given by explanations constituted by rules for the use of words. He specifically rejects the notion of 'naming' or ostensive definition as fundamental to language. In contrast, following Kripke, Swinburne adopts a philosophy of language precisely based on necessary links which are thought to obtain between names and objects.

If we take Swinburne's example of the sensation of 'pain' – when I say, 'I am in pain' I am making a necessary link (by informative designation) between the word 'pain' and a pure mental event, the sensation of pain. The mental event of sensing pain is the instantiation of a (mental) property in a (mental) substance. From Swinburne's point of view this is an epistemological issue in that I am the only person in a position to feel and name this particular pain as what it is. This again follows from the concept of privileged access and the idea of feelings as quasi-objects or events in mental space.

If we adopt Wittgenstein's (1978, PI I, 256, 257) alternative this is not an epistemological event, I am not reporting but expressing pain. When I say, 'I'm in pain' (and I am in pain) I am showing that I understand the meaning of the word and expressing it not naming it. On this account understanding just is the ability to using an expression appropriately. Such an ability or mastery is demonstrated: (a) in using the expression correctly; (b) in explaining what it means; (c) and in responding appropriately to its use. What makes it appropriate is that I am in pain (but I might be feigning). However, Swinburne

implies that privileged access implies incorrigibility – this is the import of theory of information designation.

5.2 Privacy reconsidered.

In the previous section we have seen that Swinburne maintains the view that '…each human is in a better position to know what are his or her own beliefs, intentions etc. than are others' (Swinburne, 2013, p.93). He believes that the public nature of language can be reconciled to the view that the mental is identified with the 'inner' that is with what is private and directly accessible to the subject – the proprietor of a particular mind. Private here can be interpreted as meaning 'hidden from others' (Stroll, 2000, p.137). Maybe I acquire public language and learn to name the mental objects and events. Is such a reconciling project plausible? In this section, therefore, we look in more detail into Wittgenstein's investigation of privacy and its relation to Swinburne's notion of the mental and privileged access. Is the idea that there are 'inner objects' or 'inner events' to name - cogent?

According to Anthony Kenny (1973, p.180) Wittgenstein believed that the idea of a private language rested on two mistakes. The first mistake relates to experience and is the false belief that experience is fundamentally private. The second mistake relates to language and is the false belief that bare ostensive definition can be sufficient to determine meaning. Both these beliefs, about experience and about language run counter to Swinburne's idea of privileged access and theory of designation.

5.3 Privacy as incommunicable.

What leads Wittgenstein to these conclusions about privacy and language? Why should we prefer them to Swinburne's? Anthony Kenny (1973, pp.178-202), argues that Wittgenstein develops an important distinction between, on the one hand, privacy in the sense of something *incommunicable* (an epistemological point) and on the other hand privacy in the sense of *ownership*

and possession (ultimately a grammatical point). He asks '…to what extent are my sensations [or other mental events] private?' To claim that pains are private is to say, as we have already seen, 'Only I can know whether I *really* am in pain'. Other people can only infer or surmise that I am in pain. Thus, privileged access implies: first, other people cannot 'know' that I am in pain; and secondly, I can know that I am in pain …because other people cannot know that I am in pain.

Wittgenstein attempts to falsify these implications by showing these are not products of psychology but rather of grammar (Wittgenstein, 1978, PI I, 251). They arise from confusion around language games involved in the use of mental concepts. We can contrast this approach with Swinburne's remarks about the construction of a 'science of the mental' as a kind of quasi-empirical inquiry (Swinburne, 2013, p.10). But the trick is to relocate the problems here from an empirical question to a philosophical question about what it makes sense to say.

The nub of Wittgenstein's case is that 'I am in pain' is not a declarative sentence, that is, it is not a name of a mental state or mental event (as Swinburne maintains). To use Swinburne's terminology – pain cannot be an informative designator because it is not a name. There is no inner object that the designator pain refers to. In that sense the question of 'knowing' I am in pain is not an epistemological question, 'I am in pain' is more like a cry of complaint (Wittgenstein, 1978, PI II, ix).

So, to return to the notion that privacy is essentially 'incommunicable': 'Other people cannot know that I am in pain'. Wittgenstein says that in one way this is simply false, and, in another way, it is nonsense. It is false in the obvious sense that others often 'know' when I am in pain. I hit my hand with a hammer and cry out. The claim that 'other people cannot know etc…' (or only infer it) implies 'a high redefinition' of what it is to 'know' something at all. It implies that it is only if I can introspect (mental) properties and events can I really be said to 'know' (See above, Section 2). For example, I can 'know' what my perception of the smell of coffee is because I introspect it. But you cannot know what coffee smells like to me – presumably because you cannot introspect my introspection.

Further Wittgenstein argues is that if we take the word ‘know’ to mean ‘to know in such a way that doubt is logically excluded’ the thesis is senseless (or in the form of a joke). There can be knowledge only where doubt is possible. Swinburne’s conception of privileged access certainly seems to carry with it just such a notion of certitude and, if not certitude, then ‘…each human is in a better position to know what are his or her own beliefs, intentions etc. than are others’. This must be the meaning of ‘privileged’, I can know indubitably what I am aware of thinking, believing, intending, desiring, sensing and so forth. But we ‘have’ pains rather than learn about pains. It would be odd to say that ‘I’ve just learned that I am in pain’. Doubt is not just impossible here but senseless.

Wittgenstein remarks that it is senseless to say, ‘I doubt whether I am in pain’ (Wittgenstein, 1978, PI, I, 247). Taken in this way we might want to say this is correct. But those, like Swinburne, who agree that because we cannot doubt our sensations (without counter-evidence), therefore we must ‘know’ them, are being misled by the form of the proposition. They are taking it as if it were a straightforward empirical proposition (Wittgenstein, 1978, PI, I, 251). ‘I know that I am in pain’ may be taken as a ‘grammatical proposition’.

5.4 Privacy as ownership.

The other sense of privacy which Wittgenstein brings into play is that of ownership: something is private to me if only I can have it (i.e. it is ‘inalienable’). To claim that pains are private in this second sense is certainly to mean another person can’t have my pains (Wittgenstein, 1978, PI, I, 253). It is to say something about what is possible in the real world captured by a grammatical connection between the personal pronoun ‘I’ and the possessive ‘my’. In this sense we express a grammatical fact. ‘My experience is my experience’; ‘my pains are my pains’. But what follows? We might then want to say that pains are indeed ‘inalienable’ but the upshot of this is not to make sensations any more mysteriously than overt behaviour as the following discussion indicates.

Let’s take the example of blushes and sneezes. These may equally be designated as ‘inalienable’. But this seems like a very tenuous sort of privacy –

although I guess Swinburne might classify these as impure mental events or perhaps even purely physiological although they are, it seems to me, *prima facie* sensations. Nothing seems particularly mysteriously private about these. We can agree that sensations 'are inalienable' but it does not seem to follow that this makes them any more private than (overt) behaviour.

Firstly, what does 'having' in the sense of owning mean in this context? How, for example, are different kinds of pain distinguished from one another – by locality, by intensity or similar characteristics?

> It is conceivable that I feel pain in a tooth in another man's mouth, and the man who says that he cannot feel the other's toothache is not denying *this*. The grammatical situation which we are in we shall only see clearly if we get familiar with idea of feeling pain in another person's body. For otherwise, in puzzling about this problem, we shall be liable to confuse our metaphysical proposition "I can't feel his pain" with the experiential proposition, "We can't have (haven't as a rule) pains in another person's tooth". (Wittgenstein, 1980, p.49)

As we have seen Swinburne's sense of privileged access means that the identity of my pain as a mental event is established by its instantiated in a particular (my) mental substance. This is equivalent to 'our metaphysical proposition'. Wittgenstein is arguing that it is entirely conceivable that we can in fact feel pains in another person's body. The meaning of ownership is generally taken to imply that a possession is transferrable, is alienable. In the metaphysical proposition the 'can't' is to be interpreted in the sense in which we might say an 'iron nail cannot scratch glass'. He goes on to suggest that the 'can't' in the experiential proposition might be re-interpreted as meaning 'doesn't'; 'an iron nail doesn't scratch glass'. He is making the point that in order to understand the conceivability of a person experiencing pain in another's body depends on the kinds of criteria for a pain being in a certain place. And the criteria for pains being in a certain place are entirely different from 'A has a gold tooth' (in her mouth).

Suppose Bill and I both feel sharp pain in the upper abdomen one hour after eating a vegetable curry which I have prepared. It is perfectly natural to say that we both feel the same pain. But you might want to argue that this is not strictly true in that the (same) pains are not felt literally in the same location but only in corresponding places in the same body.

Siamese Twins may feel the same pain in the place where they are conjoined. But again, you might want to press the point that pains are only specifically the same, but numerically distinct, because one pain is Tweedledum's and the other Tweedledee's. But this is to make the identity of the possessor a characteristic of the pain itself. If there are two possessors, there are two pains. This makes 'Only I can have my pains' a grammatical truth like 'If they are my pains I have them'. This is uninformative. 'Only I can have my x' is true not only of, for example of bank accounts, but of many things besides sensations (blushes and sneezes). The plausible conclusion is that pains may be inalienable; but this will not make sensations any more private than public behaviour.

Secondly, granted that if they are my pains, I must (grammatically) have them but does this exclude anyone else having them (simultaneously) which Swinburne's theory would seem to imply? In other words what is the criterion of identity of the possessor of a sensation? His answer is that the possessor of a pain is the person who gives it expression (Wittgenstein, 1978, PI, I 302). My pains are the pains I express or would express. Swinburne's answer here would be that identity is established by a mental event occurring in a particular mental substance in which this particular pain is instantiated. A property is instantiated and not 'expressed'.

We can keep pains, beliefs and so on secret – we may refrain from expressing a pain or stating a belief. Keeping something private in the sense of keeping it secret provides no grounds for calling an experience that is not kept secret a private experience. For example, if I itch but don't scratch (an act of self-restraint) or say 'I've got an itch' we may call that a private experience – but if I itch and then scratch why should we call the itch 'private'? If we consider 'private' in this way and pose the question: 'Are pains private experiences'? It is reasonable to reply – 'some are, and some are not'. And it does not follow that because some sensations are private in the sense of secret – that all experiences could be private experiences.

6 Conclusion.

In this chapter we have been concerned with Swinburne's concept of 'privileged access'. For Swinburne 'mind' is a mental substance entirely distinct from physical substance. This two-substance model identifies the mental (substance, properties and events) with what is inner and what is inner with what is private. This inner privacy is the 'site' of mental events and is directly accessible only to the 'proprietor' of a particular mind. Mental events are transparent to me but opaque to public inspection. I have access to my mental events in a special way not open to others; only I can experience my own thoughts, feelings etc. This is what is meant by 'privileged access'. Swinburne employs the contrast between privileged access of the subject with public access to support the distinction between the mental and the physical. This, he believes, is an essential move in establishing property and event dualism.

In trying to understand the meaning of 'privileged access' we have considered two alternative interpretations those of Ryle and Wittgenstein. They both aim to dispel the concept of privileged access (and by implication Cartesian dualism). If the notion of privileged access, as conceived by Swinburne, is unsound then it must weaken claims to the kind of substance dualism which he advocates. Ryle aims to dissolve the antithesis between the public and the private worlds by showing that the opposition rests on various category mistakes and the misreading of the logical grammar of our everyday mental concepts. This is revealed by the logically absurd corollaries which follow from such mistakes and misreading. According to Ryle this has the effect of showing the correct logic of mental concepts. This can be shown, in Swinburne's case, paradoxically by his tendency to observe a symmetry between the world of physical substances, properties and events and that of mental substances, properties and events as if they had the same logical status. Ryle claims to show that minds are not any kind of substance nor are they occult repositories to which I have privileged access and in which mental properties and events have their being. This is just to misunderstand the words we use in describing people's dispositions and capacities.

Wittgenstein similarly dispels privileged access as something opaque and forever beyond the range of public scrutiny. The use and misuse of language to describe mental processes is also central to his criticisms. In particular the well-known ‘private language argument’ bears down upon how we might understand the limitations of privileged access. Mental events (Swinburne’s vocabulary) are private in the very particular sense of being accessible and therefore knowable directly, only to the person who has them. The challenge of Wittgenstein’s approach is that it removes anything that we might claim privileged access to. The private/public divide is dissolved partly because anything we express must be expressed in a common, public language. Private mental objects must drop out of consideration as irrelevant. This is reinforced by rigorous analysis of the ways in which we express ourselves about our mental lives. If nothing else Ryle and Wittgenstein’s interpretations cast doubt on the fundamental schism that Swinburne wishes to enforce between the physical and the mental.

Chapter Five: Credulity and Testimony.

1 Introduction.

The previous two chapters discussed the coherence, or otherwise, of key elements in Swinburne's argument for substance dualism; a metacriterion and privileged access. These ideas underpin his claim that if we are to give a complete account of the world then we need to postulate two general types of property (physical and nonphysical or mental). Such properties must be instantiated in two radically different types of substance (physical and nonphysical or mental). He asserts, first, that most conscious events are caused by brain events which are frequently in their turn caused by exogenous physical events or events endogenous to a person's body. Secondly, conscious events cause our brain events which in turn cause our overt behaviour. But how can such radically different substances interact? We plainly have a sense of the nature of physical causation but how is it possible that immaterial, non-extended substances can have causal efficacy? To answer these questions Swinburne must demonstrate that his version of substance dualism is interactive. The principles of credulity and testimony are key components in defending interactive dualism against various counter-arguments to such interactivity.

Levin, in *Metaphysics and the Mind-Body Problem* (1979, p.80) states a standard objection when he says that substance dualism '...is prima facie incompatible with, or renders unintelligible, a wide range of familiar facts. If I bang my toe, I feel pain; if I decide to take a walk, my body moves. How are things as dissimilar as minds and bodies capable of interacting?' There is, of course, no inherent problem in dissimilar things being able to causally interact. Igniting petrol in an engine's cylinder, to cause a controlled explosion, is not 'like' or 'similar' to the explosion it causes in any ordinary sense. But our paradigm cases of causal interactions are normally understood in physical terms. The interaction of Swinburne's mental and physical substances seems of an entirely different order. It is difficult to see what can be meant by the interaction of physical and nonphysical substances.

Paradoxically Swinburne cites similar types of ‘familiar facts’ as Levin to argue the contrary case. Interactionism is probably true because we just do experience ourselves as interactive mental substances: ‘the normal view’ of body/mind relationships. The principles of credulity and of testimony have principal roles in validating the normal view. This chapter will show how these principles are deployed to meet challenges to Swinburne’s belief in the causal efficacy of mental substances. Substance dualism is probably true because, he believes, there is no sufficient counter-evidence strong enough to show that it is probably not true. Table Two, Counter-evidence to Interactionism (below), summarises the challenges to be discussed in following sections of the chapter. We finally consider the question ‘Whether the principles of credulity and testimony actually deliver the epistemic goods Swinburne believes they do’? But first we turn to an exposition of the principles themselves.

Table Two: Counter-evidence to Interactionism (Swinburne, 2013, pp.100-125).

Types of Counter-Evidence	Evidence CCP	Basis	Implications	Swinburne's Refutation
a priori	Physical Principles.	Rules out anything but physical transactions and causality.	Causal efficacy of mental events ruled out by physical laws.	Principle of credulity and the argument to 'the normal view'.
a posteriori	α-evidence Neuroscience Principles.	Neuroscience Experiments Wegner (2002) Libet (2004).	Intentions (conscious events) do not cause brain events.	CCP cannot, without contradiction, produce the testimony to substantiate the theory.
	β-evidence: Conservation of Energy.	Evidence of Brain events as a physical deterministic system.	Brain events as sufficient and necessary (physical) causes.	Quantum Theory transforms our understanding of determinist physics in line with 'the normal view'.

2 Principles of Credulity and Testimony.

2.1 Exposition.

These principles make an appearance in Swinburne's *The Existence of God* (1979). They are deployed to support claims for the evidential value of religious experience. If I claim to have a religious experience, then the probability is that this is true *in the absence of special considerations*. The principles also have key roles in Swinburne's *Mind, Brain, and Free Will* (2013) in the defence of interactive dualism. Swinburne (2013, pp.54-56) asserts the principles are 'fundamental, a priori and epistemic'. They are fundamental because, he claims, that without due regard to them our knowledge of the world would be highly restricted. They are *a priori* because they are logically arrived at as matters of reason. And they are epistemic because they help us to adjudicate between what is probably veridical and what is not.

The principle of credulity embodies the (epistemic) assumption:

> ...that any basic belief (that is, the content of that basic belief, the proposition believed) is probably true (that is, more probable than not that the belief is true) on the believer's evidence that he believes it – in the absence of evidence in the form of other basic beliefs of that believer which makes it probable that he is mistaken. (Swinburne, 2013, p.42)

If it seems that X is present to a person (P), *in the absence of special considerations*, then probably X is present. In short what we seem to experience is probably the case. He states that we must make this assumption to know anything at all. In respect of interactionism our belief that we do in fact experience ourselves as having intentions which subsequently cause, or fail to cause actions, should be accepted, as veridical, on the grounds of the credulity principle. Swinburne sees no special considerations that could reasonably overturn this judgement.

If I think I remember seeing a rainbow over Penzance last Tuesday; this is probably true. There was in fact a rainbow over Penzance last Tuesday. However, say I remember seeing fairies at the bottom of my garden last Tuesday, we would probably want to say that such a claim must be the subject to scrutiny because 'special considerations' apply. Fairies are not usually considered to be part of the furniture of our world. Perhaps my testimony is unreliable because I was drunk! But if it seems that X is, or was, present to a person, *in the absence of special considerations*, then probably X is, or was, present. In short what we seem to perceive is probably the case.

Swinburne's principle of testimony is a complement to the principle of credulity. This asserts that it is reasonable to believe that experiences that people testify to are probably as they report them. This also is subject to in the absence of special considerations which might provide counter-evidence:

> It is a second fundamental *a priori* epistemic principle that what people seem to be (i.e. apparently are) telling us that they are experiencing or remember having experienced or remember as facts, they (epistemically) probably do or did experience or probably are facts barring counter-evidence. (Swinburne, 2013, p.56)

The principle advises us that we ought to believe P's report. It is then the job of credulity principle to determine whether we should, or should not, believe P's experience to be veridical in the sense that the object X was present, and they did in fact experience it. We should believe other people's testimony unless we have some good reason not to:

> …what people seem to be (i.e. apparently are) telling us they are experiencing or remember having experienced or remember as facts, they (epistemically) probably do or did experience or probably are facts, again barring counter-evidence. (Swinburne, 2013, p.56)

According to Swinburne this is fundamental because generally what we know about history, geography, natural science and so on depends on what others have told us, i.e. the apparent written or spoken testimony of others. If P reports that they experience some object X, or remember experiencing an object Y, then our default position must be that we ought to believe that their reports are

true. We ought to accept that they are in fact experiencing X and that they did remember Y in the absence of any special compelling considerations. This seems reasonable in that, if our default position is always one of deep scepticism of our basic beliefs, then it's hard to understand how we could proceed in the business of the world at all.

Swinburne indicates the special considerations (counter-evidence) by which scepticism may be justified:

(a) If subject 'P' is unreliable;
(b) If similar perceptions are shown to be false;
(c) If there is strong evidence that X did not exist;
(d) If X can be accounted for in other ways;
(e) If any of these considerations can be shown to be true then we are entitled to reject what P claims to have experienced, for example, P was not present at the place and time X happened.

We are entitled not to believe that the experience of X or the report of such is veridical. The two principles are key to his defence of interactive dualism. His first move is to establish what he calls 'the normal view'.

2.2 The Normal View.

Swinburne's concept of interactive dualism is that there is a two way, causal, interaction between minds and bodies i.e. mental substances and physical substances. This, he claims, is 'the normal view'. He believes that, as a matter of fact, we just do experience ourselves (as mental substances) causing brain events and then actions (or failing to cause actions). Further this 'causing' is not to be understood as driven by events but by '...we humans' as pure mental substances (Swinburne, 2013, p.125). He argues that the normal view is grounded in the testimony and credulity principles because this is what we report and experience. This normal view is probably true because there is a lack of sufficient counter-evidence which might lead us to reject such interaction.

It is on these grounds that he rejects epiphenomenalism, for example, which asserts that whilst our mental states may be influenced by our physical (brain) states, our mental states cannot influence our physical states. In his seminal

survey, *The Mind and Its Place in Nature*, (2000), first published in 1925, C.D. Broad defines epiphenomenalism by four key propositions:

P1. Certain bodily events cause certain mental events.

P2. No mental event plays any part in the causation of any bodily event.

P3. No mental event plays any part in the causation of any other mental event.

C. All mental events are caused by bodily events and by them only.

Swinburne accepts P1 but denies P2, P3 and the conclusion on the probable truth of the normal view.

Swinburne's refutation takes the following form:

P1. We normally believe that our intentions cause our bodily movements, and this is just how it seems to us.

P2. On the principles of credulity and testimony: this is what we report and what we experience so we ought to believe P1.

P3. A scientific consensus is that bodily movements are caused by brain events which are preceded by the formation of intentions.

P4. If intentions cause movements then they do so by causing brain events.

C. Interactionism is probably true and epiphenomenalism probably false.

We find a precedent for Swinburne's approach, for example, in H.D. Lewis, *The Self and Immortality* (1973). Lewis is also interested in demonstrating the truth of mind/body interactions as manifest in our everyday experience. Lewis gives this homely example:

> ...If my wife calls and I go downstairs this is because I understand that dinner is ready, etc.; and surely my understanding counts, short of treating all purposive activity as some curious sort of reflex action. (Lewis, 1973, p.64)

If people report and experience themselves as having intentions and successfully causing actions or failing to cause actions, then they should be believed according to our two key principles.

To summarise Swinburne's case:

> P1. There seems to be a reciprocal relationship between mental and physical events (the normal view).
>
> P2. If there is an absence of counter-evidence then, by the principle of credulity, things are probably as they seem to be.
>
> P3. There is indeed an absence of counter-evidence that things are not as they seem to be (the epiphenomenalism position).
>
> C. We are justified in believing that things are probably as they seem to be: the normal view).

In the next section we see how similar moves are deployed to meet a further challenge to interactive dualism.

3 The Challenge of the Causal Closure of the Physical (CCP)

CCP is, according to Swinburne, an *a priori* argument that, in principle, there can be no interaction between physical events (especially brain events) and mental events. The advocates of the CCP argue that there just is no, and indeed cannot be, any satisfactory explanation of how such categorically different kinds of things such as minds and bodies might interact. CCP is logically embedded in our understanding of physical principles. Descartes himself failed to rise to this challenge and provide a cogent explanation when set by his correspondent Elizabeth of Bohemia, '*How* can the soul of man, being only a thinking substance, determine his bodily spirits and perform voluntary actions?' (Quoted in Smith and Jones, 1993, pp.54-55).

As in the case of epiphenomenalism, CCP must be ultimately false on the grounds of the principle of credulity and cannot be a sound objection to mind/body interaction (Swinburne, 2013, p.104). The form of Swinburne's refutation of the *a priori* argument is:

> P1. Brain events and conscious events are different kinds of events.

P2. Brain events sometimes cause conscious events.

P3 We just do have experiences of the effect of our mental states on physical events.

C. There cannot be a good *a priori* objection to the claim that conscious events sometimes cause brain events.

Swinburne's first counter move is that the premise of the *a priori* CCP argument is in fact false. It just is obviously not the case that unlike things do not interact:

> Humans have known for centuries how to produce substances of new kinds from other substances of very different kinds without having any remotely plausible theory of how this happens; they have known how to produce edible plants from seeds, water, and sunlight, intoxicating wine from grapes, glass from sand, and so on and so on. (Swinburne, 2013, p105)

But the analogy he employs here is with physical processes. On its own this cannot meet the *a priori* case given the uniqueness of difference between the mental and the physical. There is a world of difference between this and producing wine from grapes. His argument provides no explanation of just how corporeal and incorporeal substances are supposed to interact. Levin (1979, p.79) points out we are normally familiar with the sorts of stuff material things are made of, but we have no idea of the 'stuff' mental substances are made of; if indeed it even makes sense to talk of mental substances being made of anything at all.

Descartes' response to Elizabeth of Bohemia suggests the pineal gland is the conduit for communications between mind and body (Kenny, 1993, pp.224-225). This merely pushes the argument further back. However, the brain seems to play a similar role to the pineal gland for Swinburne; for example, it mediates between intentions (mental events) and bodily movements (physical events). How this happens remains mysterious in a way the transformation of grapes into wine is not.

Swinburne's second argument is that CCP is breached on the evidence of long established experiences of just such (undeniable) interaction between the

physical and mental. On the credulity principle such interactions are probably true.

> Sticking a needle into almost anyone does cause some brain event which causes them pain, and this happens whoever sticks the needle in, and whenever they do it. Depriving someone of liquid for many hours causes some brain event which causes almost anyone to have a very strong desire to drink. There are evident simple causal connections of this kind well known to the human race for millennia. (Swinburne, 2013, p105)

There appears to be no 'special considerations' or counter-evidence which might compel us to reject this normal view.

4 The Challenge of Empirical Evidence:

In this section, we move from an *a priori* challenge to *a posteriori* evidence for CCP (See Table Two above). This takes two forms: experimental evidence from neuroscience ('α-evidence') and evidence derived from certain physical principles ('β-evidence'). These are counter to our normal view, as described by Swinburne, in suggesting '...intentions do not cause the brain events which cause the bodily movements we suppose they do' (2013, p.106).

4.1 Neuroscience: α-evidence.

4.1.1 Wegner.

Daniel Wegner (1948-2013) in *The Illusion of Conscious Will* (2002) assembles evidence of a range of psychological phenomena which seem to show that our sense of intending, or to use his terminology 'willing', is illusory. Behaviour and intentions, on this reading, are products of unconscious, brain events. This leads him to the startling conclusion that free will is simply an illusion. Evidence for this can be found apparently in an eclectic range of phenomena such as hypnosis, automatic writing, Ouija board spelling, spirit possession, associative disorder and trance channelling. He describes a range of cases where belief in intentional behaviour seems either mistaken or unrecognised.

First, subjects 'feel' they are willing an action but are not doing so. Swinburne 'translates' this as intending some bodily movement but being mistaken about it, either because the bodily movement doesn't occur, or occurs for other reasons. In Wegner's experiments subjects are led to believe they are causing an event, such as stopping a pointer moving, when they are not. Secondly, subjects believing that their intentions are not causing events when in fact they are. For example, with 'Chevreul's pendulum' where subjects are not aware that they are making the pendulum move by their own subtle body movements. Thirdly, subjects will an act of a certain kind but do not feel that they do. For example, under hypnosis. Fourthly, the occurrence of intentional movements without subjects having conscious intentions, for example, sleepwalking. Finally, we commonly, in the present, rationalize aspects of our past behaviours. We may have claimed intentions to do things but in fact it is implausible to suppose that such intentions ever existed.

Swinburne's refutation of Wegner's claims again rests on the principles of credulity and testimony. First, Wegner's account seems to present the normal view of intentions with counter-evidence that we cannot necessarily believe subjects' experiences. Swinburne concedes that of course it is possible that a stage illusionist can manipulate us into mistakenly believing we are doing something when we are not; for example, seeing a woman being sawn in half. Psychologists can set up artificial situations in which we appear to be causing events when we are not (or *vice versa*). Cases of deception, Swinburne argues, can't really count against the fact that we do, in an everyday sense, cause our bodily movements and actions. We are not in error about the causal efficacy of our intentions. In the case of Chevreul's pendulum or automatic writing Swinburne remarks that; 'The fact that sometimes intentional-type movements occur without intending this, gives no reason to suppose that normally our intentional-type movements occur without this' (Swinburne, 2013, p. 107). Wegner's examples seem to amount to no more than pointing at odd things and then suggesting we consider the oddities to be the norm.

Secondly, Swinburne (2013, pp.42-44) argues that the principle of credulity rules out Wegner's conclusion that the weight of 'odd' cases must become the new normalcy for the good reason that we ought, special considerations apart, to believe that things are as they normally seem. Wegner's examples are

artificially contrived to deliberately deceive experimental subjects into making things look differently to how they are. Wegner gives some evidence that we may be mistaken in believing that, on some occasions, someone's intentions caused some bodily movements to occur. But this is hardly sufficient grounds for abandoning the common-sense assumption that generally our intending can lead to the action(s) intended. Swinburne is surely right in this.

Even if we had enough evidence of the Wegner sort, that our intentions did not cause our bodily movements, it would, Swinburne claims, be self-defeating. For the following reasons: (i) such evidence would have to include testimony about the subject's conscious mental events (intending something); (ii) that evidence would only be obtainable if the subject's conscious events did indeed cause brain events; (iii) and those brain events cause the subject to relate what those conscious events are. Thus, a subject's non-illusory prior, causally effective intention would have to be the grounds for showing that prior, causally effective intentions are an illusion. I will call this 'the evidence paradox'. Since Swinburne uses this pattern of refutation frequently I defer an account of how it may be resolved to a later section.

4.1.2 Libet.

Experimental evidence published by the physiologist and neuroscientist, Benjamin Libet (1916-2007) seems to show that the brain engages in (unconscious) activity prior to the subject's reporting intentions to act (Libet, 2004). If true, this might be evidence against the claim that mental states may cause brain states. There is no space here to describe in detail Libet's work. But this is hardly necessary since Swinburne's reasons for rejecting such work, as counter-evidence to interactionism, are matters of principle rather than the reliability of Libet's method. Swinburne's counter claim is that *no* experimental results could ever demonstrate this. If we apply the principle of testimony the theory leads to the evidence paradox.

Libet's most famous, and controversial, experiments are interpreted as showing that unconscious electrical processes in the brain ('readiness potential') precede the conscious decisions of experimental subjects to perform voluntary, spontaneous acts. The core of the work consists of a comparison of the timings

of neurological activity with the timings of self-reported decisions to act (say raising an arm). The results implied that unconscious neural processes preceded, and potentially caused, volitional acts. However, subjects reporting their experiences *felt* that their volitional acts had been consciously motivated. This corresponds to Wegner's claim (above) that our sense of free will is often illusory.

Libet's experimental data can be interpreted as brain events precede the formation of conscious intentions. However, to establish this, experimenters must rely on the apparent testimony of experimental subjects. This gives rise to the evidence paradox again. If the experimenters are to rely on apparent testimony they must:

> ...assume that the subjects are caused to say what they do by a belief that the conscious events occurred and an intention to tell the truth about their belief – a causal route which must go through a brain event. (Swinburne, 2013, p.119)

But for this testimony to provide evidence, to have a justified belief in the theory, we have at the same time to assume the theory is false – the formation of a belief (a mental event) must be prior to the brain event. This is precisely what the theory is supposed to forbid. The paradox is that the experimenters must rely on the falsity of the theory to have reliable testimony the theory requires for its confirmation.

A similar paradox applies to any inference from behaviour of subjects to their conscious events. Inference from behaviour to mental events presupposes that the mental events cause the behaviour rather than brain events. A similar logic, Swinburne believes, applies to apparent memory:

> ...that not merely have Libet-type experiments not so far shown that in their experimental circumstances intentions do not cause bodily movements, but that – even if the crucial predictions necessary to show this proved correct that would only show that (CCP) held in these special circumstances on the assumption that in general it was false. (Swinburne, 2013, p.120)

He concludes that no experimental results, in principle, could show CCP to be justified.

Interestingly Swinburne and Libet seem to share the view that we can isolate 'loose and separate' moments of consciousness and recognise these as

operative, independent causal factors. This is surely one of the methodological problems with Libet's experiments, and Swinburne's approach, to analysing mental events, especially in relation to subjects' reporting their volitions and their relationship to neurological activity. (For a fuller account of the methodological 'shakiness' of Libet's work see Raymond Tallis (2011, pp.54-56; pp.247-256)).

Of course, an organism being in a particular state of consciousness is a causally necessary condition for further and future 'cognitive achievements'. Suppose I intend to play the violin and I'm already a skilled performer. I have an intention to pick up the instrument, raise it to my shoulder, and raise the bow and so on. Is this one intention and action or many and if many how do we separate them? I begin to play. Is every movement of the bow a separate action with its corresponding shadowy counter part of intention? I suspect that in playing the violinist is not consciously forming multifarious intentions for every facet of the performance. Wouldn't this be endless? There seems to be something very logically odd about the models of behaviour, intention and action employed by both parties to the debate.

4.2 Physical Principles: β-evidence

In this section, we examine Swinburne's response to a further challenge to interactionism: physical events are caused by, and only by, other physical events. This opposes Swinburne's claim that mental events, such as intentions, may cause physical brain events. The challenge is rooted in two principles. First, any causal interaction must involve an exchange of energy; and secondly, energy only increases or decreases in some region of space resulting from energy input from or to a neighbouring region. These principles in turn may be derived from the well-established principle of 'the conservation of energy'. If these principles are correct then brain events may only be caused by other brain events with which they exchange energy. In such a closed physical system, there is simply no scope for Swinburne's non-physical, non-spatial mental substance to make any difference to neurological happenings in the brain.

Levin (1979, pp. 84-85) argues that relationships between mind and body, as implied by interactionism, are practically impossible. For example, consider the simple case of my moving my arm. In a resting position, it has mass *w*. If I move my arm from this motionless position a distance *d,* then enough work must be done to move *w* through *d.* Where did the energy (*e*) come from? Empirically it is a consequence of the law of the conservation of energy that the energy must come from somewhere; the creation of energy out of nothing is physically impossible. In addition, the energy involved in moving my arm must come from some preceding physical event. The contraction of muscles in my arm is the cause of *e*. The total energy in the universe would increase out of nothing if non-physical mental events caused physical events.

Smith and Jones, in *The Philosophy of Mind* (1993, pp.53-59), make a very similar case. The 'conscious mind', a mental substance, on Swinburne's account, brings about changes in the brain which lead to physical events. Bodily movements we know are neural events. If neural events cannot be caused by a pure mental substance, then the conscious mind would apparently have no role to play. Thus, the interactive dualist must be able to show that, and how, changes in brain cells are brought about by a mental substance, that is, there are biochemical and electrical changes brought about in cells not caused by prior biochemical or electrical activity. The problem is this:

> ... goes clean against a fundamental principle of the physical sciences, namely that the causes of physical changes are entirely other physical events', the CCP claim. Biochemical and electrical changes are to be explained in biochemical and electrical terms; governing laws allow no room for extraneous, immaterial, non-spatial causal influences. (Smith and Jones, 1993, p.58)

To counter such claims Swinburne mounts what we might call 'the new physics argument' to the effect that CCP does not hold universally (Swinburne, 2013, p.112). Swinburne's move is (a) to shift the issue from physical closure to one of determinism versus indeterminism; and (b) introduce 'quantum indeterminacy' to provide the necessary 'wriggle room'. (It is not clear why an indeterminate system cannot also be physically closed.) He argues that Quantum Theory has superseded classical physics and according to Heisenberg's Uncertainty Principle, all subatomic phenomena are infected by unpredictability. Small-scale indeterminacies can have big effects. The kind of classical physics on

which CCP is supposedly founded is now considered to be reduced to statistical generalisations. But none of this discussion seems to take us out of the realm of physical principles. Quantum mechanics is, after all, a physical theory about how the physical world works even if the implications of indeterminacy are somewhat mysterious. Indeterminacy doesn't invalidate established Newtonian principles at the macro-level.

Swinburne claims that quantum theory has led to various (admittedly speculative) theories about just how mental events might influence brain events. For example, the 'synaptic cleft' seems to be a favoured site for the possibility of such influence:

> The amounts of transmitter chemicals and the distances involved are very small, and so it is often a matter of natural probability of some significant finite value greater than 0 and less than 1 whether the electric potential caused by a neuron firing is transmitted to a neighbouring neuron. Various writers have argued that this the point at which mental events could affect brain events without violating any physical principles. (Swinburne, 2013, p.115)

Nothing shows that some incorporeal substances or non-physical pure moments of consciousness could be, or have been, identified as causal factors acting on physical systems. Swinburne cites Henry Stapp's work on 'Quantum Mechanical Theories of Consciousness' but even he has to admit that '...the vast majority of physicists' regard Stapp's theory '…as highly speculative and ill developed' (Swinburne, 2013, p. 117). It seems that the interactive dualist can no more look for comfort to the 'new' physics than the old.

4.3 But CCP can never be justified!

In response Swinburne moves to show that CCP can never, in principle, be justified. He poses the question: 'What would be necessary to justify a belief in such a deterministic (physical) system implied by CCP? He claims that what needs to be shown…

> ...for any random sample of brain events (and especially ones supposed to be caused by conscious events), that it follows from such a theory that each of them has some brain (or other physical event) as its immediate necessary and sufficient cause, that would

> seem to be powerful evidence in favour of CCP. (Swinburne, 2013, p.121)

He argues acquiring such evidence is deeply problematic given that it must rely on evidential testimony - a re-occurring theme.

First, if CCP is true, then the deliverances of apparent memories of past events and testimony about those events is undermined by the fact that they are not caused by experience of those events because all we have are the deterministic deliverance of brain events. They cannot be determined by previous 'experience' because Swinburne's dualist assumptions mean they are mental events and therefore ruled out as possible causes by CCP. We are not entitled to a justified belief in such a physical, deterministic theory because it precisely requires CCP to be false. Such a theory again denies the possibility of the very evidence it requires for justified belief. This is the further return of the evidence paradox.

Secondly, there is a problem about CCP's ability to deliver true predictions. We would need (a) a justified belief that certain relations between brain events occurred and (b) a justified belief that such events had been predicted by the theory. However, anyone who had not worked out for themselves the predictions of the theory must then rely on the apparent testimony of scientists to provide that evidence. But this requires that the scientists had a conscious belief ('seen') that was indeed what the theory predicted. In other words, there is dependence on the evidence of the 'conscious events', in Swinburne's dualist sense, of the scientists. However, if the deterministic theory is correct scientists would not have been able to provide their testimony by any conscious event '...neither by their intention to tell the truth nor even merely by their conscious belief about the theory predicted.' If this is right then there can be no justified belief in any reports about calculations of predictions '...since believing what the scientist reported would undermine the credibility of their apparent testimony to it' (Swinburne, 2013, p.121-122).

Swinburne claims that it is not clear that such objections even hold for physical matter (citing Quantum Theory). In summary:

> P1. There appears to be no reliable counter-evidence to the possibility that mental substances interact with physical substances.

P2. It must be the case that our intentions often cause our bodily movements.

P3. There is no reason to deny that other mental events (whether conscious events or continuing mental states) may sometimes cause brain events.

P4. There is no reason to deny that when we seem to remember past conscious events, probably they happened, unless there is some more probable counter-evidence with respect to an event.

P5. When someone tells us about their past conscious events such as their assessment of some scientific theory then their report of the occurrence of the events is true unless there is some relevant counter-evidence.

C. There can be no justified belief in CCP.

However, these moves only work if we assume that that mind/body relationships are of a quasi-mechanical, causal nature; that discussions of mental events occupy the same logical space as physical events. I suggest that this is just false.

4.4 The Evidence Paradox

In previous sections we have illustrated how Swinburne believes that the counter-evidence produced by Libet and Wegner, for example, lead to a self-refuting argument, the evidence paradox. For the claims of the neuroscientists' account to be true then their theory itself must be false. If the experimentalists' theory is true then, like every other such knowledge claim, it is a function of the determinative interaction of brain cells. This means that the testimony, the sentences that subjects utter, are the product of physiological causes. If there are always physical reasons (causes) why subjects (or any of us) make the sounds that they do make then they cannot be said, and know that they have, reasons for believing the sentence (propositions) so produced (Flew, 1975, p.59). If what I say is determined in the way the neuroscientists suggest, then we can have no good reasons for believing and saying what I do say. And what

is sauce for goose is sauce for the gander: experimenters themselves will also be the 'victims' of their own theory. There are equally no evidencing reasons for their theoretical constructions if what is going on is as we have described it. The core of the paradox is:

> ... if the sufficient reason why I believe such and such a proposition is that I am in such and such a physiological condition, then there can be no room left for me to have – or even for there to be – good evidencing reasons to warrant the belief. (Flew, 1978, p.92)

To resolve a paradox, we must be able to show that the reasoning is erroneous or there is a (hidden) flaw in the premises. Swinburne removes the paradox by reference to the fundamental and epistemic nature of the credulity and testimony principles. We just cannot reject the primacy of what persons' experience and report without epistemological disaster. These *a priori* principles show that the premises must be faulty and that things must be as they seem and as we report them; we just do, as a matter of fact, experience ourselves as making choices, having intentions and so forth (the normal view). The neuroscientific theory must be false. We are aware of our minds as mental substances interacting with our physical bodies and the world. In the following section we will show that the credulity and testimony principles cannot deliver the epistemological goods because they are not, after all, epistemological principles.

The paradox can be resolved differently to Swinburne's solution. First, his interpretation of the paradox is to make the (false) assumption that we cannot hold both descriptions at the same time; that is if the physical description is true then this must preclude a description in terms of having grounds for believing what we do believe. Secondly, I suggest, in recognising a distinction between types of reasons we can resolve the paradox. In talking of physiological or neurological reasons we are talking of reasons in the sense of causes. However, in having sound reasons for what subjects, for example, may or may not believe, is to talk about reasons in the sense of grounds for their beliefs. No reason is given to show that these two types of reasons, reasons (grounds) and reasons (causes), are necessarily competitors for the same (logical) space. They do not preclude each other. But it would 'downright fantastic', in Flew's

phrase, to suggest that because of the logical irreducibility of mental concepts to physiological one's mental characteristics must, thereby, belong to an incorporeal pure mental substance (Flew, 1975, p.60; Flew, 1978, pp.92-99).

5 Criticism of the Credulity and Testimony Principles.

I have argued in this chapter that the credulity and testimony principles are the pivotal points in Swinburne's efforts to refute challenges to what he refers to as the normal view of body/mind interactions (See Table Two: Counter-evidence to Interactionism, above). I suggest that this strategy is unsuccessful because the key principles cannot deliver the necessary epistemological and logical goods. Swinburne's assumption that they are fundamental, *a priori*, epistemic principles is simply mistaken.

5.1 Are the principles epistemic?

Swinburne states that the principles of credulity and testimony are 'epistemic'. We are to understand by this that these are principles that govern how certain sentences are known to be (probably) true. Russell Blackford (2014, p.110) in an early review of *Mind, Brain, and Free Will* (2013) asserts that the credulity principle is based on our knowledge of when our perceptions and memories are likely to be reliable or not. But suggests that:

> The Principle of Credulity is no substitute for a nuanced and well-informed understanding of when our impressions are probably reliable and when they are probably not. (Blackford, 2014, p.110)

Blackford implies that the principles do not, of themselves, adjudicate between what is veridical and what is not. It cannot be the case that we can know something to be true simply '… on the believer's evidence that he believes it' (Swinburne, 2013, p.42). This is what Karl Popper famously rejected as 'psychologism' (Popper, 1980, p.93). The fact that I believe in X's presence is undoubtedly a fact about the world, but it goes no way establishing that X is

present. We might say that your belief in X might motivate me to believe it, but this is far from it being known epistemically.

The credulity principle claims that '...what seems to us to be so, probably is so that our apparent experiences are probably our real experiences' (Swinburne, 2013, p.42). But this cannot be the basis on which sentences (propositions) are *known* to be true. The core is the ambiguity of 'experience'. There is a world of difference between: (a) accepting the truth of P's genuinely held belief that they did have an experience of X and (b) accepting whether they did have an experience of X. There is a distinction between two very different senses of the word 'experience' (Flew, 1966, pp.125-126; Flew, 1985, pp.38-39). In the first sense to have an experience simply refers to whatever the subject is undergoing psychologically. In the second sense having an experience of something implies that there must be an actual object present as well as whatever the subject is undergoing psychologically. The subject's report will be the same for both senses and would not entail the objective existence of anything. There must be a further warrant (i.e. genuinely epistemological grounds) if we are to claim epistemically that he had an experience of X and indeed X was present.

It is indisputable that people claim to have vivid experiences that convince them, without a doubt, that they have been in contact with some object corresponding to their experience. The crux of the matter is whether such subjective conviction can be furnished with the kinds of credentials to render the experience, in the second sense, true. We have no grounds for moving from the proposition that because a person can be truly said to have such and such an experience (an experience of being in the presence of an X) that we can immediately infer the experience is 'what is truly and objectively the case'. Real experiences are not necessarily experience of the real.

Swinburne's criticism of the counter-evidence rests heavily on the principles of credulity and testimony to decide issues as to whether we have experiences of ourselves as interacting mental substances, of agent causality, free will and so forth. But, as Blackford suggests, it is far from clear, '…what difference there would be in how it feels to exercise agent-causality, event caused interaction, free will or compatibilist free will or not free will at all' (2014, p.111).

5.2 Are the principles fundamental?

In Swinburne's defence we might cite his qualification to the principles that things are probably as they seem in 'the absence of special considerations'. These considerations are: (a) if subject 'S' is unreliable; (b) if similar perceptions are shown to be false; (c) if there is strong evidence that 'X' did not exist or (d) can be accounted for in other ways. If any or all of these 'special considerations' apply we may be entitled to reject the truth of what S claims to be experiencing. But if we are to appeal to such considerations then we must (logically) have knowledge from other sources as grounds for believing that S's claims are false. This suggests that the principles are not as 'fundamental' as Swinburne asserts.

It is not clear why such 'considerations' are deemed to be 'special'. Swinburne aims to place the burden of justification on the sceptic. Now this is may not altogether unreasonable. We couldn't get along in the world if we didn't accept many beliefs and knowledge claims on trust. But there are two problems here. First, from a fundamental, epistemological point of view we might reasonably want to say that there is nothing special about Swinburne's 'special considerations'. They are standard, reasonable conditions to be applied to knowledge claims especially when we are familiar with the unreliability of human testimony, for example, given bias, mistakes, lying, exaggeration and so forth.

Secondly, the issue is: 'When is it reasonable to accept the principles of credulity and testimony as default conditions and when is it reasonable to begin from the position of the so called 'special considerations'? This must (logically) depend in some way on the scope of the empirical claims being made. For it must be the case that what we claim to experience, or report, must vary greatly in the scope of their empirical and logical content. The greater the scope of the claims being made (and this must apply particularly to the claims of religious experience) the greater will be the scope for falsification and the higher the risk of just being plain wrong. We then are entitled to apply more sceptical approaches. High empirical content must, logically, imply the lower probability that what is claimed is true (Collingridge, 1987, p.20). But it is precisely in

controversial fields where the scope of the claims being made are great that Swinburne invokes the principle of credulity and plausibility: that is in the contexts of the evidential truth of religious experience and in the theory of mind/body interactions. In the former to make 'supernatural' claims must automatically invite criticism. The theory of interactionism makes claims of great scope where the probability of error is high, but the subjective nature of principles prevent adjudication in both cases.

5.3 Are the principles *a priori*?

Further I suggest that credulity and testimony are not *a priori* principles. If they have any purchase at all then it is because they are pragmatic and ethical. As Swinburne himself suggests in general if we are to acquire beliefs and knowledge at all we must trust a range of sources. Trust is an ethical relationship we have with others or a variety of justified true beliefs. We can interpret the two principles as moral imperatives; in the case of the credulity principle we ought to believe that people are experiencing what they claim to be experiencing. But this is not a test that what they are experiencing is veridical. In the case of the principle of testimony we ought to believe what people say: if person P testifies to an apparent experience of X then we ought to believe them. But, as we have already claimed, the principles in themselves do not adjudicate between what is veridical and what is not. What they tell us is that we ought to accept what people say in the immediate absence of counter-evidence or special circumstances. This does not preclude relevant special circumstances arising in future even if we don't have them now.

Finally, we might say that what we have with these two principles are '*prima facie* duties' or 'conditional duties' (i.e. we *ought* to accept what people testify to and what they say they experience). This acceptance is on ethical grounds rather than epistemic ones (Boyne, 2018). They do not provide public tests of truth or falsity. The use of the phrase *prima facie* here is in W.D. Ross's sense of an obligation that only holds subject to not being overridden by a superior obligation (Ross, 1950, p. 19). J.L. Austin observes in 'Other Minds':

> It is fundamental in talking (as in other matters) that we are entitled to trust others, except in so far as there is some concrete reason to distrust them. Believing persons, accepting testimony, is the, or one of the main points of talking. We don't play (competitive) games except in the faith that our opponent is trying to win: if he isn't, it isn't a game, but something different. So we don't talk with people (descriptively) except in the faith that they are trying to convey information. (Austin, 1970, pp.82-83)

This, of course, is common sense stuff but the use of the word 'faith' here, I believe, reinforces my point that we are dealing with an ethical relationship and not strictly an epistemological one.

6 Conclusion

I set out in this chapter Swinburne's response to counter-evidence to the doctrine of interactionism. Central to his version of substance dualism is the idea that mental and physical substances interact and mental substances may shape processes in the brain. He believes that this theory reflects what human beings experience. This he believes is 'the normal view' that we, for example, experience our intentions causing our actions. He cites evidence and counter theories which challenge the interactive dualist position such as the work of neuroscientists Wegner and Libet, and physical theories rooted in the CCP.

In each of the cases of counter-evidence the crux of Swinburne's refutation rests on his two key principles of credulity and testimony. These principles he asserts are fundamental, *a priori* and epistemic. In the absence of 'special considerations' (a) if it seems P experiences the presence of X then probably X is present to P; (b) on the testimony of person P that we ought to believe that P did apparently have an experience of X. The application of these two principles over throws the counter-evidence to interactionism. The refutation rests on the supposed paradox, that for the counter views to his preferred normal view to be true, the theories themselves must be false based on his two key principles. However, my conclusion is that the principles are not well founded because they conflate the difference between reasons as grounds and reasons as causes. I argue that they are not fundamental, not *a priori* and, crucially, not

epistemic. My alternative interpretation is that they are ethical injunctions to believe, as provisionally true, what people may testify to, claim to be experiencing or have experienced.

Chapter Six: Personal Identity.

1 Introduction

In this chapter I examine Swinburne's account of personal identity. He defends a radical account of the nature of personal identity which, by his own admission, is at odds with dominant, contemporary conceptions. Swinburne maintains that personal identity, at its most fundamental, is characterized by being 'a pure mental substance' (a soul). Persons may be thought of as essentially incorporeal and whilst our physical being may not be inconsequential it is, nevertheless, inessential to our present and continued existence. Swinburne argues that we cannot look for sources of personal continuity in our memories, character or physical characteristics. In *The Coherence of Theism* (1977, p.110) he writes: '…that the identity of a person over time is something ultimate, not analysable in terms of bodily continuity or continuity of memory or character'. This is pretty much the conclusion he defends in *Mind, Brain, and Free Will* (2013) some forty years later albeit by somewhat different means.

It is important at the outset to stipulate the sense in which we are talking about 'personal identity'. Bernard Williams (2006, pp.56-64) differentiates two senses of identity: identity as 'particular' and identity as 'type'. In the first sense we are referring to a person or a thing as being particular or uniquely itself. A particular identity answers the question of 'What is it that makes (criterial) this person or thing the same over time?' 'What is it that allows us to identify and re-identify X at time t_1 as the same X at time t_2?' The answer to this question for Swinburne, at least in relation to persons, is 'thisness' (haecceitas). In the second, and perhaps more familiar sense, identity refers to a 'type'. This answers to questions such as 'What type am I or what type is it?' 'How am I classified or how is it to be classified?' For example, questions of social identity are discussed by referring to concepts of race, class, nationality, sexuality, taste in clothes, and so forth. Suppose a valuable witness is placed in a witness protection program. They can be allotted a 'new' identity in the sense of 'type', but this will not extinguish their particular identity. They will remain the same

person as they were before they entered the program. It is this sense of particular identity with which we are concerned.

2 Swinburne's Approach

This account rests predominantly on the three principles discussed in the previous three chapters: a metacriterion, a theory of privileged access and the principles of credulity and testimony. First, Swinburne shows how, and that, competing theories are inadequate to the task of determining personal continuity and sameness. For example, the limitations of other theories of identity are thrown into relief by the application of a metacriterion; the requirement that our vocabulary ought to be able to tell the whole history of the world (Swinburne, 2013, p.10). By 'vocabulary' here Swinburne means the 'referring expressions' that will identify (or differentiate) the key constituents of the world: substances, properties, times, and events. A theory of personal identity must meet this requirement. A complete account of the world must, according to Swinburne, include two types of substance, properties, and events. The rationale for 'substance theory' is established by the metacriterion and the rationale for a two-substance theory is supported by the doctrine of 'privileged access' and by a theory of informative designation. Therefore, for Swinburne, the problem of personal identity is only to be resolved by reference to mental and physical substances. In order show this he provides a 'deeper' account of substance dualism.

Secondly, Swinburne maintains a distinction between 'humans' and 'persons'. He regards persons, in the normal sense of the word, as '…substances with a capacity (at least after normal growth) for beliefs and actions of the degree of sophistication typical of those present-day earth-inhabitants called "humans"' (Swinburne, 2013, p.141). Swinburne refers to 'humans' as '…persons who have or have had at some time a kind of body and an ancestry like those persons whom we call "humans" today'. But it is not obvious that this is what we do normally mean when we ordinarily use the word 'persons' and 'humans'. However, there are cases where the distinction might be meaningful: (a) there

are circumstances where we might doubt whether a human being is anymore a person, for example, when someone is in a terminal vegetative state; (b) cases where something is not a human being, in our normal sense, but might have characteristics of persons. We might imagine visitors from other worlds or advanced forms of artificial intelligence falling into this category. We might need to debate whether or not to extend some rights to such entities; (c) there is an analogous debate around whether foetuses are persons.

Thirdly, Swinburne's interest is in giving sense to the concept of 'disembodied persons':

> With these definitions I can raise the question whether humans are mental substances, and whether they are pure mental substances. (Swinburne, 2013, pp. 141-142)

At the outset he leaves open the questions as to 'Whether there may be persons who were, or are, never embodied? or 'Whether humans can become disembodied?' He has to demonstrate that talk about 'disembodied persons' is intelligible. The problem he must overcome is that the cases we have considered refer to physical entities of one sort or another, but disembodiment seems altogether of a different order. It seems self-contradictory to assert that we can live and have experiences without a body. The reason is that our 'person words' such as 'you', 'I', 'somebody', 'Swinburne', 'woman', and so on although different in their linguistic function are all used to refer to physical objects. And these objects are, paradigmatically, the kinds of objects that we can (physically) see, hear, touch, indicate, and, significantly, talk to. Swinburne's solution is to argue that humans are not objects at all. The groundwork for this move is already to hand in his metacriterion and his two-substance ontology.

Fourthly, the perennial challenge for advocates of non-physical, personal identity is '…supplying effective means both for the identification of such beings, and for their individual re-identification through time' (Flew, 1986, p.104). 'What identifies and individuates separate but related substances?' and most acutely 'What individuates and constitutes the unity of mental substances?' Various competing models of personal identity are tested using thought experiments. The conclusion that Swinburne draws from these tests is that personal identity is '…something ultimate, not analysable in terms of bodily continuity…'

(Swinburne 1977, p.110). In the following section we examine Swinburne's case for specific identity conditions for human persons according to his definition (above).

3 Identity Conditions

We can specify the identity conditions of objects in terms of *synchronic* (or at-a-time) identity and diachronic (across-time) identity (Dainton, 2001, p.356). Synchronic identity specifies the way an object is an object of a particular type. The conditions required to be a sparrow quite obviously differ from those required to be a metamorphic rock. Swinburne argues that the synchronic identity condition for a person is the co-experiencing of a mental substance. Defining an object's diachronic identity refers to the kinds of changes an object can undergo and remain in existence or not; different things will have different diachronic conditions. Diachronic identity, according to Swinburne, is constituted by 'thisness. In neither case are physical identity conditions required.

3.1 Synchronic unity of human persons.

Co-experiencing of mental events must, according to Swinburne, show that one subject must be experiencing all such events. (I refer to this as the 'co-experiencing thesis'). The overlapping of two or more conscious events must be events of the same substance and constitute a 'synchronic unity'. This is entailed by the very fact of 'overlapping', for example, in the simultaneous experience of heat, light and pain when being burned. He recognizes that this argument is threatened by apparent examples where this does not seem to hold:

> …occasionally in abnormal circumstances some human bodies behave in a disunified way which might seem to suggest that the conscious events connected with the bodies of those persons are not all co-experienced, that is, not events in a single consciousness. (Swinburne, 2013, p.144)

This may threaten the possibility that one body is ‘occupied’ by multiple, independent sites of consciousness rather than one unitary mental substance at the same time.

Swinburne postulates that co-experienced mental events belong to the same substance and non-co-experienced events are the properties of different substances. This constitutes the unity of the human person: ‘...humans are mental substances since their spatial boundaries are determined by a mental property’ rather than physical ones (Swinburne, 2013, p.144). Co-experience means one substance, in the present moment, has the conscious experience of hearing the waves of the sea, seeing the rain fall, smelling the sea spray, and the like. A complete description of the world must include all substances which have only and all conscious properties co-experienced at the one time. If this is the case then, in addition to these mental properties, this substance may also have physical properties and therefore physical extension in space. It must have within its borders all sections of its brain and events that are immediate causes or effects of simultaneously experienced (co-experienced) conscious events. Neuroscience suggests that the immediate, proximate causes of different kinds of conscious events (e.g. olfactory sensations) involve events in different parts of the brain. The proximate causes of different properties (such as colour and shape) should be considered as one conscious event (e.g. perception of a coloured shape) may involve events in different parts of the brain.

Lynne Rudder Baker (2014, pp.10-11) has problems in understanding and accepting Swinburne’s arguments for mental substances from synchronic unity. In particular she is concerned by Swinburne’s claim that the properties of co-experiencing certain properties at a time ‘…will delimit the physical boundaries of a substance and help to determine which physical properties it possesses’. The use of the term ‘delimit’ is equivocal in the sense that it can be interpreted causally or ontologically. To reach Swinburne’s desired conclusion, that humans are mental substances, it must be read ontologically if spatial boundaries are ‘determined’ by mental properties. But ‘delimit’ is a causal idea rather than an ontological one. She concludes on this basis that the argument is invalid or has a false premise but, on both interpretations, it is just ‘unsound’.

Swinburne has rebutted Baker's argument by claiming that 'causal facts determine the ontological facts' (2014, p.57):

> What makes a brain my brain (ontological fact) is that my mental events are caused by, or cause, events in the brain (causal fact). It then follows that if I have any physical properties (as Baker and almost all of us hold), then necessarily a mental property of mine determines that certain of my brain properties are (contingently, because of the causal connection) among those physical properties…we could regard the physical part of me as only a brain, in interaction with another substance constituted by the rest of my body.

This reply to Baker must rest on the assumption that Swinburne has already satisfactorily shown that physical substances (such as brains) can be said to interact with mental substances. I'm not convinced that this really addresses the core of Baker's argument or rather muddies the water even further. It is just hard to see how an immaterial substance 'causes' or sets physical limits.

Swinburne cites Tim Bayne's *The Unity of Consciousness* (2010) to corroborate his argument to unity from 'co-experiencing of mental events' and to undermine evidence of disunity of consciousness. Bayne, for example, describes three types of disunity: 'representational disunity'; 'access disunity'; 'phenomenal disunity'. In 'phenomenal' disunity a person's conscious events are not had as parts of one overall conscious event; a person experiences e_1, and at the same time experiences e_2, but does not experience (e_1 and e_2) (Swinburne, 2013, p.145). Other examples of 'divided consciousness', detailed by Bayne, include anosognosia ('…where patients seem unable to recognize that they have some impairment of consciousness, such as blindness'), schizophrenia and hypnosis. Swinburne follows Bayne in arguing that in all such cases, and in spite of their apparent deviation from the 'co-experiencing thesis', the conscious events referenced are in fact co-experienced. The strange behaviours can be explained in other ways by, for instance, the patients' unwillingness or inability to articulate their knowledge of relevant mental events.

Swinburne also addresses other psychological phenomena that tell against the 'co-experiencing thesis'. For example, the phenomenon of 'multiple personalities' indicates the possibility that two or even more mental substances may co-exist in one body. Bayne's explanation is that the person does indeed have two different personalities which express themselves at different times or

at the same time. However, this does not imply two independent consciousnesses. '…there is just one person who co-experiences all the conscious events causally connected with that body's brain.' (Swinburne, 2013, p.145)

The strongest case against the 'co-experiencing thesis', according to Swinburne, seems to be evidence of 'split-brain' cases. These are cases in which the corpus callosum, which links the right hemisphere of the brain to the left hemisphere, has been severed. This severing arises either as a genetic defect or from an operation, cerebral commissurotomy (thought to be a possible cure for epilepsy). The two hemispheres of the brain perform in a much more independent way than normally. For example, the right and left hands of a patient may operate in conflict. Is this evidence of two separate consciousnesses? Swinburne side steps the problem by claiming that his concern is not with which of several interpretation of the split-brain phenomenon or similar phenomena might be correct. His concern is methodological. For each interpretation or hypothesis designed to explain data there is some difficulty in explaining some aspects of the various phenomena. Swinburne concludes from the neurological evidence and the evidence of subjects themselves that the issue is just not decidable:

> ...all the evidence would be logically compatible both with all the conscious events being co-experienced, and with there being two groups of co-experienced events, members of each group not being co-experienced with members of the other group, and with some rival hypotheses. (Swinburne, 2013, p.148)

The fact that in some cases it may be difficult or impossible to decide does not shake Swinburne's conviction that the 'co-experiencing thesis' remains cogent:

> P1. A consciousness is 'phenomenally disunified' iff its conscious events (e_1, e_2) are not co-experienced as one overall experience (e_1 and e_2).
>
> P2. e_1 and e_2 are co-experienced iff the substance (i.e. the human person) who has the conscious event e_1 also has conscious event e_2, and so that person will be conscious of having e_1 while he or she has e_2 even if they do not see (e_1 and e_2) as a unified experience (and so suffers from 'representational disunity').

> P3. In the case of P2 even if the subject is not able to use all of it to guide their speech and limb movements (and so suffer from 'access disunity').

What Swinburne is seeking to show is that co-experienced events must belong to the same (mental) substance. The fact that '…we humans co-experience conscious events *entails* that we are mental substances'.

But consider the famous rebuttal by Bishop Butler (1692 – 1752) of Locke's theory of personal identity:

> And one should really think it is self-evident that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity: any more than knowledge, in any other case, can constitute truth, which it presupposes. (Butler, 1906, p.258)

Butler argues that the reason why the identity of incorporeal persons cannot be explained in terms of memory claims is that memory claims presuppose identity rather constitute it. By analogy co-experiencing of events cannot entail incorporeal, mental substances which, according to Swinburne, are the basis of personal identity because it presupposes such substances.

This is explicit even in the way that Swinburne must discuss his problem cases; there must already be someone, a human person (a mental substance) who is having experiences (mental events). Butler suggests this 'wonderful mistake' by Locke:

> …may possibly have arisen from hence; that to be endued with consciousness is inseparable from the idea of a person, or intelligent being. For, this might be expressed inaccurately thus, that consciousness makes personality: and from hence it might be concluded to make personal identity. But though present consciousness of what we at present do and feel is necessary to our being the persons we now are; yet present consciousness of past actions or feelings is not necessary to our being the same person who performed those actions or had those feelings. (Butler, 1906, p.258)

Compare Swinburne:

> For it is an evident datum of human experience that conscious mental events of different kinds (visual sensations, auditory sensations etc.) are co-experienced, that is belong to the same substance, both at one time and over time. (Swinburne, 2013, pp.142 – 143)

Here Swinburne's reverses his previous argument to the effect that co-experiencing is predicated on the logically prior existence of the 'same substance, both at on time and over time'. Thus the co-experience thesis cannot explain or entail mental substance.

3.2 Diachronic unity of human persons.

Swinburne proceeds to the concept of diachronic unity: being the same person over time. This leads to his preferred theory of personal identity: 'the simple theory'. This term appears to be derived from Parfit (1986, p.210) to refer to theories of 'the self' which assume its simplicity and indivisibility. Swinburne's interpretation is that the self is 'a separate feature of the world' which is captured in the concept of 'thisness'. He then counter poses this simple theory to competing theories broadly described as 'complex theories (or views)' apparently adhered to by 'most modern philosophers'. Complex theories characteristically ground personal identity in, for example, physical continuity, psychological continuities such as memory or character and so on (See Table Three, Competing Models of Personal Identity, below). Swinburne tests the cogency of these two types of theories using various thought experiments (See Table Four, Brain Transplant Thought Experiments, below). From this 'testing' process Swinburne claims that the simple theory emerges secure (Swinburne, 2013, pp.148 – 163).

'What constitutes the unity of the person over time?' 'What is it for a person P_2 at time t_2 to be the same person as a person P_1 at t_1?' Swinburne believes that this must involve human persons being mental substances characterized by 'thisness'. What if a succession of different substances participate in elements of the same experience? This is rather like the problem of 'multiple personalities' discussed in the previous section. Is it conceivable that '…the substance which had the first half of one experience be a different one from the experience that had the second half of the experience'? Swinburne's believes that this would just be incompatible with 'the data of consciousness' already established in 'the co-experiencing thesis': '…any awareness of a conscious

event, for example, a pain, is an awareness of oneself as co-experiencing both the first half of the event and also the second half of the event (Swinburne, 2013, p.148).

Is the co-experiencing defence plausible? It seems logically odd to propose that we experience events in fractions! We certainly all have an intuitive grasp of what it is to have (whole) experiences. Swinburne tries to explain this in the following way:

> I shall include as 'experiences', (apparent) perceptions (by means of the senses) of physical events, (apparent) 'experiences' (in the narrower sense) of our thoughts, feelings, sensations, etc. and also (apparent) awareness of truths of reason (awareness of 'seeing' some sentence to be logically possible or whatever). (Swinburne, 2013, p.55)

This may be all perfectly reasonable. But in the use of the concept of co-experiencing he tends to view experiences as somehow self-evidently bounded and discrete – for example in the very use of his nomenclature e_1, e_2, e_3 etc. But having an experience is not like having a 'pound of flesh' with or without the blood. Or we can sensibly talk about having a bath which is only half full of water; it doesn't seem intelligible to talk about only having had 'half an experience of bathing'.

Say I stop typing these words, I lean back in my chair, I look at the picture over my desk, I turn my head and look out of the window, I see the blind, the grey drizzle outside, the sky, the stone buildings, I feel hungry and so on. How many experiences am I having of how many events? Or is this just one present experience? It doesn't seem sensible to talk in quasi-quantitative terms such as 'I have just had three and half experiences of events'. Certainly Swinburne seems to think we can have 'half an experience': '…the substance which had the first half of one experience can be a different one from the experience that had the second half of the experience'! This is like a claim to be only half pregnant.

Why does this matter? Swinburne's belief in the notion of fractions of experience is tied up with his need to demonstrate that co-experienced discrete events must imply a single mental substance rather than different substances. His model for the description of mental events is analogous to his treatment of

physical events i.e. a mental event is the instantiation of a mental property in a substance at a certain time; for example, remembering an apparent memory. However, while his account of physical events seems eminently plausible because we can intuitively grasp the possible measurement of physical events in terms of duration, size, length, intensity and so forth: for example, the car took five seconds to come to a stop. It is just unclear how this might apply to the experiencing of mental events because how are such incorporeal mental events individuated?

However, Swinburne attacks complex views that ground identity in spatial temporal and physical continuity as 'animalist'. Animalism rejects substance accounts of identity in favour of a bundle theory of coexisting physical properties. For Swinburne the rejection of substance theory implies that such theories must fail his metacriterion test as incapable of giving a complete account of the world. Crucially such accounts fail to include pure mental substances. They advocate the view that '…the criteria for the existence and continuing of human beings would be the same as those for the existence and continuing of human bodies' (Swinburne, 2013, pp.142-143). Animalism treats physical properties as essential and mental properties as inessential (See Table Three, Competing Models of Personal Identity, below).

First, he claims that neuroscience, and our own experience, shows that conscious mental events of different kinds (visual sensations, auditory sensations and so on) are 'co-experienced' now (synchronically) and over time (diachronically). Together these constitute the sameness of the human person. Any account of the world would be incomplete if it did not include 'substances which have as their conscious properties all and only co-experienced properties' i.e. non-physical, mental substances. The fact of co-experience shows that co-experienced properties belong to the same (mental) substance both at one time and over time, in other words, the animalist description leaves out mental substances essential to a complete account of the world.

Secondly, given this, we would be failing to give a full description of the world if we did not count among substances with physical properties, and thus with physical parts, substances where the events in those parts are the proximate causes or effects of conscious events which are being co-experienced along

with other conscious events of the same substance. The account would be lacking if we only charted the history of each separate part of the brain and regarded each as a separate substance and thus only included events in those parts which were a proximate cause of, or caused by, the instantiation of mental properties. This would leave out truths about properties (such as co-experienced sensory properties) which would then wrongly have to be attributed to different substances. Since human beings (as co-experiencing substances) typically have physical properties then the mental property of co-experiencing will 'delimit' the physical boundaries of the substance and contribute to determining the physical properties the substance possesses. This is supposed to show that '...what constitutes such a substance which has conscious co-experienced properties is - even if it has also some physical properties – determined in part by a mental property, and so the substance is a mental substance' (Swinburne, 2013, pp.143-144). Thus, animalism must be false. As Baker (2014, p.12) points out this depends on the validity and truth of Swinburne's argument from synchronic unity. But there are strong reasons, as we saw, to think this unsound.

Table Three: Competing Models of Personal Identity (Swinburne, 2013, pp.149-151).

Theory	Characterization	Grounds for 'sameness'	Criterion of personal identity	Swinburne's objections
Complex Theory I	Animalism: physical identity model (a 'bundle theory').	Physical continuity: existence and continuance of a human body: similar properties being instantiated in the same body.	To be a person is to be a human animal; it is sufficient for x to be the same person as y iff x is the same human animal as y.	Fails the meta-criterion test: not possible to have a complete description of the world if substances are individuated by physical properties alone rendering mental life an inessential; feature.
Complex Theory II	Survivalism: psychological continuity (e.g. Derek Parfit after Locke and Hume).	P_1 'survives' as P_2 to the extent of their psychological connectedness and continuity: with similar memories, character and so forth.	P_1 at t_1 are psychological continuous with P_2 at t_2 in so far as there is a chain of persons connecting P_1 and P_2 over the intermediate period – assumed to be caused by the operation of the same brain.	Similarly relies on some notion of underlying physical cause of connectedness and continuity.

Table Three: contd.

Simple Theory	Identity as a separate feature of the world: thisness.	Being that person is compatible with having any particular mental properties or any physical. properties.	Logically possible that P_2 at t_2 can be the same as P_1 at t_1 even if: (a) P_2 does not apparently remember anything done or experienced by P_1 at t_1; (b) and /or has a completely different character from P_1; (c) has a completely different body and brain.	None. Logically possible and metaphysically possible according to the meta-criterion and theory of informative designators.

Swinburne sets out his preferred simple theory as a radical, and true, alternative to complex theories. He defines the simple view as:

> …the view that each person has a 'thisness' which makes him or her that person, a 'thisness' other than any 'thisness' possessed by the matter of their brains; and that being that person is compatible with having any particular mental properties or any physical properties (and so body) at all. (Swinburne, 2013, p.151)

It is logically possible that a person P_2 at t_2 can be the same person P_1 at t_1 even if:

(a) She or he does not apparently remember anything done or experienced by P_1 at t_1 earlier and

(b) has an entirely different character from P_1 and also has

(c) a largely different body (including brain) from P_1.

Swinburne tests the logical possibility of this simple view by thought experiments that produce various hypothetical combinations of body and brain swapping, splitting and replacement (See Section 4 and Table Four, Brain Transplant Thought Experiments, below). He concludes that it is logically possible that a person maybe the same person at different times without the continuity of physical and/or specific mental properties. His next move (Section 5) is to show how such a logical possibility may be converted into a metaphysical possibility.

4 Testing Identity Theories

4.1 Thought Experiment (TE) Tests.

Already in *Personal Identity: Great Debates in Philosophy* (1984) Shoemaker and Swinburne, the two collaborators in a 'great debate', use thought experiments to explore competing identity theories. Swinburne's conclusions are the same ones he reaches in *Mind, Brain, and Free Will* (2013): '…personal

identity is unanalysable in terms of brain and apparent memory, it is usefully defined as consisting in the continuing existence of a soul, while continuity of brain and apparent memory (backed up by continuity of character) provide fallible evidence of personal identity' (Shoemaker and Swinburne,1984, p.49).

Under various imagined conditions we are invited to consider who might count as persons given a range of more or less grotesque surgical interventions. Swinburne believes that animalist or physicalist accounts assume that personal identity is vested in the brain so moving a brain just is equivalent to moving identity. Similarly, for example, Thomas Nagel (1986, p.40) writes that…'The brain, but not the rest of the animal, is essential to the self'. The result of split-brain thought experiments is that knowing all the physical facts about what happens to brains in these transplant situations doesn't answer the question of where *I* have gone. 'Have I survived?' 'Am I now in a body A which was my left brain or in a body B which was my right brain?' Swinburne concludes that given this ignorance (based on the limitation of the mere physical facts) it must be at least logically possible that *I* may persist in the face of total change. In other words, he concludes that I am essentially a mental substance which is contingently joined to a body.

For Swinburne such thought experiments suggest that it is logically possible that a person may be the same person over time even if there is no continuity of mental or physical properties (Swinburne, 2013, p.157). Does Swinburne believe the same applies to non-persons, for example, artefacts? He argues that in the case of artefacts:

> …we normally suppose that only a small amount of replacement is compatible with continued existence of the same substance; and if – whether simultaneously or gradually – you replace many parts, the artefact no longer exists. If you replace one of the four legs of a table the same table still exists. But if you replace the top of the table and also two of its legs, what is left is a different table. (Swinburne, 2013, p.30)

In the famous example of the Ship of Theseus (Brown, 2005) there is total replacement of the original planks of the ship but at the same time the original planks are reassembled to construct a 'new' or replacement ship. Which is the original ship? Swinburne argues we can tell this story in two different ways

because our normal understanding of ‘the same ship’ is too vague. First, in the replacement of the planks the former ship has continuity over time as the same ship, but another ship is formed from the original planks. The second story we might tell is that once all the original planks are assembled into a ship this now becomes the original ship. Both stories are possible without anything being omitted in his putative complete history of the world.

But ‘What is the individuating principle?’ Swinburne’s theory is that some substances (mental substances) have ‘thisness’ which define their sameness over time irrespective of the possession of any set of mental or physical properties.

> A substance has thisness iff there could be instead of it (or in addition to it) a different substance which has all the same properties as it, monadic and relational, including its spatio-temporal relations to earlier and later substances having such-and-such monadic properties and relations. (Swinburne, 2013, p.33)

Mental substances are what they are because they have thisness and the identity conferred by thisness is a ‘brute identity’ and not further definable. Physical substances are what they are because of the properties of their ultimate parts. An individual orange is what it is in virtue of its taste, shape, colour, weight and so forth. These notions are central to Swinburne’s simple theory and he claims are corroborated by thought experiments (See Table Four: Brain Transplant Thought Experiments below). In summary:

TE_1 shows that the extent of discontinuity involved in whole brain transplants is such that the earlier original person cannot be the same as the later person. For example, animalism claims that identity is constituted by physical properties. But it is just not clear how much physical or psychological continuity is necessary or sufficient for maintain ‘sameness’. Swinburne claims, it is not decideable.

TE_2 and TE_2 a show that in cases of brain splitting the preservation of certain brain parts is neither logically necessary nor logically sufficient for personal identity.

TE_3 demonstrates that the participants are necessarily in ignorance of the outcome of the brain swapping. The question of what in these circumstances it would mean to be the same person remains again undecidable.

TE_4 and TE_4 a show that in the cases of incremental brain cell replacements it is logically possible, because of the gradual nature of the change, that P_1 is the same person. But it is also logically possible that P_1 is not the same person. Swinburne claims that this shows that any amount of brain matter does not entail any truth about sameness between P_1 at time T_1 and T_2.

James Dew (2014, pp.30-33) contests Swinburne's arguments derived from the cases of incremental brain replacements (TE_4 and TE_4 a). He argues that there are several physicalist responses that can be made to Swinburne's claims:

> If purely physical organisms can have conscious experience – which is something he [Swinburne] does not seem to say much about – then they may also be able to endure the kind of gradual brain replacement that Swinburne describes whilst remaining conscious and having 'overlapping conscious experiences lasting for the whole operation'. (Dew, 2014, p.31)

Drew suggests there are two reasons which a physicalist could deploy to account for this. The first reason is that the experimental scenario is analogous to the normal metabolic processes by which our bodies gradually replace themselves over time. There is an analogy here between Swinburne's surgical procedure and the body's natural metabolism in which there is no sense that the organism ceases to be the being that it is. This all happens whilst we are conscious and having overlapping experiences. Parts are gradually replaced over time without us ceasing to be the 'same' people. It could be argued on this model that it would just be false to suggest, as Swinburne appears to do, that it would be possible for us not to be the same person.

The second, and allied, line of thought is represented by Lynne Rudder Baker's 'Constitution View':

> Human persons begin existence constituted by human organisms that support first person perspectives. (Later, with enough bionic replacements, a human person may come to have a nonorganic body; but the person is always embodied as long as she exists.) The unity of persons, and of much else, is provided by constitution and the first-person perspective. (Baker, 2014, p.13)

For Baker there are not two kinds of properties, as Swinburne insists, but there are many types of properties but no properties which are only instantiable in immaterial (incorporeal) substances. This does not lead to reductivism often ascribed to physicalist and materialist theories. Baker argues that her non-

reductivist framework embraces both artefacts and natural objects; ‘Being a dialysis machine is as irreducible (and hence as much a part of basic ontology) as being a person or a human organism’ (Baker, 2014, p.13).

Swinburne argues that there are many disagreements between adherents of complex views about what degree of continuity is needed for a person to continue to exist. Such differences arise in classic accounts of personal identity by Locke, Kant and Hume (Swinburne, 2013, p.149, fn.). Amongst his contemporaries Swinburne singles out for criticism the work of Derek Parfit (1984) and Robert Nozick (1981). Parfit’s view of personal identity is that ‘…P_1 “survives as” P_2 to the extent to which there is psychological connectedness and continuity between them, caused by the operation of the same brain’. Personal identity for Nozick is constituted by the ‘…overall satisfaction of both physical and psychological criteria’ (Swinburne, 2013, p.150). The fact of such conflicting accounts, according to Swinburne, tends to lend support to the simple view of identity as constituted by thisness.

Table Four: Brain Transplant Thought Experiments (Swinburne, 2013, pp.151-157).

Thought Experiment	Procedure	Results	Interpretation
TE_1	Person P_1's whole brain is removed and placed into another body.	P_1's body is no longer controlled by P_1's brain; constituting a massive discontinuity.	(i)Too much discontinuity for earlier P_1 to continue to be the same person a later P_2; (ii) Undecidable just how much physical or psychological continuity is sufficient or necessary to maintain 'sameness'.
TE_2	Brain of P is split with half going to P_1 and half to P_1a.	Each will retain some memory continuity; But neither P_1 or P_1a is the same person as P. Possible that: person who has P's left brain is P; person with P's right brain is P or neither are P.	Undecidable: Whether preservation of certain brain parts is neither necessary nor sufficient for personal identity.
TE_2 a	Remove both of P_1's half brains and put each in a skull of a different clone (or twin!) from which half a brain has been removed. Other variations are also available [See p.153].	Yields P_1a and P_1b both 'moderately good candidates' for being person P_1. Possibilities: P_1a is P_1; P_1b is P_1; neither P_1a or P_1b is P_1. Which is a matter of 'arbitrary' definition.	Outcomes are not predictable. Logically possible that following a traumatic operation P_1 loses memory whilst P_1a comes to resemble P_1 in terms of memory and mental life. [N.B. Swinburne assumes P_1 is not the previous body of P_1]

Table Four contd. below.

Table Four: contd.

Thought Experiment	Procedure	Results	Interpretation
TE_3	P_1 is captured by a 'mad surgeon' and told which parts of the brain are going to be transplanted into which skulls: a resulting person will receive a million dollars and the other will be tortured. P_1 is to choose which parts and which reward/punishment. And will want to choose reward for themselves.	Parts of P_1's brain are now installed in P_2's body; and parts are now in P_3's body. But it is unclear who will remain as P_1 in terms of reward or punishment.	P_1 awaits the transplant and knows what will happen to the brain parts; but doesn't know what will happen to 'them' i.e. whether they will be the same P_1. Knowing what happens to brains is not the same as knowing what happens to persons.
TE_4	P_1 is subject to incremental removal of brain parts. P_1's brain matter is removed in increments of 10% each year for ten years (t_1 to t_2) and replaced by another P_2.	After 10 years the whole of P_1's brain has been replaced. Logically possible that because the replacement is incremental then P_1 remains the same person through the procedure. But equally logically possible that they not the same person e.g. P_2.	Hence amount of brain matter does not entail any truth about sameness of P_1 at t_1 and t_2.

Table Four contd. below.

Table Four contd.

Thought Experiment	Procedure	Results	Interpretation
TE_4a	Incremental replacement of brain tissue between t_1 and t_2. But P_1 is conscious throughout the operations.	P_1 is conscious throughout and therefore has a series of overlapping conscious experiences through the successive operations. P_1 at t_1 is the same as at t_2.	(i)Every experience may involve co-experienced events (noise, pain etc.) (ii)Every experience involves an 'owner' of that experience. (iii)When two conscious events overlap they are events entailing the same substance. (iv) We are mental substances that last longer than the specious present. (vi) P_1 is conscious and has a series of overlapping conscious experiences for the whole of the operation. Hence it is logically possible that P_1 survives as P_1 from t_1 to t_2 i.e. personhood does not depend on having a particular brain.

4.2 Limitations of such ‘experiments’.

Can we so easily make sense of this prizing apart of the mental and the physical as depicted in these thought experiments? Bernard Williams, *Problems of the Self* (1982a), imagines the swapping of the brain of an emperor with that of a peasant’s. The experiment simply assumes that memories and character traits of the peasant are transferred into the body of the emperor and vice versa. But what grounds do we have for such an assumption? Williams asks us to consider, for example, just what might happen to their respective voices?

> The voices presumably ought to count as a physical, bodily function; yet how would the peasant’s gruff blasphemies be uttered in the emperor’s cultivated tones, or the emperor’s witticisms in the peasant’s growl? A similar point holds for the features; the emperor’s body might include the sort of face that could not express the peasant’s morose suspiciousness, the peasant’s a face no expression of which could be taken for one of fastidious arrogance. These ‘coulds’ are not just empirical such expressions on these features might be unthinkable. (Williams, 1982a, p.12)

Following Williams, we might let our imaginations run riot and try to imagine the transplanting of Joseph Stalin’s brain into Marilyn Munroe’s body and her brain into his body. Can we really conceive the expressions of Munroe on the face of Stalin?

Williams suggests that in every day contexts (literal rather than imagined!) we recognize and identify people by their ways of expressing themselves involving characteristic ways of speaking, smiling, frowning, tilting the head and so forth. The strong case is that bodily interchanges are inconceivable i.e. logically impossible. We can’t conceive of the outcomes. Williams remarks that ‘…when we are asked to distinguish a man’s personality from his body, we do not really know what to distinguish from what’ (Williams, 1982a, p.12). This would imply that Swinburne’s conception of identity is not even a logical possibility. But, of course, for Swinburne psycho-physical characteristics are not constitutive of a human person’s identity.

First, we a dealing here not with fact but with fiction – products of the imagination; and there is a categorical difference between fact and fiction. We are dealing with imaginative stories rather than literal descriptions which might be true or false. If the latter were the case then we could legitimately ask 'If what you are saying is true, how do you or anyone know that it is?' But since such thought experiments do not claim to be factually true such questions cannot arise. They may have insights to offer about the logic of our concepts; they say something about how our concepts might be pulled or stretched, thickened or thinned.

Secondly, our understandings of the relevant meanings of 'person' and 'same person' are the meanings that have evolved by custom and practice in situations which our forebears have experienced and had to deal with. Swinburne himself acknowledges that there are 'normal' understandings of 'person' and 'same person'. Plainly our language did not develop to address fantastical situations of brain and body swapping experiments. So, there is no reason to expect that these fictional creations could have any stable sense of what, under such circumstances, we might mean by 'person' and 'same person'. Swinburne is probably quite right to stress that in many of the thought experiments the outcomes are undecidable. But this results not from any empirical considerations but from the inconceivability of the enterprise.

Thirdly, the bizarre predicaments envisaged in the thought experiments may in the future turn out to have some truth value. If so we might then need to adjust the meanings of our concepts to cope with new situations and understandings. But this future possibility cannot *now* be relevant or appropriate to the investigation of current meanings of the expression 'same person' or the word 'person'. Even if the speculative circumstances of the thought experiments are conceivable they simply cannot 'jeopardize' our current understandings of 'person' or 'same person'. Peter Geach (1916-2013) points out: 'Difficult cases that would jeopardize the very notion of personal identity if they did arise do not therefore jeopardize it (let us never forget this) by their mere logical possibility even if they do not arise.' (Quoted in Flew, 1987, p.124). Logical possibility is in itself insufficient to justify radical re-definitions of our relevant terms.

Fourthly, it is the case that we have ways of talking about change and 'sameness' in relationship to persons which differ markedly from Swinburne's simple view. There is a primary sense in which personal identity is associated with the physical continuance of flesh and blood creatures such as ourselves. Equally we have ways, in the vernacular, of signalling the fact that incremental and modest changes do take place in individuals. There are circumstances in which we might want to say that 'Doug was not the same person after the operation.' But we understand this as a secondary sense of 'same person' which is dependent on our primary sense. We note that Doug is now somewhat different following his double heart by-pass operation if we are certain that, in the primary sense, Doug is the same person (See Flew, 1968, pp.173-176; Flew, 1987, p.125).

Swinburne has, therefore, more work to do. This must be to turn logical possibility into metaphysical possibility; to show that 'under some circumstances' the simple theory could be true, rather than conceivable (if it is conceivable) and could be '…illustrated by examples and filled out in a more precise way' (Swinburne, 2013, p.15). This is explored in the next section.

5 The Modal Argument to Disembodiment.

As we have seen the conclusion of the thought experiments is that it is conceivable that I might exist without a body. It is not that minds or souls do exist separately from bodies but that it is logically possible that they might exist. The modal argument to the logical and metaphysical possibility of disembodiment then hinges on two key ideas: informative designators (P3) and privileged access (P5) in the summary below. However, do these concepts stand up to scrutiny and deliver what Swinburne wants them to deliver? I summarise Swinburne's argument as follows.

> P1. All the thought experiments discussed are equally applicable to each of us if we substitute 'I' for 'P' i.e. in the sense of what is logically possible.

P2. Such logical possibilities would not be metaphysically possible if we were ignorant of what 'I' referred to.

P3. A logically possible sentence (as argued in Chapter One) is metaphysically possible iff the substances designated in it may be picked out by informative designators.

P4. 'I' is an informative designator when used by me if I know how to use it.

P5. When using 'I' as an informative designator I am immune to error through misidentification. No mistake is possible when applying to a person in virtue of that person being the subject of a present experience.

C1. By being able to use 'I' the following are logically possible:

'I will exist tomorrow with a new brain' or

'I will exist tomorrow without any memory of my previous existence' or

'I will continue to exist when my brain, memory and character are replaced all at once' or

'I began to exist only a minute ago'.

C2. These are also metaphysically possible because I can know merely in virtue of knowing to what my use of the word 'I 'refers.

5.1 'I' - an informative [rigid] designator.

Swinburne argues, therefore, that not only is his conception of personal identity logically possible but it is also metaphysically possible on the basis that:

> …If all the substance properties, and so on, in a logically possible sentence are designated by informative designators, then that sentence will also be metaphysically possible. (Swinburne, 2013, p.54)

If 'I' and my proper name are informative designators, then I can infallibly identify myself as subject of experience even in the 'specious' present. It follows, he claims, that I can continue to exist even without memory,

psychological characteristics, character and so on and without continuity of brain matter.

> To be an informative designator it must be the case that anyone who knows what the word means (that is, has the linguistic knowledge of how to use it) knows a certain set of conditions necessary and sufficient (in any possible world) for a thing to be that thing (whether or not he can state those conditions in words). (Swinburne, 2013, p.10)

But is it correct to claim that 'I' is an informative designator? Kripke's (1980, p.5) original thesis is that only proper names are rigid designators. Swinburne's conditions for words to count as an informative designators expand Kripke's concept. Any word which may be said to 'pick out' i.e. 'name' the same property or substance according to the requirement of the metacriterion is an informative designator (Swinburne, 2013, p.10).

Lynne Rudder Baker (2014, pp.7-8) criticises Swinburne's claim. She argues that if it really is the case that if I know how to use 'I' means that I must know what I am talking about then everyone who uses 'I' would know themselves, by that very fact, to be a mental substance. But this, according to Baker, is clearly false. I can perfectly well deploy the first person singular without believing that I am an immaterial mental substance. Swinburne's rebuttal is that:

> …Baker has not appreciated the sentence in my definition which states that to 'know' the condition for the application of a designator '*as I shall understand this expression'* just is to be able to recognize when it applies and when it does not; knowing the necessary and sufficient conditions for a thing to be that thing is knowing the nature of what I am talking about. (Swinburne, 2014, p.52)

His defence is that 'I' as informative designator works in the specific context of the arguments and thought experiments for the logical possibility that 'I am a mental substance'. We can know certain things without necessarily fully accepting all their logical consequences. I'm not convinced that this is an adequate response. In effect it can be read as partly acknowledging the thrust of Baker's argument. What then is it to 'recognise' the correct application of a designator? Surely this implies I might be mistaken but the point of Swinburne's argument is supposed to be that I can't be mistaken.

Baker generalises her point by suggesting that we have linguistic knowledge of how to use many concepts without in the least knowing their relevant set of necessary and sufficient conditions (in any possible world). For example:

> I don't even know any necessary and sufficient conditions for a thing to be red (Swinburne's example of an informative designator). Is a reddish-orange poppy red? Is a reddish-yellow bruise red? Of the words that I use competently, there are not many (if any) for which I know any necessary and sufficient conditions. (Baker, 2014, p.8)

The utility of the idea informative designators, as Swinburne defines them, are beset by indeterminacies. This must apply particularly to the notion that 'I' is such a designator for the reason that it is relative to its use in time and relative to the speaker. And such indeterminacy, as Baker points out, must '…vitiate the usefulness of informative designators for telling the whole story of the world' (Baker, 2014, p.8). 'I' just cannot conform to Swinburne's criterion of being capable of 'picking out' the same substance or property. As many philosophers have pointed out it cannot stand for the name of some putative self nor, more widely, can it be said to pick out the 'same' things (Anscombe, 1975; Glover, 1988; Kenny, 1984; Russell, 1958).

The reasons for this are, as Jonathan Glover argues, 'I' functions as an indexical in common with words such as 'this', 'here' and 'now' (Glover, 1988, pp.66-67). It varies in what it might refer to depending on the features of the context in which it is used:

> What indexicals refer to depends on context. What makes up 'the scene in front of me' depends on the way I am facing when I refer to it. The place 'here' refers to depends on where the word is spoken. And the time 'now' refers to varies with when the word is used. 'I' is also an indexical and the person it refers to depends on who is speaking. (Glover, 1988, p.66)

So 'here' does not name a particular place, 'now' does not name a particular time and the fact that some disembodied mental substance can mysteriously utter the word 'I' cannot name a human person. If reference varies with the speaker we must already know the speaker for this to make sense. Whereas Swinburne wants the uttering of 'I' to indubitably establish the very existence of the speaker without question or corroboration.

Swinburne's defence of the notion that 'I' is an informative designator rests on the notion of privileged access and operates differently from other indexicals. He recognises that generally what is designated by indexicals may vary depending on the speaker and where and when they are being used. However, in the case of 'I' when used by me as a pure mental substance, I have privileged access to the instantiation of 'I' in my case. This means that correct use of my name or of 'I' by myself is an informative designator; I cannot be mistaken. In contrast if someone else were to use my name correctly it is an uninformative designator (Swinburne, 2013, p.165; Swinburne, 2014b, p.53). But this seems to radically suspend the normal usage of 'I' in favour of a usage heavily infused with Swinburne's ontology and epistemology.

If we think that 'I' is a name, we must ask what kind of thing does it name? Swinburne believes that it names an incorporeal mental substance. Thus we are led to the 'unhelpful invention of a mysterious self"' (Glover 1988, pp.67-68). But the grammar of self-reference does not sustain such a conclusion. For example, we refer to the self-same suitcase, to the rose itself, or my attempts to get myself fit; none of this surely gives grounds for supposing that 'self-same',' itself', 'myself' name some entity that is the self of the suitcase, the rose or me! In the latter case we do seem to want to give the reflexive pronoun a special magic. But this is Swinburne's case.

The nerve of Swinburne's argument seems to be the combination informative designation and privileged access implies that I am immune from error in identifying an incorporeal mental substance that is myself by the use of 'I'. We have already shown (Chapter Four, 'Privileged Access') that although it is correct to say 'only I experience my experiences' this fact of ownership does not come with the diploma title of 'immunity to error'. Swinburne's move using informative designators is fallacious in the way that the ontological argument is fallacious. The idea of informative designators, as defined above, play the role of 'perfection' in Anselm's ontological argument for the existence of God. Swinburne is deriving the metaphysical possibility of a fact from verbal usage. But Swinburne's claim is that '…an event is metaphysically possible iff it could happen under some circumstances' (2013, p.15). My claim is that his conception of identity cannot be true under any circumstances because it simply cannot be derived from verbal usage and privileged access.

One cannot begin with the designator and then indubitably infer the existence of some putative entity. I may know the sufficient and necessary conditions which may designate a unicorn, but this provides no warrant for the existence of unicorns. Gaunilo replies to Anselm, 'on behalf of the fool,' pointing out the fallacy with the story of the lost island. We might posit the existence of a 'perfect' island and argue that the idea of a perfect coast belongs (necessarily) to the concept of a perfect island. But none of this implies the existence of a perfect island (Cox, 2017, pp.75-80; Daniel, 2006, pp.54-56). In the same way I suggest Swinburne is arguing from the concept of a thing to its existence under some circumstances. He has no warrant for this.

For example, in order to 'pick out' a mental substance there must be a mental substance in existence for the information designator to pick out. Now Swinburne has already established to his own satisfaction that there are mental substances. But the fact that I know the sufficient and necessary conditions for application doesn't of itself entitle me to infer the existence of mental substances as he seems to be suggesting.

> A rigid designator *φ* is an 'informative designator ', iff anyone (when favourably positioned, faculties in working order, and not subject to illusion) can recognize when something is (now) *φ* and when it is not merely in virtue of knowing what the word '*φ'* means (that is having the linguistic knowledge of how to use it) … and can make simple inferences involving *φ*. (Swinburne, 2013, p.158)

The identity of 'I' with 'being mental substance' just cannot be established on the grounds that 'I' is an informative designator. And, as we have seen, the privileged access argument can't perform that function either. But this is what Swinburne is claiming: '…No mistake is possible when applying to a person in virtue of that person being the subject of a present existence'. This looks like an illegitimate attempt, using a 'stipulative' definition of a word', to deduce actual existence.

Like the ontological argument the trick is to list 'existence' as a necessary attribute of God and then deduce his necessary existence from the definition. (But no logical contradiction is involved in asserting that God does not exist and therefore the proposition that 'God exist' cannot be necessary!) In Swinburne's

argument informative designation is similarly meant to build in the notion that if I have the linguistic knowledge to apply a designator then what is denoted must exist. I am suggesting that 'informative' here plays the role of 'existence' in the ontological argument. Since 'I' and my proper name are informative designators I can infallibly identify myself as a subject of experience and thus existence. I can continue to exist even without memory and brain matter. But this is to confuse questions about the meanings of words with questions of contingent existence.

The argument rests on asserting a stipulative definition which provides meaning for Swinburne's radical new sense of what it means to be the same person over time and which supports the simple theory of identity. Having shown, at least to his own satisfaction, that a person is an incorporeal, mental substance he must show how such a vaporous concept can be identified and re-identified. Uttering 'I' in some logical, non-physical space seems to be his solution.

5.2 Paradigm cases for 'Person' and 'Same Person'.

As we have seen Swinburne believes he has established that disembodiment is a metaphysical possibility. This is intrinsic to his simple view that personal identity is compatible with having any particular mental properties or physical properties except that I am a self-reflective mental substance. Further evidence for this view he believes are paradigm case arguments for the meaning of 'person' and 'same person'. He constructs and contrasts two stories about how we come to grasp the nature of our identity. In the first story Swinburne does not deny we can construct concepts of personal identity based on some degree of continuity of our brains or (physical) behaviour. These he asserts are 'natural concepts' derived from '…paradigm examples of one earlier person being the same as some later person, when neither of these persons is oneself, together with paradigm examples of what each of us remembers in respect of the content of what the remembered person did and experienced at various past times' (Swinburne, 2013, p.159) This story is characterized by inference from a combination of public observation and personal recollection.

He proposes an alternative story based on privileged access. Our *actual* concept of personal identity must be derived from such paradigm examples of our memories but, crucially, from who remembered these things and from certain examples of present experiences (…'including experiences of conscious events lasting some time longer than that of the specious present').

> P1. Key paradigm experiences are co-experienced and such experiences must show that experience of successive overlapping events must be had by the same (mental) substance (oneself).
>
> P2. From this we derive our understanding of personal identity over time as being 'a particular subject of experiences' and hence a mental substance existing over time.
>
> C. Our sense of personal identity is not derived from physical continuities, continuity of memory or any other mental events but from being a co-experiencing self.

Swinburne's simple theory gives us a strange 'attenuated self', in Williams' phrase, without body, past or character.

> It is not entailed by a full description of the world in its physical aspect and in respect of which bundles of mental properties are co-instantiated (at a time and over time) that I, picked out as the actual subject of certain mental properties, should have the body and the particular mental properties which I have; or that the person who has that body and those properties should be me. (Swinburne 2013, p.164)

If this is meant to be an account of how we ordinarily come to understand our identity it is surely wrong. It doesn't necessarily follow from the fact that we are 'the subjects of our own experiences' that we are 'a mental substance existing over time'. More strongly this does not even seem to be logically possible in the sense that what Swinburne suggests is not a conceivable description of persons as they are defined and identified.

Dew (2014, p.36) draws out the bizarre (nonsensical) consequences of Swinburne's conception that the sameness of persons depends solely on immaterial 'thisness'. Suppose we imagine two persons: Daphne and Velma. Daphne has the mental properties, 1,2, and 3 and physical properties X, Y, and

Z, at time t_1; and at t_1 Daphne has the thisness of Daphne. But suppose at t_2 Daphne loses all these mental and physical properties and acquires a new set 4, 5, and 6, and A, B, and C. In parallel our second character, her friend Velma, at t_1 had the mental properties 4, 5, and 6, and physical properties A, B, and C, and the thisness of Velma. Subsequently, like Daphne at t_2, Velma loses her existing mental and physical properties and acquires mental properties 1, 2, and 3 and physical properties X, Y, and Z. Yet Daphne and Velma (on Swinburne's theory) would retain their respective thisness; Daphne at t_1 = Daphne at t_2 and Velma at t_1= Velma t_2 in spite of the loss of all their previous mental and physical characteristics. Dew reasonably supposes that Daphne and Velma can only be extremely confused by such a state of affairs. It seems reasonable to suggest that Daphne would now believe that she is Velma since she now has all of Velma's original and mental properties. And Velma must likely believe the same. It does seem that they would think they are someone they just are not.

The moral of the story seems to suggest that we just cannot identify ourselves across time without the coordinates of at least some continuance of our mental and physical characteristics – exposing the emptiness of the idea of thisness which offers no sense of identity. The case is markedly different from Locke's tale of the Prince and the Pauper where the bodies are switched but they retain their respective mental properties (desires, memories, dispositions and so on). As Dew argues the Prince might be confused as to how he got into the body of the Cobbler; he may not be confused about his identity; he can remember who he was as the Prince and the life he lived. No such reassurance is available to Daphne or Velma.

We could say without self-contradiction that we are physical persons with the capacity of co-experiencing. The paradigms by which we learn the meanings of 'person-word' are precisely through their application to particular, embodied, physical persons. Our sense of ourselves and others as having continuity over time rests largely on physical continuity. For example, person-words such as 'I', 'person', 'people', 'man', 'woman', 'Horner', 'Hill', although performing very different linguistic functions refer in one way or another to material objects. What else could they refer to? Well Swinburne says some none physical thisness of a mental substance. But this is hardly intelligible! Even in Swinburne's account I have to be able to articulate 'I' which is part of a public

language which I must have learnt by reference to paradigm cases of its use in referring to physical objects.

Acquiring linguistic knowledge (essential for Swinburne's theory of informative designators) involves reference to (public) objects which we can see, point at, hear, talk to and so forth:

> Both the word 'person' itself and all other person-words – such as the personal pronouns and terms like 'butcher' or 'politician' picking out members of classes of functionary – are themselves words for members of a kind – our own kind – of creatures of flesh and blood. (Flew, 1986, p.91)

Flew here also touches on how we might see the relationship between our two senses of identity: particular and type. It seems that selves cannot be the objects of discourse without reference to particular persons' states of consciousness, thoughts and emotions and these are inherently and paradigmatically states of material flesh and blood creatures. The paradigmatic case for learning to talk about cognitive and affective states is simply though the people we meet…'everyone's paradigms of what persons are are members of our own particular species of flesh and blood organisms' (Flew, 1987, p.118).

Swinburne's counter, from the principle of credulity, might be that we ought to believe that things just are as he claims they appear to be - if there is no counter-evidence. Thus:

> …it is such a strong deliverance of experience for all of us that we are aware of ourselves as continuing mental substances for very short periods of time during which we seem to recall overlapping experiences that it is hard to see that there could be any counter-evidence to rebut this result, that we really are such subjects. (Swinburne, 2013, pp.160 -161)

It would amount to saying that 'having overlapping experiences' means being 'a mental substance'. But it is not obvious that in fact we are 'aware of ourselves as continuing mental substances' (unless in the grip of theory). Elsewhere Swinburne admits as much: for example, '…we are not (normally) aware of interacting with our brains' i.e. we just do not experience ourselves as pure mental substances interacting with brains (Swinburne, 2013, p.173).

We are more likely to think of ourselves as flesh and blood creatures with memories, personal characteristics, a biography, aspirations about the future

and so on. On the basis of the credulity principle in fact we ought to reject Swinburne's claim that '…we are aware of ourselves as continuing mental substances'.

It is not clear either that the theory of informative designation, nor the argument from paradigm cases of personhood, establish the metaphysical possibility of the simple theory of identity. Swinburne's definition of identity is that each person is a pure mental substance having a thisness which is non-physical and which we may call a soul. This further stage is briefly examined in the next section.

5.3 Re-defining 'Person' and 'Same Person'.

Swinburne corroborates his earlier distinction between 'persons' and 'humans' (Section 2 of this chapter) by formulating a high-redefinition of personal identity.

> The continued existence of a person over time consists in the continued existence of a mental substance; and it is metaphysically possible that that substance acquires a totally new body, totally new apparent memories and character. I and each of my readers, being persons, are mental substances in this sense; and having at some time a kind of body and an ancestry similar to those persons called "human" today, we are human persons. (Swinburne, 2013, p.163)

He has already established to his own satisfaction the preconditions for this version of substance dualism.

> P1. Pure mental events do not supervene on physical events;
>
> P2. Mental events last for a period of time and that what makes it one event is that it is constituted by two half events that are co-experienced by the same person;
>
> P3. A person can have successive pure mental events; and
>
> P4. A person's identity is independent of physical continuity and any particular bundle of mental properties.

It is logically possible that I can lose all contact with the physical world and yet continue to have thoughts and feelings. And because 'I' is an informative

designator this logical possibility can be metaphysical possibility; a mental property is all that is necessary for existence.

Our uniqueness as persons is captured by thisness because I can 'pick out the same myself' irrespective of any particular sensations, visual, tactile etc., and I would be the same. Thisness is independent of, and indifferent to, any particular body and physical properties and any that might be possessed by physical matter. Swinburne qualifies this by asserting that '…having a body and bodily well-being are not unimportant'. Having a body is important for having a worthwhile human life with its attendant pleasures. Normally everyone 'on earth' is composed of a body and a soul (Swinburne, 2013, pp.170 – 173).

The word 'soul', in Plato's sense, denotes the essential part of a human being.

> P.1 Given that it is metaphysically possible that I can be disembodied;
>
> P2. and given that some earlier person who has all the same physical parts as me as well as all the psychological characteristics as me could equally well be me *or not me*;
>
> P3. The difference must consist in the presence or absence of some non-physical part – which is a pure mental substance which makes me, me;
>
> P4. I need only a pure mental part (given that I can be disembodied) my soul, in order to exist;
>
> C1. My soul carries my thisness.

Swinburne's view of the soul here differs from his earlier formulation in *The Evolution of the Soul* (1987, p.171). There the soul was characterized as the sort of stuff which is potentially divisible. He now believes souls must be indivisible ('simples') (Swinburne, 2013, p.173). He responds to traditional questions about substance dualism: 'How are the body and the soul related? 'Where is the soul located given that it is incorporeal?' First, the soul interacts with the body and hence the world by operating via the brain and can only function if the brain, with which it relates, functions normally. To the second question he acknowledges that non-physical substances cannot (by definition) 'occupy' space i.e. they cannot 'fill' space nor exclude other (physical) things.

But, he claims it is ‘natural’ to identify the soul’s location by where it interacts with the world.

> The soul of a human being living on earth is located (in one sense) in that human’s brain and (in another sense) in the whole of that human’s body. But under abnormal circumstances (e.g. at death) [sic] my argument shows that it is (metaphysically) possible that the soul could become connected to (i.e.) interact with a new body or exist (and even have a conscious life) without a body. (Swinburne, 2013, p.173)

But this hardly addresses the force of the problems. For example, if the soul does not occupy space what sense can we make of it being ‘located’, if only temporarily in the brain or to say it is located where it interacts with the brain? Similarly, just how does a soul occupy a new body’? ‘What could possibly be the mechanism for such a transition of an ethereal self? If thisness is the guarantor of a person’s particular identity, stripped of tangible characteristics, just how are we meant to first, identify it and then re-identify it? The incoherence of all this, I believe, renders his picture of the soul (mind) as something neither logically possible much less metaphysically possible.

6 Conclusion.

Swinburne radically redefines the meaning of ‘person’ and ‘same person’. What justifies us in saying that person Y at time t_2 is the same as person X at time t_1? Broadly Swinburne says that Y at t_2 is the same as X at t_1 iff X and Y are both humans and are pure mental substances (souls) characterized by a particular thisness. This answer assumes as settled his ontology and epistemology. The hinge of the argument is the move from logical possibility of persons as pure mental substances to their metaphysical possibility. First, this is accomplished by the use of ‘I’ as an informative designator; we can know that we are the same persons merely in virtue of knowing to what our use of the word ‘I ‘refers’. Secondly, our grasp of the meanings of ‘person’ and ‘same person’, derive from paradigm cases, showing that our identity must derive from a co-experiencing self, a pure mental substance.

Swinburne relies on his metacriterion to clinch the argument. A complete history of the world must contain histories of human bodies and human souls. It is metaphysically possible, he concludes, that there can be souls without bodies and bodies without souls. However, '…under normal circumstances of earthly life' we need not suppose that souls and bodies come apart. I have argued that the move from logical possibility to metaphysical possibility through informative designators and privileged access can be shown to be fallacious.

Further Swinburne's view of personal identity, the simple view, is radically at odds with our normal understandings. We argued that people are characteristically those whom we meet in physical space and our understandings of self and selfhood are constituted by the public language that has evolved to describe our (physical) human situation. The two-substance account, that our bodies are contingent and our incorporeal souls essential (our guarantors of sameness), gives rise to what would seem to be the unsolvable problems of continuity, identification and reidentification of bodiless persons. Our normal paradigm examples of continuity just do ultimately rest on reference to physical continuity; a person at time two is the same person as at time one on the basis of the former being materially continuous with the latter.

Swinburne's description of the soul leaves us with a highly attenuated sense of personal identity. His picture of the residual self seems to strip it of all the normal characteristics which we would take to be the very constitution of a person. His essential self has no permanent physical or psychological characteristics: it is an 'I' which can acquire a totally new body, totally new apparent memories, and a new character. It is, in effect, mere empty form. This seems the very negation of what it is to be a person; it is difficult to see just what, in these circumstances what 'thisness' amount to. Why would this entity (if such it be) be of any interest or concern to *me* since its minimalism can hardly be construed as *personal* in the sense that we understand it?

Chapter Seven – Summary and Conclusions

1 Aim and Objectives Revisited

My aims and objectives, set out in Chapter 0ne, were to evaluate Swinburne's arguments for substance dualism developed in the first six chapters of *Mind, Brain, and Free Will* (2013). In the Preface Swinburne reflects that the conclusions he reaches remain broadly the same as those reached in *The Evolution of the Soul* (1987). One difference, he notes, is in *Mind, Brain, and Free Will* he concludes that the soul is a 'simple' that is a non-physical, indivisible, pure mental substance. In the earlier book the soul is a non-physical, mental substance but is divisible, in other words, it may have parts (mental characteristics). However, the main differences between the later and the earlier book lies, he believes, in a fuller, deeper treatment of the philosophical arguments for substance dualism. For example, the identity criteria of events and substances have new underpinnings from a metacriterion and the theory of informative [rigid] designation. Another significant innovation is the grounds by which certain states of affairs (such as disembodiment) can be shown not only to be logically possible but also metaphysically possible. These theoretical innovations, combined with his privileged access doctrine and principles of credulity and testimony, constitute his case for substance dualism and a radical view of personal identity. The argument of this dissertation is that in spite of the 'fuller' and 'deeper' treatment the arguments fail to be compelling.

2 The Mind/Body Problem.

Swinburne's intention in *Mind, Brain, and Free Will* is to show that a person is pure mental substance: and '...we may call this non-physical part of me 'my soul' (Swinburne, 2013, p.170). The individual soul is our 'thisness' constituting our identity and continuity over time. The soul must then be the basis of our

identification and re-identification. Our identity, according to Swinburne, is not dependent on any physical properties or indeed any specific mental properties. I can pick myself out as such a pure mental substance by the informative designator ‘I’ or a name which is my name for myself. Although our normal human state takes an embodied form Swinburne’s case is that, at the very least, it is metaphysically possible that I, as a pure mental substance, can exist in disembodied form.

Swinburne criticizes a variety of trends in contemporary philosophy of mind. These are, on the whole, tokens of physicalist type theory including identity theory; functionalism; eliminativism; instrumentalism; and event identity. Swinburne rejects such theories on ontological and epistemological grounds. Physicalism is inadequate for a complete description of the world in not registering a distinction between physical and mental substances, properties and events (a metacriterion). Epistemologically phenomenal truths are not reducible to, or identified with, physical or microphysical truths. The world is not a causally closed physical system and there is a causal nexus between phenomenal and microphysical properties. I identify three principal elements Swinburne relies upon to ground substance dualism. These are a metacriterion, a privileged access principle and the credulity and testimony principles. These elements combine to produce Swinburne’s vision of personal identity as ultimately pure mental substance.

3 A Metacriterion.

The criterion is ‘…the requirement that our vocabulary should enable us to tell the whole history of the world.’ We must be able to fully describe the data of human mental and physical life to provide a coherent explanation of human identity. Swinburne believes that the establishment of a set of canonical descriptions provides an account of the world in terms of substances, properties, events and times. These are assumed to be the fundamental ontological constituents of the world. The terms in this canonical vocabulary will be such that a competent language user would be able to identify which substance (mental or physical), property and event is involved. The words

which pick out substances, properties, and events, according to Swinburne, are informative [rigid] designators which designate the same object ‘in every possible world’ in which it exists. Objects are characterized by some essential property. Identity conditions established in this way underpin substance dualism.

First, even if we identify things as mental or physical why should we assume that they are: (a) independent of each other; and (b) that one might exist without the other? Suppose we can informatively designate lungs on the one hand and the process of absorbing oxygen on the other as distinct. But lungs and the absorption of oxygen are clearly related in a functional way. The process of absorbing oxygen can happen because it is realized in a physical structure with appropriate causal abilities. It doesn’t follow that somehow the absorbing of oxygen can happen in some ethereal way without lungs or other similar physical structures. Similar if we consider the example, much beloved of the philosophers of mind, the firing of C-fibres and the experience of pain, this seems analogous to the lung and absorption of oxygen example. On Swinburne’s model the firing of C-fibres and the experience of pain are clearly informatively designated as mental and physical concepts. But it would, on that basis, appear equally foolish to suggest that pains can be identified floating mysteriously free.

Secondly, this strategy depends on a contentious theory of designation. For Swinburne informative designators are a means of reforming natural language by making identity conditions more precise. However, this fails to make the distinction between increasing precision and the representational power of natural language evolved through dealings with the ordinary, practical matters of life. The theory of designation is a reduction in the power of language because it trades on a false model of language consisting of naming of subjects/predicates (substances/properties). ‘I’, ‘you’, ‘me’, ‘this’, ‘that’, for example, are not names but index words. This may be part of the story but if the intention is to provide a complete history of the world required by the metacriterion don’t we have to take account of interrogative, imperative and optative sentences and much else besides?

Thirdly, informative designation seems to rest on the notion of naming substances because they are ‘natural kinds’. This is the claim that the essence of an object (broadly interpreted) can be identified by its composition. But this is surely false. It is a far too simplistic account of how language evolves, and usage defines meaning. There is no reason to suppose that what a thing consists of constitutes the essence of a thing and there is no reason to suppose that the structure of language maps the structure of the world. There are no essences in the sense that Swinburne needs to support the project of a canonical vocabulary which would ‘…tell the whole history of the world’. The theory of informative designation presents a picture of the way language works which is greatly oversimplified if not erroneous. It is at best contentious and at worst simply false.

4 Privileged Access.

According to a widely accepted principle, privileged access, we each have unique access to our own mental goings on. Swinburne interprets this as an epistemological principle to the effect that we each know our mental states in a special way not open to others. Only I know what it is to experience my thoughts, feelings, intentions, sensations and so forth. Others can only infer such things about me. Swinburne uses this principle to elaborate the nature of mental substance. The epistemological privilege of this mode of access supposedly guarantees that we know our ourselves with some certainty in a way we cannot be known by others. This is supposed to be further evidence for the dualism of mind and body; mind is a non-physical space in which I witness the occurrence of my uniquely private mental events.

First, we can say, following J.L. Austin (1970), that this way of thinking confuses a distinction between two sorts of questions: (a) epistemological and (b) of ownership. Questions of the first kind are different in meaning from questions of the second kind. To ask, ‘How do I know that Bill is upset?’ (because sometimes I surely do know that Bill is upset), does not mean the same as ‘How do I introspect Bill’s mental states?’ In the former case I can pretty much know

that Bill is upset in the way that I know what it is for me to be upset. Ordinary language provides us with a plenitude of ways of describing and talking about our own and others' mental states. We 'see', for example, a person's emotional state written raw across their face or in their gestures. This doesn't necessarily involve an act of inference in the way that Swinburne suggests. In the latter case when we introspect naturally only I can introspect my thoughts. I am usually aware that these are my thoughts but that is straightforwardly a matter of ownership; grammatically, only I can experience 'my' thoughts'. It doesn't mean that I am infallibly certain of what I am feeling or thinking. We can be, and often are, mistaken or confused about our mental states. Others may be in a better position, say if I am grieving, to understand what is happening to me. Ownership simply does not guarantee, to use a colloquialism, that 'I know my own mind'.

Secondly, Gilbert Ryle (1976) robustly attacks the privileged access doctrine by dissolving the antithesis between public and the private worlds in showing that the opposition rests on a misreading of the logical grammar of our everyday mental concepts. Ryle's work is particularly useful when applied to Swinburne's tendency to observe symmetry between the world of physical substances, properties and events and that of mental substances, properties and events. For example, Swinburne's commitment to mental substance leads him to the notion that we can talk about 'the first half' of an experience so presumably he believes it makes sense to talk of fractions of experience of mental events. I can ask for 'half a bucket of water' but it's nonsense to ask for 'half a bucket of experience'. Experience is not in the category of things we can measure. I argue that we ought to prefer Ryle's account of the mind/body problem precisely on the grounds that he sets out. He shows through a detailed analysis of mental concepts, that '…the official theory does rest on a batch of category mistakes by showing that logically absurd corollaries follow from it. The exhibition of these absurdities will have the constructive effect of bringing out part of the correct logic of mental concepts' (1976, p.24).

Ryle's case is that the mind is not any kind of substance like the body. To assume this substance dualism is to end up in all of sorts of nonsense much like the Cheshire cat in Alice which gradually disappears leaving only its grin behind. To treat the mind as a kind of occult, incorporeal repository to which I

have privileged access is to commit a fundamental category mistake or indeed many such mistakes. Mental events are not some category of events in the way we talk about physical events. (This, for example, leads Swinburne to important errors in treating 'mental' causation as on all fours with physical causation.) The word 'mind', I think Ryle successfully shows, is not an informative designator for a substance. After all we do not stumble across mental substances in the way we might stumble across a body (or at least only metaphorically). 'Mind' is a kind of short hand for describing, and knowing about, people's dispositions, abilities and capacities in much the same way that Austin suggests. I conclude that Swinburne's sense of 'privileged access' falls foul of the kind of category mistake that Ryle suggests.

Thirdly, Wittgenstein famously advances the 'private language argument' to contest the solipsistic implications of the privileged access doctrine. Mental events (Swinburne's term) are private in the very particular sense of being accessible and therefore knowable directly, only to the person who has them. This would appear to lead to a notion of 'the self' locked into the circle of its own thinking (but what would thought be in such circumstances?). Wittgenstein's analysis removes anything that we might claim privileged access to. Anything we think about and express in language must be expressed in a common, public language which is publicly acquired. Having my own private language is logically impossible because there simply is no way of checking whether I'm using terms now in the way I used them before. Private mental objects must drop out of consideration as irrelevant. Swinburne attempts to challenge Wittgenstein's argument by suggesting that I could check against apparent memories. But this fails to address the core problem that there is no standard against which such apparent memories might be checked. Austin, Ryle and Wittgenstein's show the claims of privileged access doctrine to be deeply problematic.

5 Credulity and Testimony.

An important problem for Swinburne, having by his own lights established the cogency of a dualist account of the physical and mental life of human beings, is ‘How can mental and physical substances interact given their categorial difference?’ ‘How can an incorporeal, non-physical, non-extended substance (the mind) interact with corporeal, physical, extended substance (the body)? Swinburne’s claim is that interactive dualism just does reflect our practical everyday experiences referred to as ‘the normal view’. For example, it is a commonplace of our experience that our intentions frequently lead to actions which affect the world; this constitutes evidence of interaction. He attempts to show that the normal view can be demonstrated to be correct by defeating counter-theories. On principle non-physical minds (mental substances) cannot plausibly have energetic interactions with physical bodies without breaching fundamental physical laws: the principle of the causal closure of the physical (CCP). A further challenge to the normal view is evidence gathered from experimental subjects that brain activities may precede the conscious awareness of an intention to act.

Swinburne argues that alternative theories to the normal view fall foul of his principles of credulity and testimony and consequently the normal view must be correct. These principles are fundamental, *a priori*, and epistemic. In the absence of ‘special considerations’: (a) if it seems P experiences the presence of X then probably X is present to P; (b) on the testimony of person P we ought to believe that P did have an experience of X. Swinburne believes that the two principles show the alternative theories to be self-refuting. In effect the counter-theories, such as CCP and those of Libet, imply that there are always (and only) physiological reasons (brain events) *determining* the experiences we have or the utterances we make. If this is the case then it seems to follow that we cannot have (and know that we have) sound reasons for believing the experiences and the testimony of, for example, experimental subjects. A similar logic also will apply to the observations and deliverances of experimental neuroscientists. The only reason for believing the propositions of our testimony would be if the claim to physical determinism (which is at the heart of the counter-theories) is false. The evidence required to show the alternative theories to be true can only be acquired if the theories are in fact false: the evidence paradox.

However, I argue that the two key principles do not have the sway that Swinburne gives to them. In short, they are not fundamental, nor *a priori* and, crucially, not epistemic. Given that they operate with the caveat 'in the absence of special considerations' it is hard to accept them as fundamental. The special considerations discussed are hardly 'special' but rather 'reasonable' and are likely to apply in most circumstances where the principles are being applied in contentious or complex contexts. They are not *a priori*, that is not derived from reason alone, but rather based in pragmatic motives to lubricate the wheels of discourse. We need to start somewhere. Most importantly they do not seem to be epistemological at all in the sense that they are not directly concerned with establishing the grounds for justified, evidentially true beliefs. They are rather *prima facie* moral principles which urge, reasonably enough, that we *ought*, in the first instance, to accept what people testify to or claim to be their experiences. But they are not themselves the necessary and sufficient grounds for knowing something to be the case. They cannot do the job that Swinburne thinks they can do in overturning the counter-evidence and establishing 'the normal view'.

6 Personal Identity.

The metacriterion, the privileged access doctrine and the principles of credulity and testimony converge in a radical redefinition of personal identity. What makes a 'person' the 'same person' over time? Swinburne argues that person Y at t_2 is the same as person X at t_1 iff X and Y are both humans and are pure mental substances (souls) characterized by the same thisness. Each person's thisness is particular and distinct. This he refers to as 'the simple view' as compared with 'complex views' which grounds identity in various combinations of psycho-physical characteristics. Swinburne believes that the logical possibility of persons as pure mental substances (the simple view) can be further established by a variety of thought experiments which show that disembodiment is conceivable.

The challenge then is to show that logical possibility of persons as disembodied, pure mental substances, is also a metaphysical possibility. This is effected using the metacriterion and 'I' as an informative designator. A complete history of the world must contain histories of human bodies and human souls. It is metaphysically possible, he concludes, that there can be souls without bodies and bodies without souls. However, '…under normal circumstances of earthly life' we need not suppose that souls and bodies come apart.

Two kinds of attack are open here. First, the move from logical possibility to metaphysical possibility is mistaken. Swinburne is wrong to assume that 'I' is an informative designator and that 'I' can name something, thisness, which is beyond our ordinary physical and mental characteristics. Pronouns like 'I', 'you', 'me' are index words and do not function as names. 'I' just cannot be the name of some elusive, ghostly selfhood in the way that 'this' cannot be the name of a place. 'I' is not an alias for David Horner or Richard Swinburne. Its linguistic function is to indicate the person who David Horner and Richard Swinburne refer to when they use 'I' to indicate themselves. And David Horner and Richard are all too tangible creatures of flesh and blood and known to be such.

Secondly, it is plausible to suggest that it is not even conceivable (logically possible) that persons are 'pure mental substances'. It is difficult to understand just what this means. Talk of non-physical (mental) substances seems like talk about 'immaterial material'. Swinburne's view of personal identity, the simple view, is radically at odds with our normal understandings of persons and the words we normally use to characterise persons. Persons are characteristically flesh and blood creatures whom we meet in physical space. Our conceptions of ourselves and others are constituted by a public language that has evolved to describe our all too human situations. This is not to suggest that our ordinary way of talking about people is to some extent flexible and capable of incorporating talk about persons and bodies as if they might be separate. But such talk is often of a purely metaphorical kind. My argument is that if we try and take seriously the notion of persons as somehow existing independently from their bodies I can't even begin to conceive of how that can make sense. In spite of his elaborate defence of dualism Swinburne gives us no inkling of just how we (here and now in this world) might identify and re-identify bodiless persons. Of course, none of this prevents our imaginations dreaming of

uploading ‘selves’ into computers or transporting ‘persons’ instantaneously through space. But the fact, rather than the fiction, is our standard paradigms of identification and continuity rest on reference to physical continuity; a person at time two is the same person as at time one based on the former being materially continuous with the latter.

Swinburne’s conception of the soul as pure thisness is of something without identifiable characteristics. It is an empty ‘I’, a formal, fluid self which might ‘acquire’ a new body and acquire new apparent memories, and a new character (but what can ‘acquire’ mean in this ethereal context?). This seems the contradictory of what we mean by ‘being ourselves’. This can be of no interest to *me*. This minimalism can hardly be understood as *personal* or conceivable as a person.

For these reasons it seems Swinburne takes us on an interesting journey we end up in a blind alley. The arguments he makes for his brand of substance dualism are not compelling. It seems to me that the situation is not retrievable. In retrospect I suggest that his first step on the journey is the perhaps the fatal one. This is his fundamental and unshakable assumption that it makes sense to talk about immaterial (mental) substances. In spite of the plethora of argumentation he fails to make this claim stick with the consequence that his ultimate concept of what it is to be a person seems utterly formal and without content. Of course, this doesn’t necessarily imply that there may not be other approaches to mind/body dualism, argued from different premises, that may be more successful. For example, the work of Barry Dainton (2014) is an imaginative attempt to continue the dualist tradition albeit in a very different form to that of Swinburne.

Swinburne uses the arguments for substance dualism to derive a concept of free will. His definition of free will is that an agent has free will ‘…insofar as the agent acts intentionally without their intentions being fully determined by prior causes’ (Swinburne, 2013, p.202). The definition, it seems to me, is at variance with a range of other views in the field. This interpretation of the idea of free will draws specifically on arguments, especially concerning intention and causation (see my Appendix: Causation and Free Will. It especially hinges on a crucial but contentious distinction between intentional and nonintentional

causation. It would be interesting perhaps to pursue Swinburne’s account of free will and moral responsibility in the light of criticism already made of his arguments for substance dualism. Can it survive?

Appendix – Causation and Free Will

Swinburne's ontology leads him to interpret mental events paradigmatically as if they are physical events. This is what Ryle (1976) famously calls 'the para-mechanical hypothesis'. Swinburne's commitment to a para-mechanical interpretation of interactionism is reflected in his commitment to substance-powers-liability (SPL) theory of physical causation and agent causation. This is the basis on which he subsequently goes on to ground his concept of free will and moral responsibility in Chapters Seven and Eight of *Mind, Brain, and Free Will* (2013). In this Appendix I want to suggest briefly the consequences of this approach for his interpretation of free will.

Problems with Causation.

Swinburne believes that standard theories of (physical) causation do not easily fit with a his two-substance ontology. Standard theories rest on 'event-driven' interpretations of causality rather than 'substance-drive' ones. He rejects Hume's account of causation, 'regularity theory', because, although 'coherent', '...it does not fit many paradigm examples of our ordinary talk about "laws" and "causes"' (Swinburne, 2013, p.127). He suggests that regularity theory does not fully capture what might be meant when we say, 'Newton's laws are a "law of nature" or the position of Saturn caused the recession of Jupiter from the Sun' (Swinburne, 2013, p.127). His key point is that regularity theory lacks the explanatory power of two other theories he goes on to compare: relations-between-universals (RBU) and substance-powers-liabilities (SPL).

According to RBU theory, laws of nature express physical necessitation or probability relations between universal properties. For example, '... all photons have a rectilinear velocity of 299,729 km/sec in empty space'. This is a relation of natural necessity that connects the universals 'photon' and 'having a rectilinear velocity of 299,729 km/sec in empty space'. This means that when a

photon is instantiated (i.e. a photon exists) it 'brings with it' by natural necessity the universal property 'having a rectilinear velocity of 299,729 km/sec in empty space' (Swinburne, 2013, p.129).

Alternatively:

> On the SPL account the fundamental laws are regularities; but not regularities of succession of events, but regularities in the causal powers and liabilities of substances – powers to naturally necessitate (or to make naturally probable), that is, to cause effects; liabilities of natural necessity to exercise those powers under certain circumstances or under all circumstances. (Swinburne, 2013, p. 129)

If we translate the photon example above into the language of SPL then: 'all photons have a velocity of 299,729 km/sec in empty space' is the regularity that each photon has the same power of every other photon *to cause itself* [italics mine] to be 299,792 km distant after *n* sec, and the *liability inevitably* [italics mine] to exercise this power when in empty space (Swinburne, 2013, p.129). This account fits with Swinburne's two-substance ontology.

Swinburne goes on to make a distinction between two types of causation: 'non-intentional causation' and 'intentional causation' (agent causation). These are two species of the same genus because in both cases causation is rooted in the powers and liabilities of substances. Intentional 'events' are not really events at all. For intentional causation, the person whose intention it is, is a substance not an event: 'A person having an intention (acting) is simply the person exercising causal power'. And a person, Swinburne goes on to argue, (on his general theory of substance dualism) is a pure mental substance who exercises causal powers. Substances with specific capacities (powers) describe the world as it is – rather than (abstract) laws of nature – the RBU approach. Non-intentional and intentional causation is thereby rendered under a 'simpler' account. This apparently explains why so many effects can be explained by a combination of intentional and non-intentional causes, for example, a (thrown) rock crashing through a window.

Swinburne's Preferred Theory.

Swinburne sets out to show why the SPL account should be preferred. The superiority of SPL (i.e. it is more probably true) can be established by certain epistemic principles (Swinburne, 2013, pp.40-66). When comparing scientific theories, we ought to prefer the one which is the simplest and/or makes the evidence more probable.

> The simplicity of a scientific theory is a matter of it postulating few kinds of entities (i.e. substances), few properties, few kinds of properties, properties closely related to observation, few laws relating substances to each other and mathematically simple relations in the laws. (Swinburne, 2013, p.59)

He claims that we ought to prefer the SPL account on the following grounds:

> R1. It is readily understandable: the SPL account may be rendered in terms of readily accessible concepts i.e. powers and liabilities. The RBU theory is in this sense more complex in that it requires the postulation of 'strange things' such as the relations of natural necessity between universals.
>
> R2. It is parsimonious: since powers and liabilities are properties of substances we do not need an extra category such as laws of nature – beyond substances, properties, and times to give a complete account of the world.
>
> R3. It is not event driven: in the SPL account 'causing influence' is a basic category not analysable in terms of any other category. This provides a more economical theory contrary to Regularity and RBU theories because it is not events causing other events but substances causing events.

Laws of nature are contingent generalisations about the powers of substances and their liabilities to exercise those powers in different, varying circumstances. (Swinburne, 2013, p.131). The principal advantage of the SPL theory is that it creates the space for the metaphysical possibility of a type of substance (a mental substance) with intentions and capable of intentional causation or agent

causation (See Swinburne, 2013, pp.132-133). This supports interactionism as characteristic of mind/body relationships. In summary:

> P1. There are three possible, candidate accounts of causation: a theory of regularity (Hume), the RBU theory and the SPL theory.
>
> P2. Only the RBU and SPL theories constitute plausible accounts of what is involved in our talk about laws of nature in an explanatory sense.
>
> P3. It is possible to judge the explanatory theory which is most likely to be true according to criteria of relative simplicity or complexity.
>
> P4. The SPL account meets the criterion of simplicity in the sense that it can be expounded through readily accessible concepts and in terms of our actual experiences of what it is to cause something to happen i.e. events do not cause other events, but substances do.
>
> C. It is metaphysically possible for there to be substances which have the power to cause things to happen but no deterministic liability to do so.

This denies the determinism which seems inherent to the RBU theory i.e. the idea of necessitating and universal laws. SPL provides a foundation for intentional causation which is also compatible with non-intentional causation. Swinburne's next move is to 'flesh' this out to show that humans are the type of substances of the kind indicated by the SPL account: substances which have possibilities *to intend* to cause effects (actions). Swinburne gives the following argument that humans are capable of intentional causation and thereby the possibility of exercising free will:

> P1. The principle of credulity states that we should believe that things are as they seem to be in the absence of any (compelling?) counter-evidence i.e. things are generally as we believe they are.
>
> P2. When we try to do an action, we believe we can exercise a causal influence, maybe with the co-operation of other causes, and bring about an intended effect.
>
> P3. There cannot be any counter-evidence to falsify the claim that normally intentions cause, with the mediation of brain events, effects

> both within and outside our bodies. This constitutes the performance of actions.
>
> C. 'Trying' to do something, performing actions, just is (normally) exercising causal power. (Swinburne, 2013, pp.134-135)

The first thing to note is P3 is deeply ambiguous. If the claim here is to the effect that it is not logically possible for there to be counter-evidence, then this ceases to be a contingent proposition. It can then only be made true by definition and thus makes no direct empirical claim upon the world. If it is meant to be a claim that there is little or no practical possibility of counter-evidence, then the claim must be a weak claim. The less that is claimed the less there is at risk. The greater the empirical content of a proposition then the wider is the front disclosed to cavil, contradiction and falsification. Either interpretation must weaken the argument (Collingridge, 1987, p.49).

Secondly, intentional 'events' are not really events at all (on the SPL theory) and there is something in this. The person whose intention it is, is a substance not an event. 'A person having an intention (acting) is 'simply' the person exercising causal power' (Swinburne 2013, p.2). It is from our own experiences of our selves causing effects that our concept of causation arises. This is fine if what we mean are persons in our normal, everyday understanding of persons as flesh and blood creatures acting in the world. But the water is surely muddied by Swinburne's distinction between persons (mental substances) and persons (humans). He writes:

> I shall understand by 'persons' (in what I believe is the normal sense of the word, but nothing turns on that [*sic*]) substances with a capacity (at least after normal growth) for beliefs and actions of the degree of sophistication typical of those present-day earth-inhabitants called 'humans'. I shall understand by 'humans' any persons who have or have had at some time a kind of body and ancestry similar to those persons whom we call 'humans' today. (Swinburne, 2013, pp.141-142)

The problem here is that 'person' in this incorporeal sense is not a person in any every day, recognisable sense as Swinburne seems to think. It is not an expression for something readily and certainly identifiable or a known sort of thing (like 'humans'). Person in Swinburne's highly restricted sense is a purely negative term which seems impossible to supply appropriate means of

identification. Persons may persist minus a body. While it makes sense, in using Swinburne's terminology, to say that humans exercise causal powers, it isn't clear what it might mean for (bodiless) persons, in his sense, to exercise such powers. How would we identity, and then re-identify, a bodiless person acting at one time and then another? (See Flew, 1966, p.32)

The meaning of 'intentional' and 'non-intentional causation'.

As we have seen Swinburne draws a distinction between intentional and non-intentional causation. The former is characterised as 'agent causation'. But non-intentional and intentional causation seemed to be conceived according to a causal paradigm. I suggest that this is not only barking up the wrong tree but operating altogether in the wrong wood. It is a category mistake. As Ryle (1976) suggests this is to conflate our descriptions and explanations of human actions with those of what we might more accurately call 'natural', rather than physical, causation. One of Swinburne's weaknesses is his all-encompassing use of the word 'physical'. We must, I think, speak of not only physical causation (with its hinterland suggestive of physical process and mechanism) but also a range of causations appropriate to the many species of scientific explanation: chemical, biological, genetic, and so on each with their own special characteristics.

On Swinburne's account we experience ourselves as mental substances as deciding to cause things to happen. This is 'the normal view' supported by the credulity principle. We believe ourselves to be deciding and implementing intentions (not necessarily successfully) as if intentions were some 'inner movement' or impulsion which causes brain cells to fire leading to physical events bodily movements or physical events in the world. But where do these 'impulsions' come from? Intentions are the source of a causal chain leading from mental events to physical events. Swinburne acknowledges, of course, that the reverse is also true, physical events may lead to brain events which impact our minds (mental substances). This is why Swinburne can suggest, fallaciously, that I am 'caused' to have certain beliefs. But this is surely false. I may have the beliefs that I do not for reasons-as-causes but because I have reasons-as-grounds for believing them or rejecting them.

Of course, it must be the case that if we are talking about physical, bodily movements then causative mechanisms, such as muscle contractions, nerve

impulses and so forth, are involved. But it is less clear that it makes sense to talk about intentions or motives on the physical paradigm. The matter becomes clearer if we think about intentions as the reasons for doing things as Elizabeth Anscombe (2000) suggests. She tackles the problem by defining intentional actions as 'those to which a certain sense of the question "Why?" is given application' (Anscombe 2000, p.8). The 'certain sense' is the sense in which, if there is a positive answer, it is the giving of a reason for acting. 'Why did you go to the gym?' 'I went because I wanted to lose weight'. The question 'Why?' is 'refused application', according to Anscombe, if the answer is along the lines of 'I did not know I was doing that.' There are many descriptions which may be applied to the same action some of which the agent is aware of and some not (Anscombe, 2000, p.11).

In answering the question 'Why' in a certain sense we give our reasons for acting that is explaining our motives. Giving a reason in the sense of providing a motive is not the same as giving a reason in the sense of giving a causal explanation:

> Motives may explain actions to us; but that is not to say they 'determine', in the sense of causing actions. We do say: 'His love of truth caused him to…' and similar things, and no doubt such expressions help us to think that a motive must be what produces or brings about a choice. But this means rather 'He did this in that he loved the truth'; it interprets his action. (Anscombe, 2000, p.19)

Intentions and motives are not mental causes in the sense that Swinburne would have us believe. It is important to distinguish reasons as motives and reasons as causes. A mental cause is what is described when we answer the question 'What produced this action or thought, or feeling on your part: what did you see hear or feel, or what ideas or images cropped up in your mind which led to it?' (Anscombe, 2000, pp.17-18).

It may be natural to pass from the notion of motive as suggestive of something that *moves* which is then interpreted as what *causes* a person's actions. This is surely what we find in Swinburne in that what causes our actions is a mental event created by a mental substance that brings about brain events which bring about actions. But as Anscombe (2000, p.18) suggests how this is supposed to happen is obscure. Is it a kind of 'pushing' in some other (incorporeal) medium

or by some other means? There are altogether other grounds, or ways of interpreting, how it is that persons have intentions and make choices without Swinburne's notion of some ghostly 'movings' of an incorporeal mental substance.

Finally, I suggest that Swinburne's theory of causation leads him to define 'free will' in a curious and erroneous way. An agent has free will '…insofar as the agent acts intentionally without their intentions being fully determined by prior causes' (Swinburne, 2013, p.202). He claims this is a 'normal understanding' of free will recognised by non-philosophers. But is this a normal 'understanding'? We are more likely, in everyday discourse, to talk about being free to choose this or that rather than 'causation' and 'free will'. It seems to me, although I cannot develop the point here, that this is far too restrictive a way of looking at the matter in seeming to suggest that the explanations of intentions and actions can be sought either on a physical model of causation or by some mysterious 'pushing' in some other (incorporeal) medium. In seeking explanations for actions we talk about reasons-as-causes; reason-as-motives; and reasons-as-grounds'. Swinburne simply fails to acknowledge these distinctions and their relationship to appropriate explanations.

Glossary

Terms as defined by Richard Swinburne in *Mind, Brain, and Free Will* (2013). Appropriate page references appear in brackets.

Terms	**Definition/Description**
Animalist View of Identity	'...the criteria for the existence and continuing of human beings [are] the same as those for the existing and continuing of human bodies.' (p.42)
A Simple	'...is something which has essentially only one indivisible part; then the existence of that part is necessary and sufficient for the existence at any time of that substance.' (p.30)
Basic Beliefs	'...propositions which seem to be true to us but not on the basis of other propositions (always allowing that we may come to believe on the basis of other evidence that they are not true after all) ...They will include beliefs about our mental life (I have a toothache).' (p.41)
Bundle Theory of Identity	'...substances are merely bundles of co-instantiated properties...the identity of a substance S_1 at one time t_1 with some substance S_2 at another time t_2 depends (as well as on both substances having the essential

	properties of the same kind) on the extent to which the properties of each bundle are similar, and whether the bundles are linked over the intervening time by a (spatio-temporal) chain of similar bundles, each of which causes the existence of the next member of the chain.' (p.149)
Causation	Under the RBU account of laws of nature: '...causation is regarded as the relation between two events'. (pp.7-8)
	Under the SPL account of laws of nature: '...it is a substance or several substances together which cause effects; they cause effects normally in virtue of their liabilities to do so under certain conditions.' (p.8)
Coherent Sense	'...a sense such that some existential sentence containing that word used in that sense does not entail a contradiction. (By an 'existential sentence' I mean one that affirms the existence of something) ...If the sense of a word is coherent, then it designates a coherent concept.' (p.50)
Complex View of Identity	'...personal identity is analysable in terms of degrees of continuity ... of other things, that is, of properties and /or matter.' (p.150)
Conceivability	'If we understand by a sentence being "conceivable" that it does make sense to suppose it to be true (i.e. that it does not entail a contradiction), then

	“conceivable” means the same as “logically possible”.’ (p.47)
Credulity Principle	‘...that any basic belief (that is the content of that belief, the propositions believed) is probably true (that is, more probable than not that the belief is true) on the believer’s evidence that he believes it – in the absence of evidence in the form of other basic beliefs of that believer that makes it probable that he is mistaken.’ (p.42) Or...’what seems to us to be so probably is so, that our apparent experiences are probably real experiences.’ (p.42)
Defeaters	‘...beliefs acquired by apparent experience, memory, and testimony are...open to counter-evidence. Such counter-evidence is often called a ‘defeater’. There are two kinds – undermining and overriding. If we have inferred (whether consciously or without thinking about it) the occurrence of some event ε from evidence γ, then an undermining defeater is evidence (making it probable) that γ did not occur or is not good evidence for ε whereas an overriding defeater is new evidence that ε did not occur.’ (p.57)
Epistemic Assumption	‘…A justified belief in a scientific theory (which is not itself a consequence of any higher-level theory in which the believer has a justified belief) requires a justified belief that the theory makes true predictions.’ ‘…A justified belief that a theory makes true predictions is (unless this is a consequence of some

	higher-level theory in which the believer has a justified belief) provided by and only by the evidence of apparent experiences, memory and testimony that the theory predicts certain events and those events occurred.' (p.64) '…Such justification is undermined by any evidence that apparent experience was not caused by the event apparently experienced, any apparent memory was not caused by an apparent experience of the event apparently remembered, or any apparent testimony was not caused by the testifier's intention to report his or her apparent experience or memory.' (p.64)
Event	Formally '...either the instantiation of a property in some substance or substances, or event or events at a certain time, or the coming into existence or the creating to exist of some substance at some time.' (p.6) In contrast to ordinary usage: '...only those instantiations of properties which involve change are called events.' (p.6)
Experiences	'...(apparent) perceptions (by means of the senses of physical events, (apparent) 'experiences' (in a narrower sense) of our thoughts, feelings, sensations etc., and also (apparent) awareness of truths of reason (awareness of 'seeing' some sentence to be logically possible or whatever).' (p.55)

Externalists	'...hold that the justification of a belief depends on the kind of process which produced it: e.g. "reliabilism" which holds that a belief is justified iff it is produced by a reliable process, a process which normally produces true beliefs'. (p.50)
Fundamental Criterion	'...justified belief that some event occurred requires the assumption that that event is accessible (in a privileged way) to the believer or causes an event thus accessible...' (p.65).
Humans	'I shall understand by 'humans' any persons who have or have had at some time a kind of body and an ancestry similar to those persons whom we call 'humans' today.' (p.142)
Identity: Events	'Events which involve the instantiation of the same properties in the same substances at the same time will be identical...We should count any two canonical descriptions as picking out the same event if and only if any possible world which contained an event picked out by one would contain an event picked out by the other one, and conversely.' (p.27)
Identity: Property	'...It follows from properties being individuated by the informative designators which pick them out...it is a purely a priori matter whether or not property a is identical with property b (depending on whether having the one entails having the other, and conversely).' (p.24)

Identity: Substances	‘For a substance S_2 at time t_2 to be the same substance as S_1 at an earlier time, t_1, two kinds of criteria have to be satisfied. First the two substances have to belong to the same minimum kind (defined by essential properties). The second requirement ...is that the two substances should consist of largely the same parts – or parts obtained by gradual replacement for the former substance.’ (p.29)
Informative [rigid] Designators	‘...words which always pick out the same property and so on and tell us fully which property that is.’ (p.10) ‘They will pick out the same property or whatever iff that property or whatever can be designated by the same informative designator; and so two properties or whatever will be the same iff their informative designators are logically equivalent (that is each entails the other).’ (p.10)
Impure Mental Property	‘...derives its mental nature from there being an aspect or element of it which consists of a pure mental property; and the subject’s privileged access to the instantiation of a mental property then consists in her privileged access to the instantiation of the pure mental property.’ (p.68)
Internalists	‘...hold that a belief is justified iff it is made probable by evidence accessible to the believer.’ (p.50)
Justified Belief	‘A belief is justified iff the belief is rendered

	(epistemically) probable by the evidence available to the believer and the believer believing it because of the evidence.' (p.41)
Laws of Nature	Interpreted as relations-between-universals (RBU): 'Laws are (logically contingent) relations of natural necessitation (or probabilification) between universals, that is between properties in my [Swinburne's] sense.' (p.6) Or interpreted as: substance-powers-liabilities (SPL): '... regularities, not of actual (past, present, and future) successions of events, but regularities in the causal powers and liabilities, which are properties, of actual substances – powers naturally to necessitate (or make naturally probable), that is, to cause effects; and liabilities of natural necessity to exercise those powers under certain circumstances or under all circumstances.' (p.7)
Logically Contingent Sentence	'...is a logically possible sentence which is not necessary (logically).' (p.19)
Logically Impossible Sentence	'...is one which entails a contradiction'. (p.19)
Logically Possible Sentence	'...is one which does not entail a contradiction.' (p.19)

Logically Necessary Sentence	'...is one iff its negation entails a contradiction.' (p.45)
Mental Event	'It is now natural to define a mental event along the same lines as one to which the substance involved has privileged access.' (p.70)
Mental Property	'...as one to whose instantiation in it a substance necessarily has privileged access on all occasions to its instantiation.' (p.67)
Metacriterion	'The requirement that our vocabulary should enable us to tell the whole history of the world... determining when we should deem that two referring expressions pick out the same property, substance or whatever.' (p.10)
Metaphysical Contingency	'Some event is ...metaphysically contingent iff it is metaphysically possible but not metaphysically necessary.' (p.15) 'A metaphysically contingent sentence depends not merely on what the sentence claims but how the world is, independently of how we describe it,' (p.20)
Metaphysical Impossibility	'An even is metaphysically impossible iff it did not occur and would not have occurred even under any ... circumstances.' (p.15)
Metaphysical	'An event is metaphysically necessary iff, whatever

Necessity	else is the case, it must happen.' (p.15)
	'When we know fully what we are talking about (e.g. that in talking about 'water', we are talking about H_2O, mere a priori considerations can show whether some sentence is metaphysically necessary or impossible for this kind of reason. Hence there is available a definition of a sentence as metaphysically necessary (impossible or possible) iff it is logically necessary (impossible or possible) when we substitute co-referring informative or uninformative designators, that is designators which pick out the same substance or property.' (pp.19-20)
	'Truth or falsity of a metaphysically necessary or impossible sentence depends solely on the conventions of language.' (p.20)
Metaphysical Possibility	'An event is metaphysically possible iff it could happen under some circumstances.' (p.15)
Non-rigid Designator	'...is a word which applies to something only as long as it has some non-essential properties. The Prime Minister, as normally used, is a non-rigid designator; it refers to whoever is Prime Minister at the time in question'. (pp.10-11)
Normal View	'...it is very probable indeed that there is a two-way interaction between conscious events and physical events.' (p.101)

	‘We just do, as a matter of fact, cause physical events to occur by forming and carrying out intentions (i.e. pure mental events). We normally believe that our intentions guide our bodily movements, and this is just how it seems to us. The scientific consensus is that bodily movements are caused by brain events which are preceded by the formation of intentions. If intentions cause movements, they do so by causing brain events’. (p.103)
Objective History of the World (or a segment thereof).	‘...all the events which have occurred, are occurring, or will occur.’ (p.8) ‘...everything that has happened, is now happening and will happen ever or anywhere (or in that segment) …this history can be told with the familiar categories of substances, properties and times.’ (p.4) ‘A full history of the world will have to include (or entail the occurrence of) pure mental events as well as physical events; sensations, thought, intentions, beliefs, and desires as well as brain events and behaviour.’ (p.87)
Persons	‘I shall understand by ”persons” (in what I believe is the normal sense of the word, but nothing turns on that) substances with a capacity (at least after normal growth) for beliefs and actions of the degree of sophistication typical of those present-day earth-in habitants called “humans”.’ (p.140)
Physical Event	‘...as one to which the substance involved does not have privileged access.’ (p.70)

Physical Property	'...as one to whose instantiation in it a substance necessarily has no privileged access on any occasion of its instantiation.' (p.67-68)
Place	'...talk about place is reducible to talk about substances and their relations to other substances. A place is the place it is in virtue of the spatial relations (of distance and direction) to particular physical objects (substances).' (p.8)
Properties	'A property may be a monadic property of one substance (possessed by that substance by itself quite independently of its relationship to other substances), or a relation between two or more substances...properties (monadic and relational) are universals, that is, they could be possessed by different substances from the ones they are possessed...' (p.5)
Properties: Essential	'...if a substance did not have these it could not exist'...occupying space is an essential property of my desk.' (p.5)
Properties: Contingent	…some properties of a substance are contingent (i.e. non-essential) properties of that substance. Being brown is a contingent property of my desk; if it were painted red instead, the desk would still exist. (p.5)

Principle of the Identity of Composites	'...there cannot (logically) be two things which have all the same parts having all the same properties, arranged in the same way'. (p.35)
Principle of the Identity of Indiscernables	'...maintains that there cannot be two things (substances or any other things) which have all the same properties, monadic and relational.' (p.33)
Pure Mental Event	'...a mental event which does not entail a metaphysically contingent physical event as that substance.' (p.70)
Pure Mental Property	'...as one whose instantiation in a substance does not entail the instantiation of any metaphysically contingent physical property in that substance.' (p.68)
Rigid Designator	'...is a word which in every possible world, designates the same object.' (p.10)
	'For a rigid designator to be an informative designator it must be the case that anyone who knows what that word means (that is has the linguistic knowledge of how to use it) knows a certain set of conditions necessary and sufficient (in any possible world) for a thing to be that thing (whether or not he can state these conditions in words).' (p.12)

Same Year	'…that there are two possible senses of 'same year'; in one sense an event happens in the same year as another event iff they occur at the same temporal interval after the original event, and in another sense an event happens in the same year as another event iff they both occur at the present moment or at the same temporal interval before or after the present moment...So it is intelligible to suppose – that is, it is logically possible – that every type event might have occurred a "year" earlier than it did, in the sense of "year" in which the year is the year it is in virtue of its temporal distance from the present year.' (p.53)
Simple View of Identity	'...is the view that each person has a "thisness" which makes him or her that person, a "thisness" other than any thisness possessed by the matter of their brains; and that being that person is compatible with having any particular mental properties or any physical properties (and so body) at all.' (p.151)
Simplicity Criterion	'The simplest explanation [of human behaviour] ...is...probably the correct one…That simplicity is evidence of truth is the inductive criterion shared by all humans which give access to 'other minds' as well as all our knowledge of events beyond our immediate experience.' (p.89)
Soul	'...is non-physical and indivisible and possesses essentially only pure mental properties.' (p.173) ',,,a simple…' (p.173)

Substance	'...physical substances are what they are in virtue of their properties, including the causal and spatio-temporal relations of their parts to those earlier substances.' (p.38)
	'An informative designator of a substance will then pick it out in virtue of its having the properties which make it a substance of a certain kind, and the spatio-temporal or other relations of its parts to those earlier substances.' (p.38)
	'...a particular concrete object: my desk, that person, the photon, (particle of light) emitted from this light source which landed on this screen.' (p.4)
	'Substances may have other substances as parts and... can exist (it is logically possible) independently of all other things of its kind (i.e. all other substances, apart from its parts) ...Substances exist all at once'. (p.4)
Supervenience of Properties	'Loosely, properties of kind A supervene on properties of kind B, iff which substances have which (if any) A – properties is entailed by which (if any) substances have B-properties. More precisely, a kind of property A supervenes on a kind of property B iff 'Necessarily' for any property F of kind A if any substance x has F, there exists a property G of kind B such that x has G, and necessarily any substance γ having G has F.' (p.22)

Supervenience of Events	'…events consisting in a substance x having a property F of a kind A, supervene on events consisting in x having property B of kind G, iff necessarily for any x which has property F of kind A, x has some property G of kind B, such that necessarily if x has G, x has F.' (p.22)
Testimony Principle	'...a second a priori epistemic principle [the first being the Credulity Principle – DSH] that what people seem to be (i.e. apparently are) telling us that they are experiencing or remember having experienced or remember as facts, they (epistemically) probably do or did experience or probably are facts, again barring counter-evidence.' (p.56)
Thisness	'A substance has thisness iff there could exist instead of it (or in addition to it) a different substance which has all the same properties as it, monadic and relational, including its spatio-temporal relations to earlier and later substances having such-and-such monadic properties and relations.' (p.33) 'Identity for non-physical substances which have thisness would be a 'brute' identity, not further definable.' (p.34)
True Sentence	'...entails no contradiction, and if it is obvious that some sentence (e.g. 'my desk is brown') is true, then it is obvious that it is logically possible.' (p.45)

Bibliography

ANON.,1975. The Mandela particle is threatened. *New Scientist*, 7 August, p.310.

ANSCOMBE, G.E.M.,1975. The First Person. *In*: S. GUTTENPLAN, ed. *Mind and Language*. Oxford: Oxford University Press, 1975, pp.45-65.

ANSCOMBE, G.E.M., 2000. *Intention.* 2nd ed. Cambridge, Mass.: Harvard University Press.

ARISTOTLE, 1928. *The Works of Aristotle. Volume I.* W.D. ROSS, ed. London: Oxford University Press.

ARMSTRONG, D.M., 1968. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.

AUSTIN, J.L., 1970. Are there a priori concepts? *In*: J.O. URMSON and G.J. WARNOCK, eds*. J.L. Austin, Philosophical Papers*. 2nd ed. Oxford: Oxford University Press, 1970, pp.33-54.

AUSTIN, J.L., 1970. Other Minds. *In:* J.O. URMSON and G.J. WARNOCK, eds*. J.L. Austin, Philosophical Papers*. 2nd ed. Oxford: Oxford University Press, 1970, pp.76-116.

AUSTIN, J.L., 1970. A plea for excuses. *In:* J.O. URMSON and G.J. WARNOCK, eds*. J.L. Austin, Philosophical Papers*. 2nd ed. Oxford: Oxford University Press, 1970, pp.175-204.

BAYNE, Tim, 2010. *The Unity of Consciousness*. New York: Oxford University Pres Inc.

BAKER, Rudder, Lynne, 2014. Swinburne on substance dualism. *European Journal for Philosophy of Religion*, 6 (2), Summer, pp.5-15.

BALDWIN, Thomas, 2001. *Contemporary Philosophy: Philosophy in English since 1945.* Oxford: Oxford University Press.

BERLIN, Isaiah,1979. Introduction. In: Isaiah BERLIN, ed. *The Age of Enlightenment: the 18th Century Philosophers.* Oxford: Oxford University Press, 1979, pp.12-29.

BLACKBURN, Simon, 1999. *Think: A compelling introduction to philosophy.* Oxford: Oxford University Press.

BLACKBURN, Simon, 2005. *Dictionary of Philosophy*. Oxford: Oxford University Press.

BLACKFORD, Russell, 2014. Richard Swinburne *Mind, Brain, and Free Will*. Oxford University Press, 2013. 237 pages. $34.95 (Paperback ISBN 9780199662579). *Philosophy in Review*, XXXIV (3-4), pp.110-112.

BOYD, Graham, W., 2016. Not Us: The Brain's Risk/Reward System Makes Our Choices. *Philosophy Now*, (112), February/March, pp.10-12.

BOYNE, C., 2018. Personal Communication. 6th April.

BROAD, C.D., 2000. *The Mind and Its Place in Nature*. 1st ed 1925. London: Routledge and Kegan Paul.

BROWN, Christopher, M., 2005. *Aquinas on the Ship of Theseus: Solving puzzles about material objects.* London: Continuum.

BUTLER, Joseph, 1855. *The Analogy of Religion to the Constitution and Course of Nature: also Fifteen Sermons*. 4th ed. Joseph Angus ed. London: Religious Tract Society.

CHALMERS, David, J., 1996. *The Conscious Mind.* Oxford: Oxford University Press.

CHALMERS, David, J., 2002. *Philosophy of Mind: Classical and Contemporary Readings.* Oxford: Oxford University Press.

CHURCHLAND, Paul, M., 1995.*The Engine of Reason, the Seat of the Soul: A philosophical journey into the mind.* Cambridge: Massachusetts, The MIT Press.

CAMPBELL, Keith, 1971. *Body and Mind.* London: The Macmillan Press.

CLAIR, Kassia, St., 2016. *The Secret Lives of Colour.* London: John Murray.

CLARKE, Desmond, M., 2005. *Descartes' Theory of Mind*. Oxford: Oxford University Press.

COLE, Peter, 1999. *Philosophy of Religion.* London: Hodder and Stoughton.

COLLINGRIDGE, David, 1987. *Criticism: its philosophical structure.* Lanham, MD: University Press of America.

CONEE, Earl and SIDER, Theodore, 2005. *Riddles of Existence: A Guided Tour of Metaphysics.* Oxford: Clarendon.

CORNFORTH, Maurice, 1971. *Marxism and the Linguistic Philosophy.* London: Lawrence and Wishart.

COX, Gary, 2012. *The God Confusion: why nobody knows the answer to the ultimate question.* London: Bloomsbury.

CRAWFORD, Sean, 2005. *Aspects of Mind.* Milton Keynes: The Open University.

DAINTON, Barry, 2001. *Time and Space.* Chesham: Acumen Publishing.

DAINTON, Barry, 2008. *The Phenomenal Self*. Oxford: Oxford University Press.

DAINTON, Barry, 2014. *Self: Philosophy in Transit.* London: Penguin.

DANIEL, David, Mills, 2006. *Briefly: Anselm's Proslogion.* London: SCM Press.

DAVIDSON, Donald, 1980. The Individuation of Events. *In:* D. DAVIDSON, *Actions and Events*. Oxford: Clarendon Press,1980, p.179.

DAVIDSON, Donald,1979. *Mental Events.* In: Ted HONDERICH and Mysles BURNYEAT, eds. *Philosophy as It Is.* London: Allen Lane,1979, pp.213—236.

DENNETT, Daniel, C. 2002. True Believers: The Intentional Strategy and Why It Works. *In:* David J. CHALMERS, ed. *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press, 2002, pp.556-567.

DEW, Jnr., James, K., 2014. Swinburne's new soul: A response to Mind, Brain, and Free Will. *European Journal for Philosophy of Religion*, 6 (2), Summer, pp.29-39.

DILMAN, Ilham, 2011. *Philosophy as Criticism: Essays on Dennett, Searle, Foot, Davidson, Nozick.* London: Continuum.

FARRIS, Joshua, 2015. What's So Simple About Personal Identity. *Philosophy Now,* (107), April/May, pp.30-33.

FINCH, H.L., 1995. *Wittgenstein.* Shaftesbury, Dorset: Element Books.

FLEW, Antony, 1953. *A New Approach to Psychical Research*. London: Watts.

FLEW, A. and MACINTYRE, A.C., eds., 1956. *New Essays in Philosophical Theology*. London: SCM Press.

FLEW, Antony, 1961. *Hume's Philosophy of Belief: a study of his first Inquiry*. London: Routledge & Kegan Paul.

FLEW, Antony, 1964. Introduction. In: Antony FLEW, ed. *Body, Mind, and Death*. London: Collier-Macmillan, 1964, pp.1-28.

FLEW, Antony, 1966. *God and Philosophy.* London: Hutchinson.

FLEW, Antony, 1968, Locke and the Problem of Personal Identity. *In:* C.B. MARTIN and D.M. ARMSTRONG, eds. *Locke and Berkeley*: *a collection of critical essays.* London: Macmillan, 1968, pp.157-178.

FLEW, Antony, 1975. *Thinking about Thinking.* Glasgow: Fontana/Collins.

FLEW, Antony, 1978. *A Rational Animal*. Oxford: Clarendon.

FLEW, Antony, 1979. *A Dictionary of Philosophy.* London: Pan.

FLEW, Antony, 1986. *David Hume: Philosopher of Moral Science.* Oxford: Basil Blackwell.

FLEW, Antony, 1987. *The Logic of Mortality*. Oxford: Basil Blackwell.

FLEW, Antony, 1992. The Cartesian Assumption. *In*: Paul EDWARDS ed. *Immortality*. New York: Macmillan, 1992, pp.220-224.

FLEW, Antony, 1994. *An Introduction to Western Philosophy*. Rev. ed. London: Thames Hudson.

FRANKLIN, Christopher, Evans, 2015. Mind, Brain, and Free Will. *Philosophical Review,*124 (2), April, pp.255-258.

FREGE, 1988. *Translations from the Philosophical Writings of Gottlob Frege*. 3rd. ed. Peter GEACH and Max BLACK, eds. Oxford: Oxford University Press.

GEACH, Peter, 1992. Immortality. *In*: Paul EDWARDS ed. *Immortality*. New York: Macmillan, 1992, pp.225-234.

GLOVER, Jonathan, 1988*. I: the philosophy and psychology of personal identity.* London: Allen Lane.

GOETZ, S. and TALIAFERRO, C., 2011. *A Brief History of the Soul.* Chichester. West Sussex: Wiley-Blackwell.

GRAYLING, A.C., 1990. *An Introduction to Philosophical Logic*. Oxford: Blackwell.

HEIL, John, 2013. *Philosophy of Mind*. 3rd ed. New York: Routledge.

HOLDSWORTH, Nick, 2014. Reviews *Mind, Brain, and Free Will* by Richard Swinburne Oxford University Press, 2013, pp.242 ISBN: 978-0-19-966245-2 (Hbk.) £50 ISBN: 978-0-19-966245-9. *Philosophy*, 89 (4), October, pp. 635-642. [Online publication 24th Jan.,2014] Accessed 11/10/15.

HUNT, David P., 2014. Swinburne on conditions for free will and responsibility. *European Journal for Philosophy of Religion*, 6 (2) Summer, pp.39-49.

JARWORSKI, William, 2014. Swinburne on Substances, Properties and Structures. *European Journal for Philosophy of Religion*, 6 (2), Summer, pp.17-28.

KALDAS, Samuel, 2015. Descartes versus Cudworth: On the moral worth of animals. *Philosophy Now*, (108), June/July, pp.28-31.

KENNY, Anthony, 1975. *Wittgenstein.* Harmondsworth, Middlesex: Penguin.

KENNY, Anthony, 1990. *The Metaphysics of Mind*. Oxford: Clarendon Press.

KENNY, Anthony, 1993. *Descartes: a study of his philosophy*. Bristol: Thoemmes Press.

KRIPKE, Saul, A. 1980. *Naming and Necessity.* Cambridge, Massachusetts: Harvard University Press.

KRUEGER, Joel, 2012. Seeing Mind in Action*. Phenomenology and Cognitive Science*, 11, pp.149-173.

LATHAM, Sally, 2014. Swinburne's Separations. *Philosophy Now*, (102), May – June, pp.35-38.

LEVIN, Michael, E, 1979. *Metaphysics and the Mind-Body Problem.* Oxford: Clarendon Press.

LEVY, Joel, 2016. *The Infinite Tortoise: The Curious Thought Experiments of History's Great Thinkers.* London: Michael O'Mara Books.

LIBET, Benjamin, 2004. *Mind Time – The Temporal Factor in Consciousness*. Harvard: Harvard University Press.

LITTLEJOHN, Clayton, 2013. Review: *Mind, Brain, and Free Will* by Richard Swinburne, Oxford, Oxford University Press. *Metapsychology* [online review], 17 (35). Accessed 10/10/15.

McGUIRE, Mara and HAUSE, Jeffrey, 2014. *Mind, Brain, and Free Will*. *The Review of Metaphysics*, 67 (3), p.670. [Arts and Humanities Full Text]

NAGEL, Thomas,1986. *The View from Nowhere.* Oxford: Oxford University Press.

NAGEL, Thomas, 2012. *Mind and Cosmos: Why the Materialist, Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford: Oxford University Press.

NOWELL-SMITH, P.H., 1965. *Ethics.* Harmondsworth, Middlesex: Penguin.

NOZICK, Robert, 1981. *Philosophical Explanations*. Oxford: Oxford Univesrity Press.

O'HEAR, Anthony, 1991. *What Philosophy Is: An Introduction to Contemporary Philosophy.* London: Penguin Books.

PALMER, David, 2013. Richard Swinburne *Mind, Brain, and Free Will*, Oxford University Press, 2013, 242PP. $35.00 (pbk.), ISBN 9780199662579. *Notre Dame Philosophical Review*, 9 (13) [e-journal]. Available at: Ndpr.nd.edu/news/42687-mind-brain-and-free-will/ [Accessed 10/10/15].

PAPINEAU, David, 2002. *Thinking about Consciousness.* Oxford: Clarendon Press.

PARFIT, Derek, 1986. *Reasons and Persons*. Oxford: Oxford University Press.

PINDER, Mark, 2017. Experimental Philosophy versus Natural Kind Essentialism. *Philosophy Now,* (120), June/July, pp.30-32.

POPPER, Karl, R, 1980. *The Logic of Scientific Discovery.* London: Hutchinson. [1st. ed, Vienna, 1935].

PUTNAM, Hilary, 1975. The Meaning of 'Meaning'. In: Hilary PUTNAM, ed. *Mind, Language and Reality: Philosophical Papers, Vol, 1*. Cambridge: Cambridge University Press, 1975.

PUTNAM, Hilary, 1995. *Reason, Truth and History.* Cambridge: Cambridge University Press.

PUTNAM, Hilary, 2002. The Nature of Mental States. *In*: David J. CHALMERS, ed. *Philosophy of Mind: Classical and Contemporary Readings.* Oxford: Oxford University Press, 2002, pp.73-79.

REA, M., 2014. *Metaphysics: the basics*. London: Routledge.

ROBINSON, Richard, 1972. *Definition.* Oxford: Clarendon Press.

ROSS, W.D., 1950. *The Right and the Good.* Oxford: At the Clarendon Press.

RUSSELL, Bertrand, 1958. Mind and Matter. *In: Bertrand Russell, Portraits from memory and other Essays.* London: George Allen & Unwin, 1958, pp.140-160.

RYLE, Gilbert, 1963. Systematically misleading expressions. In: Antony FLEW, ed. *Logic and Language (First Series)*. Oxford: Basil Blackwell, 1963, pp.11-37.

RYLE, Gilbert, 1976. *The Concept of Mind.* Harmondsworth, Middlesex: Penguin Books.

SCHAEFFER, Peter, n.d. Review of *Mind, Brain, and Free Will*. By Richard Swinburne. Available at: peterschaefer.files.wordpress.com [Accessed 11/10/15].

SHOEMAKER, S. and SWINBURNE, R., 1984. *Personal Identity.* Oxford: Blackwell.

SCRUTON, R., 1994. *Modern Philosophy: an introduction and a survey.* London: Stevenson.

SMITH, Peter and JONES, O.R., 1993.*The Philosophy of Mind: An Introduction.* Cambridge: Cambridge University Press.

STROLL, Avrum, 2000. *Twentieth Century Analytic Philosophy*. New York: Columbia University Press.

SWINBURNE, Richard, 1977. *The Coherence of Theism.* Oxford: Clarendon.

SWINBURNE, Richard, 1987. *The Evolution of the Soul*. Oxford: Clarendon Press.

SWINBURNE, Richard, 2004.*The Existence of God.* 1st ed. 1979. Oxford: Oxford University Press.

SWINBURNE, Richard, 2013. *Mind, Brain, and Free Will.* Oxford: Oxford University Press.

SWINBURNE, Richard, 2013a. *Mind, Brain, and Free Will*. *The Montréal Review* May. Available at: www.themontrealreview.com. Accessed 11/10/15.

SWINBURNE, Richard, 2014a, Précis of *Mind, Brain, and Free Will*. *European Journal for Philosophy of Religion*, 6 (2), Summer, pp.1-3.

SWINBURNE, Richard, 2014b. Response to reviewers. *European Journal for Philosophy of Religion*, 6 (2), Summer, pp.51-63.

TAYLOR, Richard, 1983, *Metaphysics.* 3rd. ed. Englewood Cliffs, New Jersey: Prentice-Hall.

VEALE, Graham, 2013. Review: *Mind, Brain, and Free Will* by Richard Swinburne. Apologetics 315.com, Saturday, May 11, [Accessed 10/10/15]

WEGNER, Daniel, 2002. *The Illusion of Conscious Will*. Cambridge, MA: The MIT Press.

WILLIAMS, Bernard, 2006. Identity and identities. *In*: Bernard WILLIAMS, *Philosophy as a humanistic discipline*. A.W. MOORE, ed. Princeton: Princeton University Press, 2006, pp.56-64.

WILLIAMS, Bernard, 1982. *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press.

WILLIAMS, Bernard, 1982a. Personal identity and individuation. *In:* Bernard WILLIAMS, *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press, 1982, pp.1-18.

WILLIAMS, Bernard, 1982b. Bodily continuity and personal identity. *In*: Bernard WILLIAMS, *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press, 1982, pp.19-25.

WILLIAMS, Bernard , 1982c. Imagination and the self. *In:* Bernard WILLIAMS, *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press, 1982, pp.26-45.

WILLIAMS, Bernard, 1982d. The self and the future. *In:* Bernard WILLIAMS, *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press, 1982, pp.46-63.

WILLIAMS, Bernard, 1982e. Are persons bodies? *In:* Bernard WILLIAMS*, Problems of the Self: Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press, 1982, pp.64-81.

WITTGENSTEIN, Ludwig, 1978. *Philosophical Investigations.* Trans. G.E.M. Anscombe. Oxford: Basil Blackwell.

WITTGENSTEIN, Ludwig, 1980. *The Blue and Brown Books: Preliminary Studies for the Philosophical Investigations.* 2nd ed*.* Oxford: Basil Blackwell.

WITTGENSTEIN, Ludwig, 2008. *On Certainty*. G.E.M. ANSCOMBE and G.H. von WRIGHT eds. Denis Paul and G.E.M Anscombe trans. Oxford: Blackwell.