

The digital version of the doctoral thesis *Death and the Self. A Metaphysical Investigation of the Rationality of Afterlife Beliefs in the Contemporary Intellectual Climate* by Johan Eddebo, is published without poems on pp. 13, 194 and 195.

For the full version, see ISBN 978-91-513-0137-2 (printed version).

Acta Universitatis Upsaliensis  
Uppsala Studies in Philosophy of Religion

Cover picture: Maria Näsström.

“if i believe”. Copyright 1923, 1951, (c) 1991 by the Trustees for the E. E. Cummings Trust. Copyright (c) 1976 by George James Firmage, “why did you go”. Copyright 1923, 1951, (c) 1991 by the Trustees for the E. E. Cummings Trust. Copyright (c) 1976 by George James Firmage, from COMPLETE POEMS: 1904–1962 by E. E. Cummings, edited by George J. Firmage. Used by kind permission of Liveright Publishing Corporation.

Johan Eddebo

# Death and the Self

A Metaphysical Investigation of the Rationality of Afterlife  
Beliefs in the Contemporary Intellectual Climate



UPPSALA  
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Geijersalen, Engelska parken, Thunbergsvägen 3, Uppsala, Friday, 15 December 2017 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Charles Taliaferro (Department of Philosophy, St. Olaf College, Northfield, Minnesota).

### **Abstract**

Eddebo, J. 2017. Death and the Self. A Metaphysical Investigation of the Rationality of Afterlife Beliefs in the Contemporary Intellectual Climate. *Uppsala Studies in Philosophy of Religion* 3. 206 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0137-2.

This dissertation's purpose is to test the hypothesis that beliefs in the possibility of post-mortem survival can be rationally held within the context of the contemporary scientific and philosophical environment. In terms of criteria of rationality, a basic evidentialism is assumed, such that propositions which are sufficiently supported by the available evidence can be rationally held. With regard to the compatibility with contemporary science and philosophy, it follows as a further criterion that the relevant evidence must be satisfactorily anchored within the framework of these traditions.

The relevant evidence concerns two levels. First, the basic level of the conceptual coherence of afterlife beliefs is addressed, so that the logical possibility of post-mortem survival can be established. Secondly, the viability of the metaphysics which are implied in the support of the logical possibility (i.e. the metaphysics needed to actualize post-mortem survival) is defended, establishing the metaphysical possibility of post-mortem survival. At this stage, reductive physicalism, which is the only position that effectively undermines post-mortem survival, is criticized, and the problem of interaction which burdens several of the survival-enabling ontologies is addressed.

As for the criterion of scientific compatibility, it is further shown that contemporary physics are compatible with the survival-enabling metaphysics, and that contemporary physics can be argued to provide a moderate positive relevance with regard to these positions.

The conclusion drawn is that belief in the possibility of post-mortem survival is not only rationally permissible within the framework of contemporary science and philosophy, but also rationally obligatory, i.e. that this possibility cannot rationally be denied with regard to the reviewed evidence.

*Keywords:* death, subjectivity, selfhood, afterlife, irreducibility, critique of naturalism, immaterialist ontologies, dualism, idealism, philosophy of religion, quantum theory, interpretations of quantum mechanics, modernity

*Johan Eddebo, Studies in Faith and Ideologies, Philosophy of Religion, Box 511, Uppsala University, SE-751 20 Uppsala, Sweden.*

© Johan Eddebo 2017

ISBN 978-91-513-0137-2

urn:nbn:se:uu:diva-332097 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-332097>)

*To St. Edith Stein*



# Contents

Acknowledgments.....	11
1. Introduction .....	15
1.1. Introduction and overview.....	15
1.1.1. General remarks on the purpose and subject .....	17
1.2. Purpose .....	20
1.3. Empirical framework.....	22
1.4. The current state of research on the viability of afterlife beliefs.....	22
2. Method, goals and prospects.....	26
2.1. The cumulative case .....	29
2.1.1. The cumulative case for the rationality of afterlife beliefs .....	31
2.2. Foundations of the case .....	31
2.2.1. Intelligibility .....	32
2.2.2. Validity and soundness .....	32
2.3. Detailed structure of the cumulative case.....	33
2.4. Summary .....	34
3. The intuition of post-mortem survival .....	35
3.1. Introduction .....	35
3.2. Intentional burial and primordial forms of belief in an afterlife.....	37
3.3. An argument for the historical presence of a continuity intuition.....	38
3.3.1. The intuition of post-mortem continuity.....	41
3.3.2. The intuition of post-mortem continuity and the first-person perspective .....	47
4. Death and post-mortem survival .....	49
4.1. Introduction .....	49
4.1.1. Definition of death and post-mortem survival .....	49
4.2. Legacy problems of death and the afterlife .....	51
4.2.1. The problems of personal identity .....	52
4.2.2. The synchronic and diachronic problem in view of the phenomenological criterion .....	52
4.2.3. The problem of individuation .....	57
4.2.3.1. The phenomenological criterion and spatio-temporal individuation .....	59



4.3. Basic possible responses to the legacy problems from the major metaphysical positions .....	61
4.3.1. Reductionist physicalism: mnemonic and bodily continuity .....	62
4.3.2. Dualism, idealism and non-reductive materialism: immaterial continuity, composite continuity and phenomenal presence .....	65
4.3.2.1. Immaterial continuity .....	65
4.3.2.2. Composite continuity .....	67
4.3.3. Brief assessment of the remaining metaphysical positions.....	70
4.4. Prominent arguments for post-mortem survival: physicalist-friendly arguments.....	72
4.4.1. Zimmerman's case.....	72
4.4.2. Van Inwagen's argument .....	76
4.4.3. Trenton Merricks' account .....	77
4.5. Immaterialist arguments .....	78
4.5.1. Swinburne's argument .....	79
4.5.2. Hasker's argument .....	80
4.6. Conclusions .....	80
5. A basic metaphysics of mind: the ownness model of the minimal self ....	84
5.1. Introduction .....	84
5.2. The basics of the model.....	85
5.2.1. On the presence and character of subjective consciousness .....	86
5.2.2. On the indivisibility and unity of subjective consciousness .....	87
5.3. The ownness model: the first-person perspective as the minimal self ..	92
5.4. The ownness model in relation to alternative conceptualizations .....	93
6. A critique of reductive physicalism .....	95
6.1. Introductory remarks .....	95
6.2. What is reductive physicalism? .....	96
6.3. Why does reductive physicalism emerge? .....	98
6.4. Which arguments effectively undermine reductive physicalism? ...	102
6.4.1. On eliminative materialism.....	102
6.4.2. On non-eliminative reductive physicalism .....	107
6.4.2.1. Conceptual reductionism.....	107
6.4.2.2. The identification-problem of analytical reductionism ....	108
6.4.2.3. Non-eliminative metaphysical reductionism.....	109
6.4.2.4. Pure agnostic functionalism .....	110
6.4.3. Broad-scope arguments .....	111
6.4.4. Indispensability arguments .....	112
6.4.5. The physical's phenomenal dependency .....	113
6.4.6. Functions and intentionality .....	117
6.4.7. Incommensurability arguments.....	123
6.4.8. Baker's argument from phenomenal self-reference.....	124
6.5. Conclusions .....	129

7. The problem of interaction and non-monist metaphysics .....	131
7.1. The Cartesian origins of the interaction problem .....	132
7.2. Cartesian substance dualism .....	133
7.3. The problem of interaction .....	135
7.3.1. Basic responses to the problem of interaction .....	138
7.3.1.1. Negating the principle: Swinburne's response .....	138
7.3.1.2. Broadening causality .....	140
7.3.1.3. The compromise of emergent dualism and property dualism .....	140
7.3.1.4. The hylemorphic response .....	142
7.4. The problem of interaction and the metaphysics of causation .....	144
7.4.1. Four basic models of causation .....	145
7.4.2. Causal-model approaches to the problem of interaction .....	149
7.4.2.1. The proxy approach .....	149
7.4.2.2. The primitivist response .....	151
7.4.2.3. Causality-modifying direct interaction .....	154
7.5. Conclusions and further considerations .....	155
8. Quantum mechanics, phenomenal consciousness and transsubstantial interaction .....	157
8.1. Introduction .....	157
8.2. The metaphysical implications of quantum mechanics .....	157
8.2.1. Planck, Einstein and the quantum stone age .....	158
8.2.2. Wave-particle duality and the measurement problem .....	160
8.2.3. Von Neumann, Stapp and the Copenhagen orthodoxy .....	162
8.2.4. Non-orthodox approaches .....	167
8.2.5. Neorealism and hidden variables .....	169
8.2.6. Bell's interconnectedness theorem and non-locality .....	170
8.2.7. Criticisms of Bell's theorem: superdeterminism, the time- symmetric theories and the transactional interpretation .....	175
8.2.8. Remaining alternatives to the old orthodoxy .....	178
8.3. Consensus and consequences .....	179
8.3.1. Ontological implications .....	181
8.3.2. The unorthodox path towards immaterialism .....	183
8.4. Conclusions .....	186
9. Conclusions .....	189
9.1. On natural immortality .....	189
9.2. Evaluating the case .....	190
9.3. Concluding remarks on theism .....	192
10. Bibliography .....	196



# Acknowledgments

Thanks be to whom thanks are due. As my old vicar Leif Nordenstorm stated at the beginning of his own dissertation, it is entirely a grace of God that this work has been completed.

Sometimes an entirely pleasant journey, sometimes even a harrowing one, it would never have brought me to its destination were it not for the loving support, encouragement, scholarly advice and sharp criticism I have received from a large number of individuals. First and foremost of these are my supervisors Mikael Stenmark and Karin Johannesson, who for many years have played invaluable roles in my education as well as the genesis and development of this research project, and unlike myself have had an unwavering faith in my ability to bring this project to fruition. Of an equal importance has been my wife, Rebecka, for whose insightful comments on nearly every weighty issue throughout the dissertation I am thoroughly indebted – not to speak of her patience with the philosopher’s monomania to which I have exposed her and the steadfast support she has afforded me throughout the project. Another figure the value of whose contributions are inestimable is Ulf Jonsson, who criticized a previous draft of this dissertation during my concluding seminar, greatly helping to amend it.

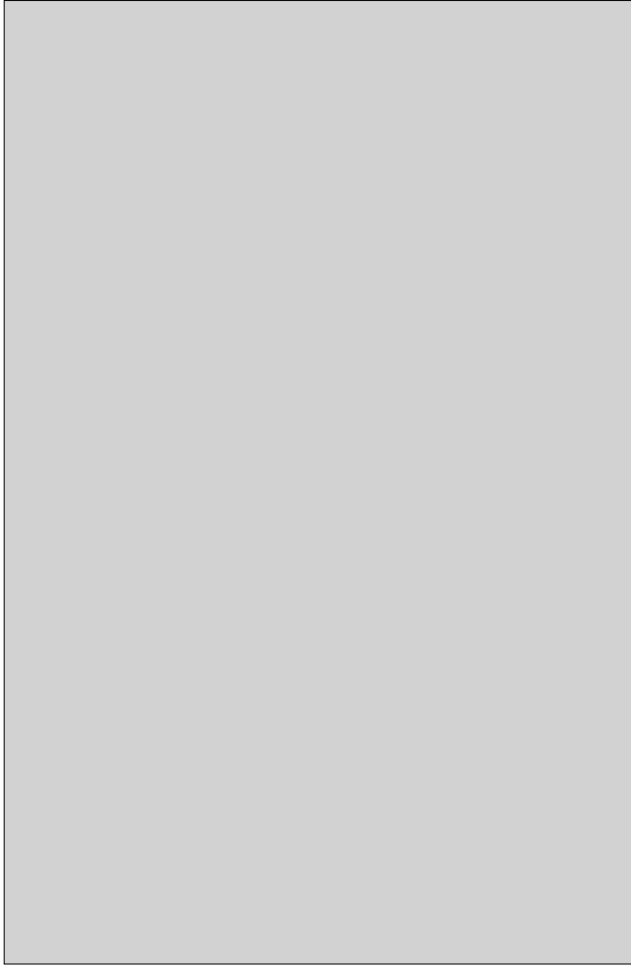
Many thanks are due to Uppsala University, The Carnegie Foundation and The Olaus Petri Foundation for funding my research.

I would as well like to express my gratitude towards the higher seminar in Philosophy of Religion at Uppsala University for the countless discussions and critical comments which have aided me in this work. Noteworthy participants thereof are Mikael Leidenhag, Oliver Li, Mikael Söörhus, Francis Jonbäck, Ulf Zackariasson, Maria Klasson Sundin, Lotta Knutsson Bråkenhielm, Carl-Reinhold Bråkenhielm, Elena Kalmykova, Christoffer Skogholt, Ingrid Malm Lindberg, Dan-Johan Eklund, Clare Mellqvist and Lina Åhlfeldt.

Other persons I wish to thank are Leemon B. McHenry for providing me with valuable material and comments on metaphysical idealism, Kicki Näslund and Trox Sjödin for the occasional room and board and the latter for

a sobering if sometimes scathing criticism of my work, Emil Lundin for spurring me to fully embrace the truths I had rationally acknowledged, Mattias Lundman and Eberhard Herrman for providing, in quite different ways, much of the inspiration and encouragement that originally convinced me to embark upon this journey, and finally my parents Hans and Yrsa Eddebo, as well as my mother-in-law Miriam Sundqvist, for allowing me to stay with them and work on this project during the endless Nordic summer days.

Gävle, late September 2017



---

<sup>1</sup> E. E. Cummings, G. J. Firmage (ed.) “why did you go”, in *Complete Poems, 1904 – 1962*, New York: Liveright Publishing Corp. 1992 p. 30, first published as "V" in “Seven Poems”, in *The Dial*, Vol. 68, No. 1, January 1920



# 1. Introduction

## 1.1. Introduction and overview

Ideas pertaining to death, a possible afterlife and the search for a coherent understanding of the two concepts, have seemingly occupied the human imagination ever since our prehistory. The very oldest preserved written texts give testimony of a preoccupation with this fundamental mystery of human existence, which evidently has remained more or less constant up until today.<sup>2</sup> But also such artifacts and remains from widely different locations all over the world, likely created before mankind's conceptual worlds had crystallized into anything approaching coherent bodies of tradition, carry obvious signs of behavioural patterns of the meaning-making variety linked to the demise of specific individuals.<sup>3</sup> Confronted with one's own as well as significant others' seemingly inevitable mortality, human beings have throughout history attempted to compose strategies with the purpose of making it possible to if not fully comprehend the relentless finitude of life itself, at least cope with it in a remotely meaningful and worthwhile manner. A significant part of such strategies has throughout the ages manifested in the form of ideas concerning some variant of an afterlife – the oldest preserved fragments of formal systems of belief arisen within the framework of complex human societies, seems to have comprised these ideas from the very beginning. Remains from the Predynastic period of ancient Egypt bear witness that the dead were buried with such peculiar care, and before their departure equipped in such a manner that the explicit ideas regarding an afterlife which the much later writings tell us of, very well may have been a living mythology in the minds and practices of the people as early as the very first permanent settlements in the region, established during the Late Stone

---

<sup>2</sup> The complex mythology reflected in the so called "Pyramid Texts", one of the oldest extant religious texts, provides a clear example of such notions which likely also indicates their presence in older oral tradition. (James P. Allen, *The Ancient Egyptian Pyramid Texts*, Atlanta: Society of Biblical Literature 2005).

<sup>3</sup> Cf. *The Paleolithic Origins of Human Burial* by Paul Petitt for an overview of the phenomenon. Pp. 162-163 in Philip Lieberman's *Uniquely Human*, Cambridge: Harvard University Press 1991 describes the oldest known instance of ceremonial burial performed by humans, discovered in the caves of Qafzeh and Skuhl in what is today northern Israel. The findings are just shy of 100,000 years old, and contains a number of interesting ritual objects, there among the jaw bone of a boar (a remarkable kill in regard to the contemporary hunting practices) which have been placed in the hands of one of the dead.



Age.<sup>4</sup> Traces of the Sumerians' religious worldview remaining in the oldest extant writings in the history of the world, which thereupon informed later Mesopotamian mythology, tell of vivid and complex conceptions of an after-life according to which the separation of a human's two independent aspects, body and spirit, which death entailed, preceded a new type of existence for the individual, entrance into a realm of the dead wherein her essence would meander throughout eternity.<sup>5</sup>

In prehistoric China, we also find burial practices similar to those of ancient Egypt, in which articles for everyday use were regularly buried together with the dead as early as 6,000 BC, suggesting of the later articulated practices of ancestor-worship.<sup>6</sup> Later remains, still preceding the advent of writing, indicate well-developed ideas regarding a reciprocal relationship between the living and those passed on, characterized by an exchange of gifts and services in a kind of honor-based, cultic economy, transcending the boundaries between the respective spheres of the living and the dead.<sup>7</sup>

Models of the afterlife similar to these are likewise seen among the vast number of pre-Colombian cultures who have left any tangible remains of their culture behind.<sup>8</sup> Oral traditions local to the area formerly dominated by the Aztec Empire, recorded by Spanish explorers during the latter part of the 16<sup>th</sup> century, give us a picture of a popular mythology describing an afterlife in which the dead were reunited with their relatives and loved ones, together with whom they finally faced their creator. The roots of this narrative are found in the ancient Meso- and South American civilizations' rich bodies of myth, where a colourful set of ideas regarding a life after death are to be found.<sup>9</sup>

As can be seen from a cursory review of the literature, human beings in virtually every culture that has existed throughout history, with few exceptions, preserved peculiarly similar narratives pertaining to the afterlife of the individual. These narratives have often been associated with ceremonial burial and other rituals associated with bereavement, and has thus very likely functioned as a tool for coping with the trauma that the demise of one's kin nec-

---

<sup>4</sup> Françoise Dunand, Roger Lichtenberg, *Mummies and Death in Egypt*, Ithaca: Cornell University Press 2006, pp. 7-8

<sup>5</sup> Alexander Heidel, *The Gilgamesh Epic and Old Testament Parallels*, Chicago: University of Chicago Press 1949, pp. 138-139

<sup>6</sup> Randall Nadeau (red.) *The Wiley-Blackwell Companion to Chinese Religions*, Malden: Blackwell Publishing Ltd. 2012, p. 29

<sup>7</sup> Ibid.

<sup>8</sup> See Elizabeth Benson, *Death and the Afterlife in Pre-Columbian America*, Washington: Dumbarton Oaks Library and Collection 1975

<sup>9</sup> Gregory Shushan, *Conceptions of the Afterlife in Early Civilizations*, London: Continuum Publishing 2009, pp. 135-136

essarily entails, but they have evidently also been part of the wider meaning-making practices through which mankind as an interpretative species generally try to make sense of, and relate to a capricious and often unpredictable environment.

However, the fact that such conceptualizations intended to describe some form of afterlife in certain respects seem to have similar basic traits, must not be interpreted as indicative of any common, specific ontologies, i.e. that people throughout history who in separate contexts conceived of a life beyond death, because of this necessarily shared a particular metaphysics. Our human consciousness has indeed approached ultimate reality exhibiting a great diversity of interpretations, as well as the questions regarding what it means to be human; where the boundaries between the self and others really lie; and exactly what it is in a person actually may transcend death.

What we in a very general sense may surmise from these ostensibly common aspects of meaning-making practices seen in widely different cultures, is therefore little more than that some aspect of a subject in some sense seems to have been considered capable of transcending the death of the body.

#### 1.1.1. General remarks on the purpose and subject

This possible transcultural intuition of the subject's post-mortem continuity, which seems to be present in our common ideological heritage since before even the emergence of full behavioural modernity, is the focal point of this dissertation. Can the implication of this intuition, i.e. the subject's survival of bodily death in any form, possibly take place? Is it rational to hold beliefs of this kind with regard to modern science and philosophy, and can claims to the effect that there actually is some form of afterlife, be made plausible in this context? These kinds of questions will be at the center of this dissertation.

Obviously, a vast amount of material has been produced on this topic throughout history. Some of the very earliest written works ever produced already address the possibility of survival of death, in ways that with regard to their complexity and ingenuity have been surpassed by few subsequent thinkers. For this reason alone, it may seem incredibly presumptuous to imagine that anything whatsoever could be added to the discussion. Yet, to paraphrase Oswald Spengler, every culture, at the dawn of every new era, faces the task of embracing and affirming a coherent view of the human person and its nature, whether or not it desires to. This seems to be something that the global, westernized civilization in the borderlands between

stagnant late modernity and an embryonic post-globalized is yet to take upon itself.<sup>10</sup>

Therefore, there seems to be at least some opportunity to provide potentially fruitful insights in support of this process of self-reflection with the analysis that this dissertation intends to provide. In view of this, it is hoped that this work may contribute in some small sense to better anchor the discussion regarding the inner life of the human person, and our commonly inherited intuition that this inner life in some sense is not necessarily limited by our inevitable physical demise, in the important traditions of knowledge relevant to contemporary thought and worldviews – without therefore disregarding the historical continuity and contextuality of these traditions.

There are other important reasons to consider the relationship between our understanding of the subject at hand and contemporary traditions of knowledge to be imperfect. One is, for instance, the limited extent to which the 20<sup>th</sup> century's developments within natural science have been fully utilized in the relation to philosophical theories on the human subject, which is still often approached on post-Cartesian materialist terms, i.e. on the assumption of a reductive physicalism according to which intentionality and subjectivity have no place in a basic ontology.<sup>11</sup> Another obvious reason is the various debilitating philosophical problems reductive materialism has faced, especially in relation to the mind-body problem.

The persistence of these specific early-modern metaphysical foundations, in spite of prominent alternatives, has been addressed in various ways. Adorno and Horkheimer and the tradition of Critical Theory considered the instrumental rationality of the modern era to be a prime cause of the situation. They argued that Enlightenment metaphysics were given a particular ideological function, which entrenched and preserved it. The economization of thought and the utilization of philosophy mainly in search for profit and practical benefits, rather than truth as such, are taken to have provided these

---

<sup>10</sup> Oswald Spengler, *The Decline of the West*, Oxford: Oxford University Press 1932 pp. 158-160

<sup>11</sup> This seems to be the only plausible analysis of physicalism unless it is to collapse into some variant of non-reductive physicalism or property dualism. It is also the standard understanding of physicalism, cf. "Materialism", in Donald M. Borchert, (ed.) *Encyclopedia of Philosophy*, vol. 6, Thomson Gale 2006, or "Materialism", in Edward Craig (ed.) *Concise Routledge Encyclopedia of Philosophy*, New York: Taylor & Francis Routledge 2000. Alternative formulations such as global supervenience physicalism, cannot successfully rule out dualistic ontologies, cf. Jaegwon Kim, "'Strong' and 'Global' Supervenience Revisited", in *Supervenience and Mind, Selected Philosophical Essays*, Cambridge: Cambridge University Press, 1995

metaphysical foundations with an ahistorical, insular and opportunistic character that preserved them from disrupting influences.<sup>12</sup>

The inertia of the Cartesian-Newtonian paradigm, together with the insular character of the traditions of knowledge it engendered, could also be understood in terms of the thoroughgoing penetration of the paradigm's general narrative at the time of our first globally interconnected civilization of the modern era. The subsequent reluctance of late modernity to erect any equivalent meta-narratives as a comprehensive replacement, may thereupon, to some extent, explain the persistence of the paradigmatic metaphysical foundations, even after their theoretical rationale has been challenged.<sup>13</sup>

Critical-theoretical and structuralist analyses provide a further possible explanation for the inertia of the Cartesian-Newtonian paradigm, or what from these perspectives is often approached as the ideological nexus of progress, instrumental reason, materialism and disinterested objectivity. In essence, such analyses have suggested that the significant ideological role of the institutions of natural science in modern Western society has tended to preserve a certain worldview. The importance of these institutions is, among other things, founded upon their ability to provide stable, cohesive and otherwise useful narratives within a particular economic framework. The narratives, in turn, rest upon certain paradigmatic discourses, such as the Cartesian-Newtonian ones. These paradigmatic discourses then tend to function as a corrective with regard to alternative theoretical positions that challenge or destabilize the ideologically important narratives provided by the institutions of natural science.<sup>14</sup>

Notwithstanding how this aspect of the contemporary zeitgeist is characterized, it can be argued that it has narrowed the prospects for the exploration of unconventional theoretical perspectives in science and philosophy. This would relate to this dissertation's subject of inquiry as something of a hindrance to a comprehensive, empirically and philosophically informed examination of the nature of the self, and of what may be considered a transcultural intuition of the post-mortem persistence of the self. Hopefully, then, this work will in some small way contribute to furthering the discussion of the self and post-mortem survival beyond these obstacles.

---

<sup>12</sup> Max Horkheimer, Theodor Adorno, *Dialectic of Enlightenment*, Stanford: Stanford University Press 2002, pp. xvii-xviii; p. 232; see also Stanislav Grof, *Beyond the Brain: Birth, Death and Transcendence in Psychotherapy*, New York: State University of New York Press 1985

<sup>13</sup> Robert Boyd, Peter Richerson, *Culture and the Evolutionary Process*, Chicago: University of Chicago Press 1985, pp. 56-57

<sup>14</sup> Cf. e.g. Althusser's concept of the ideological state apparatus, Louis Althusser, *Lenin and Philosophy and Other Essays*, London: Monthly Review Press 1971, p. 127ff

## 1.2. Purpose

This dissertation will examine the hypothesis whether afterlife beliefs can be considered plausible, and held rationally, in the intellectual, philosophical and scientific climate of the contemporary Western world. In assessing this hypothesis, the coherence of the basic idea of post-mortem survival will first be examined, and a response to criticism of its coherence will be offered. In thus establishing the basic coherence of afterlife beliefs, one main account of post-mortem survival which is based in the phenomenological presence of the subjective first-person perspective will be suggested and developed. It will be argued that this account is the most viable option, in that it provides the best prospects of grounding a response to the basic coherency problems of post-mortem survival. Another benefit is that it's also compatible with the conceptual core of most traditional afterlife beliefs within religious contexts. Further on, the metaphysical commitments implied by this main account of post-mortem survival will be defended.

The key issues to be addressed with regard to coherence are the problems of identity: the diachronic and synchronic problem, and the problem of individuation. In terms of the metaphysical commitments, they will be negatively defended with a critique of reductive physicalism, the set of positions which are the least amenable to the account of post-mortem survival that is presented. Moreover, the interaction problem of dualism and various forms of non-reductive physicalism will be engaged with, and the indirect support for these “non-materialist ontologies” with regard to the metaphysical implications of quantum theory will be examined.

A thereby implied assumption of this dissertation is that of the epistemic significance of scientific practices, such that a philosophical position can to some extent be supported by its congruity with regard to well-supported data from the sciences. That is, philosophical positions, while obviously not being invalidated by an incompatibility with scientific data and theories, are considered to be supported by being compatible with, or implied by them. Still, this assumption will be bracketed, so as to merely stating that in a context where science is assumed to be epistemically significant in relation to worldviews and metaphysics, the plausibility of worldview-related beliefs and the rationality in holding them, is supported if they are congruent with generally accepted scientific theory. But there is no necessary connection: empirical science does not make or break metaphysics. In other words, we will assume the position that a coherent philosophical position based in the entailments of basic logic, can neither be decisively refuted nor proven by interpretations of empirical data, notwithstanding their epistemic significance.

However, since contemporary philosophy was characterized as somewhat insular, and as contemporary science is not fully theoretically coherent, it's not entirely clear exactly what the dissertation ought to relate to, with the purpose of examining whether such congruency can be established or not. In other words, how should we best go about to "anchor the discussion in the important traditions of knowledge relevant to contemporary thought and worldviews" as was previously stated?

This will be addressed in detail in the chapter on method. In terms of a few brief opening remarks, it can be said that this scientific anchoring of the evaluated hypothesis will be made in relation to well-established and universally accepted theory, and plausible interpretations thereof, which can be supplied with a sound philosophical backing. This can be formulated as a basic criterion of scientific compatibility. Without glossing over the conceptual, ideological and theoretical incoherence of modern science in terms of a set of institutions as well as a system of theories, it will thus be assumed that any well-established theory or set of theories from natural science, which possesses an adequate philosophical backing, can be utilized as such a point of anchoring.

This would arguably open the door for a *reductio* criticism on the grounds of eclecticism, also possibly formulated as the more obvious charge of irrelevance. Arguing along these lines, it could be claimed that a criterion of scientific compatibility understood in the way described, implies that any theory from the domain of science, no matter how philosophically implausible, unethical or otherwise problematic, would support the rationality of holding philosophical positions that are implied by them or are compatible with them, if the theories were majority positions or even only influential. This would certainly be true, however, due to the bracketing of the assumption of the epistemic significance of science, this support only holds contextually. That is, specifically in a society such as ours, where science is generally assumed to be epistemically significant, the congruity between science and a particular philosophical position renders the latter more rational to hold. Moreover, the assumed epistemic significance of science has an auxiliary rather than primary function. The philosophical coherence of the positions under scrutiny is the foundation of their plausibility, which in turn is supported by their congruity with descriptions of the world provided by science.

An assumption of the positive epistemic significance of such scientific descriptions is not at all necessary, and arguably not even preferable: if science fails to support the metaphysics implied by our best philosophy – so much the worse for science. However, in attempting to mirror the basic assumptions of contemporary Western society and culture, the hypothesis we are

examining will nonetheless implicitly assume the positive epistemic relevance of natural science.

### 1.3. Empirical framework

As previously stated, the intent of this project is to explore the possibilities of creating a relevant, philosophically viable and scientifically compatible description of post-mortem survival – the post-mortem continuity of subjectivity. This necessitates an anchoring in such scientific data and findings which has clear ontological significance in a way that permits us to consider and assess the relevant ideas and their conceptual frameworks; i.e. the data and findings must enable fruitful theorizing regarding the metaphysics of death and the afterlife; the relationship between mind and body; the ontology of mind etc.

Quantum theory is the primary field of scientific research to which the investigation relates. To be more precise, an important part of this investigation is associated with data and theory which has arisen within the theoretical structures of quantum mechanics. The reasons for this are fairly obvious; the scientific revolutions which have taken place within this and interrelating fields during the 20<sup>th</sup> century, has severely impaired the theoretical foundations for the yet lingering, reductionist “clockwork model” of the physical world, describing it as deterministically interrelated, separate physical objects, where little conceptual space for non-physical approaches to agency or consciousness is to be found. In removing the foundations of this model, quantum theory has thereby enabled a greater space for ontological theorizing than was previously available.<sup>15</sup>

### 1.4. The current state of research on the viability of afterlife beliefs

The topic of ours has historically been addressed by every generation in some way or another, our period being no exception to this general rule. Indeed, it may be argued that we’ve seen a renewed interest in the discussion as such during the last few decades, in some contrast to the preceding period wherein the tendency was either towards a negation of the possibility, or a dismissal of the issue.<sup>16</sup> More recently, several best-selling titles exploring

---

<sup>15</sup> Nick Herbert, *Quantum Reality*, New York: Anchor Books 1985 pp. xi-xv

<sup>16</sup> <http://www.slate.com/bigideas/is-there-life-after-death/essays-and-opinions>, accessed 2016-11-11

the issue have surfaced on the lists of major news outlets, and a three-year project funded by the Templeton Foundation aptly named *The Immortality Project* ran from 2012-2015, producing a good deal of material on the subject, covering everything from neuroscientific research to matters of philosophy and theology.<sup>17</sup>

The previous research relating to the dissertation's topic is massive. Some of the most important work can be exemplified by Flew's work on conceptions of immortality, of Hick's research on post-mortem survival in relation to a multireligious framework, as well as Habermas & Moreland's compilation of evidence for immortality in different conceptions.<sup>18</sup> Van Inwagen's work on the viability of resurrection, as well as William Hasker's support of physicalist resurrection models, have also been important in relation to the current discourse on the topic.<sup>19</sup> Lynne R. Baker's recent overview of the topic, which introduces the ontologically neutral criterion of phenomenal persistence which plays an important role in this dissertation, must also be mentioned.<sup>20</sup>

The main dividing line within research in the field at large, runs between the more empirically oriented work wherein issues such as near-death-experiences or NDEs are explored with regard to their implications, and the metaphysically oriented investigations wherein questions of the conceptual coherence of post-mortem survival and the nature of consciousness and human (and sometimes non-human) persons are at the forefront. *The Immortality Project* was a broad-front effort exploring many orthodox as well as more unusual interpretations of immortality, which clearly bridged this divide to some extent. In the actual publications resulting from the project, however, the divide remained relatively intact, with a clear separation between the science-based approaches and the theological and philosophical projects.<sup>21</sup>

---

<sup>17</sup> <http://www.sptimmortalityproject.com>, accessed 2016-11-11

<sup>18</sup> Cf. Anthony Flew, *Merely Mortal: Can You Survive Your Own Death?*, Amherst: Prometheus Books 2000; John Hick, *Death and Eternal Life*, Louisville: Westminster John Knox Press 1994; J. P. Moreland, Gary Habermas, *Beyond Death: Exploring the Evidence for Immortality*, Eugene: Wipf and Stock Publishers 2004

<sup>19</sup> van Inwagen, Peter, "The Possibility of Resurrection", *International Journal for the Philosophy of Religion*, 9 1978; William Hasker, "Materialism and the Resurrection: Are the Prospects Improving?", *European Journal for the Philosophy of Religion*, 3:1, 2011

<sup>20</sup> Lynne R. Baker, "Death and the Afterlife", in William J. Wainwright (ed.), *The Oxford Handbook of Philosophy of Religion*, New York: Oxford University Press 1997

<sup>21</sup> John Martin Fischer, "Results of funded research projects", *The Immortality Project*, July 2012, retrieved from <http://www.sptimmortalityproject.com/wp-content/uploads/2012/07/RESULTS-OF-FUNDED-RESEARCH-PROJECTS.pdf> 2016-11-11



Of the popular best-sellers in question, the focus has generally been on NDEs (near-death experiences) and the empirical or scientific support for afterlife beliefs. An example of a prominent such work is Eben Alexander's *Proof of Heaven* in which the author explores his own near death-experience in the light of his competence in the field of neurobiology,<sup>22</sup> another is Dinesh D'Sousa's *Life After Death* which combines empirical data with musings on the metaphysics of post-mortem survival.

With regard to the wider academic setting outside of the aforementioned project, an important recent event was the publishing of *The Oxford Handbook of Philosophy of Death* in 2012. This work is an anthology that in some depth explores the metaphysics and the ethical and philosophical problems associated with death and post-mortem survival, a groundbreaking compilation of contemporary research in the field.<sup>23</sup> Shelly Kagan's *Death*, also from 2012, based on his popular open course at Yale University, has become an important work on the topic in as well, arguing against the viability of most traditional forms of afterlife belief, based chiefly in his conceptualization of the human subject.<sup>24</sup>

Michael Sudduth's *A Philosophical Critique of Empirical Arguments for Postmortem Survival* from 2016, is one of few recent works that combines the popular focus on empirical issues with a more incisive philosophical analysis of the problems related to this particular line of inquiry. Sudduth focuses on the empirical data related to NDEs, mediumistic communication and reincarnation accounts, and finds the likelihood and Bayesian-style explanatory arguments wanting with regard to establishing positive immortality, i.e. unconditional survival of corporeal death, or "natural immortality" as Hegel phrased it.<sup>25</sup>

This dissertation differs from the above in that it will attempt to discern whether affirmative beliefs in post-mortem survival (i.e. the probability of our possible survival) can be considered rational with due regard to the contemporary philosophical and scientific environment, and in that it undertakes a closer look upon the metaphysical implications of contemporary science, rather than just the supposed empirical support for post-mortem survival such as NDEs or reincarnation accounts. What is also novel is the coupling of the issues of post-mortem survival with the broader assessment of the

---

<sup>22</sup> Eben Alexander, *Proof of Heaven*, New York: Simon & Schuster 2012

<sup>23</sup> Ben Bradley, Fred Feldman, Jens Johansson, *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012

<sup>24</sup> Shelly Kagan, *Death*, New Haven: Yale University Press 2010

<sup>25</sup> Michael Sudduth, *A Philosophical Critique of Empirical Arguments for Postmortem Survival*, New York: Palgrave MacMillan 2016 281ff

contemporary metaphysical landscape as such, and the discussion of possible empirical support for the survival-enabling ontologies found within quantum mechanics.

## 2. Method, goals and prospects

The purpose of this dissertation is to critically evaluate the hypothesis that the contemporary philosophical and scientific environment can provide a context within which beliefs in the metaphysical possibility of an afterlife can be rationally held, in accordance with a basic evidentialist framework.

Two major questions immediately arise from such a statement. What exactly does “the contemporary philosophical and scientific environment” refer to, and given our response to this question, in what way can the stated hypothesis best be evaluated?

Concerning the first question, the core idea is to ascertain whether such support for the rationality of holding afterlife beliefs can at all be provided in terms of convincing (plausibly sound) arguments acceptable within, and to some extent anchored in the mainstream of our multi-faceted contemporary scientific and philosophical environment.

One might contend that this is a vague declaration, that this supposed “mainstream” is an ever-changing entity with few if any persistent characteristics over time, and that at least some epistemic support for whichever fringe theory one chooses to espouse necessarily can be produced in relation to it. To avoid such criticism, it will be assumed that the evidence is contextually coherent *if* the following two sub-criteria are fulfilled:

1. the relevant arguments in support of afterlife beliefs are acceptable as valid and probably sound within the wider philosophical tradition (assuming the truth of the relevant premises)
2. at least some of said arguments are supported by well-established scientific theory or data

Having established this, the second basic question remains – the matter of how to best evaluate the hypothesis as such. How can we ascertain whether or not evidence successfully rendering afterlife beliefs rational is to be found, which simultaneously satisfies the above criteria for being anchored in the scientific and philosophical environment?

It would seem that the most reasonable avenue of approach is that of a general evidentialism, which would entail identifying the most promising arguments supporting the relevant propositions, and to evaluate their collective weight and whether or not some of them can be considered sufficiently supported by well-established scientific theory or data. On a basic formulation of evidentialism, a person's particular belief is rationally justified if it is supported by the evidence available to the person in question.<sup>26</sup> Also, in the specific context of philosophy of religion, arguments are considered to be the main form of justifying evidence (which does not exclude e.g. phenomenal experiences or empirical data, upon which arguments can be constructed).<sup>27</sup>

A belief that is justified according to the available arguments, the set of which also qualifies according to the two sub-criteria presented above, will thus be considered rational.

To identify the most promising arguments supporting the relevant propositions, and evaluate whether or not they can be considered to render them rational, the support for afterlife beliefs will be approached in terms of a cumulative set of arguments operating on two different levels. This approach intends to consider the collective weight of the evidence in terms of epistemic support, rather than constructing one major argument or a few arguments with a more narrow focus. This two-level cumulative case is, simply put, a type of cumulative argument involving the broader metaphysical framework as well as the particular support for the coherence of afterlife beliefs. The two levels refer to the support for the coherence of afterlife beliefs, and the support for the necessary metaphysical framework, respectively.

The reason for such an approach is that this particular form of a cumulative case, which functions by integrating arguments from two conceptual levels, provides broader warrant for the relevant type of propositions when compared to arguments with a more narrow focus. A combination of evidence operating on several levels arguably provides better prospects for providing the "supportive context" stated in the purpose, than would one single major argument, or a set of separate arguments with a more narrow focus.

Importantly, it should be noted that the form of cumulative reasoning employed in this situation has important basic differences in relation to what in philosophical discourse is commonly referred to as "the cumulative case argument", which generally is taken to employ a large number of non-

---

<sup>26</sup> Cf. Earl Conee and Richard Feldman. *Evidentialism: Essays in Epistemology*, Oxford: Clarendon Press, 2004

<sup>27</sup> Alvin Plantinga, *Warranted Christian Belief*, Oxford: Oxford University Press, pt. 2 2000

decisive arguments which together is taken to produce sufficient warrant, for instance by way of a probability calculation.<sup>28</sup> In contrast, the case presented in this dissertation involves purportedly decisive arguments in support of the metaphysical possibility of post-mortem survival given certain ontological commitments, as well as, in turn, a set of decisive arguments for the broader metaphysical framework enabling such ontological commitments. Thus, since several of the arguments presented on each level are by themselves sufficient to grant the truth of the relevant proposition, the “leaky bucket” response advanced by Flew in opposition to certain cumulative-type arguments is irrelevant.<sup>29</sup> Nonetheless, the provision of several ostensibly decisive arguments on each level is taken to imply a greater epistemic support than the isolated arguments can themselves contribute, due to the ensuing redundancy.

The dissertation’s purpose is thus not to examine the validity and soundness of one single decisive argument, or a few arguments with a similar focus, supporting the metaphysical possibility of post-mortem survival. The point is rather to examine whether beliefs in such a possibility ought to be considered rational in the aforementioned particular context, given the state of a set of arguments regarding the metaphysics of the human subject, the conceptual coherence of the notion of an afterlife, and the broader metaphysical framework.

By “operating on two levels”, the case is intended to provide two distinct avenues of support for the type of propositions in question. For the first level, the ostensibly “narrow” arguments for the veracity of afterlife beliefs will be examined. Examples of this would be arguments purporting to show that the very notion of the continued existence of persons beyond corporeal death is at all coherent, or arguments toward the coherence of a philosophical anthropology which renders such continuity possible. For the second level, the broader ontological background will be addressed, and arguments indicating the viability of such general metaphysics which are necessary if we are to understand the ontology of persons in such a way that a post-mortem existence is conceivable. As an example of the importance of the second level, it may well be the case that specific arguments for e.g. a dualist ontology of the person (assuming that dualism is congenial to afterlife beliefs) convincingly show that a person may possibly be separated from her body, but that the general metaphysics needed to support such an anthropology for various reasons seems untenable, for instance, because it is taken to be contraindicated by all ontological implications of natural science, or because of the

---

<sup>28</sup> Evans, C, *Pocket Dictionary of Apologetics & Philosophy of Religion*, Downers Grove: Intervarsity Press: 2002

<sup>29</sup> Antony G. N. Flew, *God and Philosophy*, London: Hutchinson 1966 p. 63

severity of the problem of interaction. Therefore, the general rationality of an ontological background congenial to the first-level arguments needs to be supported in turn, if the case as such is to be entirely viable.

Finally, we will distinguish between *rational permissibility* and *rational obligatoriness*, such that a belief (e.g. in the possibility of post-mortem survival) is rationally permissible if the notion as such is coherent and there is a reasonable possibility that the necessary background metaphysics are in place (and agnosticism can be evaded due to e.g. Stenmark's presumptionist principle),<sup>30</sup> whereas the belief is rationally obligatory if the notion is coherent, and the truth of the necessary background metaphysics cannot be denied without contradiction, or is supported by a very strong degree of probability.

## 2.1. The cumulative case

How, more exactly, does a cumulative case-argument in general look; how can it be defined and what are its criteria? The kind of reasoning it embodies was arguably popularized as a philosophical method by Richard Swinburne and Basil Mitchell, who with the concept basically referred to an interconnected set of limited arguments providing support for a "wider theory", a complex narrative different aspects of which is supported by various pieces of evidence.<sup>31</sup> The paradigm example is a less-than-trivial theory regarding some set of ancient historical events, such as the process of Julius Caesar's usurping of the Roman dictatorship after the Civil War. The factuality of these events is from the perspective of contemporary society's anchored in a multitude of historical and archeological evidence, such as preserved written records like Caesar's *Commentari De Bellum Civile*, Cassius Dio's *Historia Romana*, early secondary material such as Tacitus' *Historiae* and various archaeological material corresponding with information regarding important battles. The point is that no single piece of evidence by itself can prove the conventional account of the entire process: Caesar's own written commentaries alone would render the account dubious, to say the least, and merely the archaeological remains from a few key battles or troop movements would make a narrower theory more probable in comparison to the full-blooded conventional account. Taken together, however, the cumulative weight of the evidence renders the wider theory more probable than not, notwithstanding the fact that no single piece of evidence can be considered conclusive.

---

<sup>30</sup> Mikael Stenmark, *Rationality in Science, Religion and Everyday Life*, Notre Dame: University of Notre Dame Press 1995, p. 212 ff

<sup>31</sup> Richard Swinburne, *The Existence of God*, Oxford: Oxford University Press 2004, p. 19; Basil Mitchell, *The Justification of Religious Belief*, London: The Macmillan Press 1973

A further important difference besides not using “limited” arguments between Swinburne’s understanding and usage of the cumulative case, and the different one we are employing here, however, is that the propositions whose plausibility is under scrutiny do not constitute a “wider theory” of the kind that could not conclusively be proven by one or two pieces of evidence or arguments, like the conventional account of Caesar’s rise to power or the General Theory of Relativity. On the contrary, afterlife beliefs as they are considered in this dissertation (see chapter 4 on *Death and Post-Mortem Survival*) rather constitute a “narrow theory”, that is, a fairly limited set of propositions for which a few pieces of evidence can conceivably provide conclusive support. Neither will an explicit, detailed probability calculation be necessary in ascertaining whether there is functional support for the relevant propositions, since the formal implications of the available evidence will be more or less trivial to assess in relation to such a “narrow theory”.

Before we move on to examine the structure of the case in detail, it must finally be remarked that, arguably, a majority of cases made for particular philosophical positions in some general sense must be regarded as cumulative cases. This would be due to the fact that they almost without exception are built upon an interconnecting set of arguments towards the central proposition or propositions.

However, the particular version of the argument type this dissertation intends to present differs from a cumulative case in such a general sense, in that it intends to explicitly operate on the two levels previously described (evidential justification for the coherence of the beliefs as such, and the provision of arguments in support for a broader metaphysical context that enables the metaphysical commitments that the beliefs require), and in that it does not attempt to perform any explicit and detailed probability calculation, but, again, rather intends to render the relevant conclusions rational to hold with regard to the separate arguments on each level, further supported by the redundancy in supplying several decisive and/or supporting arguments.

### 2.1.1. The cumulative case for the rationality of afterlife beliefs

It is very hard for a man to defend anything of which he is entirely convinced. It is comparatively easy when he is only partially convinced. He is partially convinced because he has found this or that proof of the thing, and he can expound it. But a man is not really convinced of a philosophic theory when he finds that something proves it. He is only really convinced when he finds that everything proves it.<sup>32</sup>

G. K. Chesterton

We have assumed an evidentialist position, according to which a belief is rational to hold if it can be supported by the relevant evidence, broadly considered, that is available to a situated subject. Since particular evidence and arguments may be evaluated differently with regard to certain background assumptions, the importance of the broader framework in terms of worldviews or metaphysics is emphasized.

As stated, we will assume the general context of the contemporary Western intellectual environment with regard to science and philosophy. This can be regarded as the wider context in relation to which the case must provide warrant for beliefs in the metaphysical possibility of an afterlife. As previously mentioned, the need for anchoring our case in the contemporary scientific context can be considered in terms of a criterion of scientific compatibility, while the case will naturally be addressed in relation to contemporary philosophical research.

## 2.2. Foundations of the case

With the above remarks in mind, we may recall the central goal of our cumulative case – that is, to evaluate the hypothesis that a generalized belief in an afterlife can be rational when asserted within the contemporary philosophical and scientific environment. Afterlife beliefs will be understood in the sense of the affirmation of some form of post-mortem continuity of subjectivity (cf. chapter 4 for details), yet we have not set any criteria for the cumulative case-argument as such. Given the definition of rational belief presented, two essential criteria are obviously implied, that is, evidential justification for the coherence of the beliefs as such, and the provision of arguments in support for a broader metaphysical context that enables the metaphysical commitments that the beliefs require. However, before elaborating on exactly how this is to be attained, we will first address the more basic criteria of the cu-

---

<sup>32</sup> G. K. Chesterton, *Orthodoxy*, Stornoway: Reformation Press 2002, p. 46



mulative case-argument. When does an argument actually work within the specified environment?

Simply put, an argument will be taken to function, or be philosophically successful, if it's coherent and comprehensible, if the conclusions follow from the premises (i.e. if it's *valid*), and if the premises are true or can be rationally believed to be true (i.e. if it's *sound*). Similarly, the case will on the whole be considered successful if the conclusions can be established with regard to the arguments employed as parts of the cumulative case.

### 2.2.1. Intelligibility

First of all, we have what might be considered the most basic criterion, that of conceptual coherence, intelligibility or understandability. One must be able to state the argument as well as the proposition it's intended to make plausible in such a manner that it can actually be grasped, and formulated and held without any contradictions or obvious inconsistencies. Considering the central hypothesis of the dissertation, that the belief in a post-mortem continuity of subjective experience can be considered rational within a particular context, then the key components of this type of proposition as well as the justifications presented must be coherent and understandable. That is, all concepts relevant to the case and its implied conclusion, such as those of subjective experience and physical death, must be presented with sufficient clarity and without contradiction that a rational human being will be able to understand what is being stated and contended.

### 2.2.2. Validity and soundness

Following the intelligibility criterion, the arguments which make up the case must also function as valid and ostensibly sound arguments to establish the relevant conclusions. Validity is trivial to ascertain, an argument is valid if the conclusion is entailed by the premises, i.e. if the conclusion cannot be false if the premises are true.<sup>33</sup> Soundness, on the other hand, is oftentimes not, due to the issues of verifiability and general epistemology which inevitably accompany the concept. As it's not feasible to construct a coherent theory of truth, nor to conclusively solve the problems associated with determining when a non-analytic proposition or premise actually is true within the confines of this dissertation, our demands upon the cumulative case-argument will, as was just stated, be restricted to *ostensible soundness*. What this means in practice, is that the key components of the cumulative case must function as valid arguments, the truth of whose premises can be ration-

---

<sup>33</sup> L. T. F. Gamut (Johan van Benthem, Jeroen Groenendijk et al), *Logic, Language and Meaning*, Chicago: University of Chicago Press 1991, pp. 1-3

ally believed, and is not in turn strongly and plausibly contradicted by any sound defeating arguments (or implicit defeaters the negation of which is incompatible with the positive arguments).

### 2.3. Detailed structure of the cumulative case

As previously stated, the cumulative case, if it is to be considered successful, must render the purported conclusions plausible. This, as we have seen, entails evidential justification for the coherence of the beliefs as such, and the provision of arguments in support for a broader metaphysical context which enables the metaphysical commitments that the beliefs require.

In terms of direct and indirect evidential justification, arguments provided for two main positions will be assessed. First, whether a post-mortem continuity of subjective experience is in itself a coherent concept will be examined, and whether the philosophical problems associated with the notion of an afterlife in some form, can be overcome. Secondly, it will be assessed whether the metaphysical commitments regarding the ontology of the human person associated with the surpassing of the conceptual difficulties, can in and of themselves be defended – i.e. post-mortem survival may well be logically possible iff we assume the *prima facie* coherent metaphysical position M, yet if this position implies something metaphysically impossible, then so is post-mortem survival.

The provision of a broader context also has two main parts. First will be provided an analysis of the compatibility between the notion of the post-mortem continuity of subjective experience and the major basic views on the metaphysics of the human person, explaining how such continuity can be conceived with regard to these views, and which seem the most congenial to it. Secondly, a general defense of the category of metaphysical perspectives which are compatible with an affirmation of post-mortem continuity will be attempted.

An important corollary to both the evidential justification as well as the provision of the broader context is the proposition that the problem of interaction can be meaningfully addressed from the point of view of dualist, non-reductive or pluralist metaphysics. This proposition partially intends to support the position that non-physical subjects can actually exist as such, as well as to strengthen the general viability of dualist or pluralist metaphysics, which arguably to some extent have been undermined by the interaction problem.

## 2.4 Summary

The dissertation will examine the hypothesis that the contemporary philosophical and scientific environment can provide a context within which beliefs in the metaphysical possibility of an afterlife can be rationally held. The basic coherence of the notion of post-mortem survival will mainly be defended in chapter 4, and intends to establish the general logical possibility of post-mortem survival. The metaphysical possibility is dependent upon the viability of the metaphysical commitments entailed by our defense of the coherence of post-mortem survival. The extent to which the viability of said metaphysical commitments can be supported, determines whether or not post-mortem survival can be said to be metaphysically possible.

We will maintain that belief in the metaphysical possibility of post-mortem survival is *rationally permissible* if the notion as such is coherent, and if there is a *reasonable possibility* that the metaphysics implied by our defense of the coherence of the notion are true. A metaphysical position will be taken to have a reasonable possibility to be true if it's logically possible, and if there are no conclusive defeating arguments against the position that cannot be effectively responded to. In this case, agnosticism can be evaded by employing Stenmark's presumptionist principle, and it is thus rationally permissible to maintain any belief in the metaphysical possibility of post-mortem survival.

We will moreover maintain that belief in the metaphysical possibility of post-mortem survival is *rationally obligatory* if all the above holds, as well as if the truth of the necessary background metaphysics can either not be denied without contradiction, or if it is unambiguously supported by a very strong degree of probability. A metaphysical position will be considered probably true if it is supported by a majority of the relevant arguments that also lack conclusive defeaters.

### 3. The intuition of post-mortem survival

#### 3.1 Introduction

Bereavement and its cognate, the realization of one's own mortality, arguably the defining evils of our common human condition,<sup>34</sup> have shaped human culture ever since the dawn of behavioural modernity some 150,000 years ago.<sup>35</sup> The strategies employed for addressing, interpreting and understanding these problematic facets of existence are deeply intertwined with the roots of all complex human societies and are accordingly enormously varied, but with a few noteworthy exceptions (e.g. classical Stoicism, certain parts of ancient Judaism or modern secular materialism), they've generally tended towards some form of negation of the finality of a person's death. In this chapter, it will be argued that this tendency, traced from the earliest stirrings of organized symbolic culture during the Middle Paleolithic, to the great religious systems of civilization, in the broadest possible terms can be described as a *common intuition of the post-mortem continuity of subjectivity*. Such a metaphysically minimal description seems to be preferable before more pregnant alternatives, since it thereby is more likely to be relevant in regard to a multitude of worldviews and/or ontologies which possibly could not accommodate more precise definitions of being, of death, of the spirit or the soul, etc. On the basis of this intuition, arguably common to human experience, we will then in later chapters develop a suggestion of what, metaphysically, post-mortem survival must entail and eventually discern whether or not the affirmation of its possibility can be reasonable.

But wherein, one may ask, lies the value in identifying such a phenomenon, possibly common to the majority of human cultures, in regard to a discussion of the nature of human consciousness and the possibility of its post-mortem continuity? Surely, people throughout history have believed the most peculiar things, and the mere occurrence of an idea (or in this case, rather some-

---

<sup>34</sup> However, observations have indicated that several kinds of animals other than higher primates, do indeed also exhibit signs of emotions similar to those accompanying grief, cf. John Archer, *The Nature of Grief – The Evolution and Psychology of Reactions to Loss*, London; Routledge 1999, pp. 53-56;

<sup>35</sup> The oldest known use of ochre as a dye is dated to 164 kya before present, see C. W. Marean, M. Bar-Matthews, J. Bernatchez et al, "Early human use of Marine resources and pigment in South Africa during the Middle Pleistocene", in *Nature*. 2007; 449:905–908.

thing like the outline of an idea), no matter how widespread, could in itself never provide sufficient justification for its plausibility. Truth isn't a matter of popular opinion, most would claim. This is hard to contest, and the support for the possibility of some form of post-mortem survival to be found in the mere presence of even universal belief in such a notion is of the very weakest kind, unless perhaps one manages to erect something like an argument from religious experience on this basis.

This is, however, not the purpose. The utility of the intuition of continuity is in this case that it provides a conceptual basis for a broad, minimal notion of post-mortem survival of the greatest possible relevance across religious and cultural boundaries. The notion is put forth as a common denominator for ideas of the afterlife of human beings attested to in various religious traditions, and its purpose is twofold. First and foremost, it intended as a basic argument for the historical and contemporary importance of ideas similar or relatable to this intuition, for human cultures globally. Even if one does not accept that the features presented, taken together, paints an adequate approximation of the basics of every religious tradition's idea of post mortem-survival, the hope is that this "intuition of continuity" may still indirectly affirm a wide variety of ideas of an afterlife (such as the traditions of post mortem-continuity of the Dharmic traditions, wherein the metaphysics of the person is quite different in comparison to the Cartesian paradigm, and where a sharp distinction between salvation and many forms of "afterlife" is emphasized)<sup>36</sup>.

This is to say that the continuity-intuition presented as a broad, inclusive notion, is intended to help make the contributions of this dissertation accessible and relevant to as wide a spectrum of religious traditions as possible, and in line with this hopefully could be relatable to any worldview that includes an analogue to, or derivate of this intuition in its particular context.

Secondly, the intuition of continuity functions as the starting point for the main investigation of this dissertation, namely the connection between the phenomenal subjectivity of human beings and the possibilities of an afterlife, and as we shall see, helps us bring the issue of post-mortem survival into the domain of philosophy of mind in a clear and intuitive manner.

---

<sup>36</sup> Kedar Nath Tiwari, *Comparative Religion*, Motilal Banarsidass & Shri Jainendra Press 1983, p. 208-211

### 3.2. Intentional burial and primordial forms of belief in an afterlife

Our story of this primordial tendency of the human heart by a suggestive coincidence happens to begin in northern Israel; in Galilee and on the slopes of Mount Carmel, where in the early 20<sup>th</sup> century, a number of surprising archaeological findings were discovered. These consisted of ancient human remains, which contemporary methods have dated to between 80,000 – 120,000 years before present<sup>37</sup>, in circumstances indicating that some of them were deliberately buried.<sup>38</sup> This likely makes these remains found at the two adjacent sites of Qafzeh and Skhul (where Skhul is probably oldest) the very earliest known manifestation of symbolic human thought of any type, predating other prehistoric art forms such as cave paintings or the so-called “Venus figurines” by 40,000 years, at the very least.<sup>39</sup>

The Skhul cave is a shallow rock shelter located in the Nahal Meaorot canyon in the Mount Carmel range. Here, most findings were made in layer B, an archaeological layer characterized by Mousterian types of tool-making (a level of stone tools-technology wherein flint was used to produce smaller, knife-like tools, more advanced than the preceding hand-axes of the Acheulean industry that generally ended around 100,000 years before present).<sup>40</sup> The human remains found in this layer were associated with what seemed to be intentional ornamentations in the form of shells taken from sea snails, as these weren't found locally and thus probably were not used as a source of food.<sup>41</sup> One of the dead males were found holding a mandible of a wild boar, an artifact with obvious contemporary significance, if for no other reason than that of boars being ferocious and lethal adversaries even in the face of organized hunters from technologically advanced societies of our day.<sup>42</sup> The Qafzeh cave is likewise a rock shelter, found in today's Lower Galilee near

---

<sup>37</sup> Roger Lewin, Robert Foley, *Principles of Human Evolution*, Malden: Blackwell Publishing 2004, p. 385

<sup>38</sup> Erella Hovers, Steven Kuhn, *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, New York: Springer 2010, pp. 171-188

<sup>39</sup> Jean Clottes, *Chauvet Cave: The Art of Earliest Times*, Salt Lake City: University of Utah Press 2003, p. 214; Universität Tübingen, "Ivory Venus Figurine From The Swabian Jura Rewrites Prehistory", *Science Daily*, May 14 2009, retrieved from <http://www.sciencedaily.com/releases/2009/05/090514084126.htm> 2013-04-02

<sup>40</sup> Ofer Bar-Yosef, "The chronology of the Middle Paleolithic of the Levant", in T. Akazawa, K. Aoki, and O. Bar-Yosef (eds.), *Neandertals and Modern Humans in Western Asia*, New York: Springer-Verlag 1998, pp. 39-56

<sup>41</sup> Mariah Vanhaeran, Francesco d'Errico, Chris Stringer, Sarah L. James, Jonathan A. Todd, Henk K. Mienis, "Middle Paleolithic Shell Beads in Israel and Algeria", *Science* 312 (5781): pp. 1785–1788. doi:10.1126/science.1128139. PMID 16794076, retrieved 2013-04-02

<sup>42</sup> Paula Hammond, *The Atlas of the World's Most Dangerous Animals*, Tarrytown: Marshall Cavendish 2010 pp. 189-191

Nazareth, on one side of Mount Precipice.<sup>43</sup> At this site, six of the fifteen discovered human corpses clearly appear to have been purposefully buried.<sup>44</sup> Here we see a greater diversity of grave goods, although of a similar nature – a variety of decorative sea shells were found, together with remains from a number of large mammals, including horse and rhinoceros. The shells show possible indications of having been strung, as well as traces of ochre, and were transported to the site from the Mediterranean Sea some 35 kilometres away, implying deliberate ornamental use.<sup>45</sup> One of the most conspicuous burial decorations is the red deer antlers in the hands of a young boy of approximately 13 years of age, interred in a pit in the bedrock.<sup>46</sup>

### 3.3. An argument for the historical presence of a continuity intuition

These instances of ceremonial burial represent some of the very earliest indications of full human behavioural modernity.<sup>47</sup> It's noteworthy that such comparatively elaborate burial practices predate virtually every other manifestation of symbolic culture, which indicates the possible presence of what one is inclined to call a type of foundational metaphysics which is then later elaborated in subsequent forms of art and oral tradition. To speculatively postulate an explicit metaphysics beneath these products of what very likely was ritualistic behaviour is of course apt to be misleading due to the very limited data available, as well as the lack of any obvious interpretative basis. The simple act of intentional burial can by itself signify a great multitude of ontological convictions (or lack thereof), and may merely indicate the emotional attachment of the bereaved, and thus their desire to care for their deceased loved ones, without thereby implying any particular perspective on the nature of reality as such.

Nonetheless, it seems likely that the early human beings who performed these burials as members of a reflective, gregarious and communicative species, in relation to said practices at the very least entertained certain ideas

---

<sup>43</sup> Erella Hovers, *The Lithic Assemblages of Qafzeh Cave*, Oxford: Oxford University Press 2009, pp. 29-33

<sup>44</sup> Hovers & Kuhn 2010, pp. 185-188

<sup>45</sup> Daniella E. Bar-Yosef Mayer, Bernard Vandermeersch and Ofer Bar-Yosef, "Shells and ochre in Middle Paleolithic Qafzeh Cave, Israel: indications for modern behaviour" *Journal of Human Evolution*, Volume 56, Issue 3, March 2009, pp. 307-314

<sup>46</sup> Hovers & Kuhn 2010, pp. 185-188

<sup>47</sup> The concept of *behavioural modernity* is a bit ambiguous, but generally incorporates evidence of tool-making, use of pigment, cooking, self-ornamentation and burial (see Silvana Condemi, Gerd-Christian Weniger (eds.) *Continuity and Discontinuity in the Peopling of Europe*, New York: Springer-Verlag 2011, pp. 275-277)

which were regarded as significant. Given this, we may also safely assume that these ideas functioned much the same as important beliefs do in social settings we know from history as well as contemporary life – that is, they were passed on among friends, relatives and acquaintances, and if accepted or influential, they played a part in shaping the participants' view of the world as well as their behavioural patterns (it may be instructive to note that hominids have been physically capable of complex spoken language since the time of *Homo erectus*, the direct ancestor of modern humans, which is seen to have existed between 18,000,000 and 300,000 years ago)<sup>48</sup>. Markings of such symbolic behaviour is evident in these burial remains, even if the stories that once may have surrounded them are not to be heard today.

If we then for a moment leave the region of northern Israel, and move some 1,500 kilometers eastward to the Zagros mountains of Kurdistan, we are approaching the site of the oldest known burial of an ostensibly ceremonial nature performed by *Homo neanderthalensis*.<sup>49</sup> Contrary to what one might assume, this site at Shanidar cave in mount Bradost of Kurdistan is likely a bit more recent than those of Qafzeh-Skhul, and is dated to between 60 – 80,000 years BP.<sup>50</sup> Pollen samples were found in the vicinity of one of the ten burial sites, indicating the presence of entire flowering plants within the grave, which has been used to argue for relatively advanced symbolic behaviour among the Neanderthal, although there is a distant possibility that the site was contaminated by burrowing animals.<sup>51</sup>

The propagation of deliberate and possibly ritual forms of individual burial around this time period is seemingly not limited to modern humans, which tends to shroud its quite sudden origins in an even greater aura of mystery. The possible explanations for these findings are independent development of the behaviour, cross-cultural communication, or perhaps a combination of the two – but no matter the exact source of this phenomenon, it's nevertheless obvious that a certain behavioural pattern is taking root among contemporary populations, which in all likelihood is accompanied by an understanding of *why* these actions are performed, an idea of their purpose.

---

<sup>48</sup> Thomas Wynn, "Did *Homo erectus* speak?" *Cambridge Archaeological Journal* 8 (1), 78-81, 1998

<sup>49</sup> At Krapina cave in Croatia, bone fragments from about 70 individual Neanderthals were discovered in the early 20:th century, a possible sign of ritualistic behaviour in association with human remains, dated to about 130.000 – 110.000 BP. (Robert Jurmain, Lynn Kilgore, Wenda Trevathan, *Essentials of Physical Anthropology*, Belmont: Wadsworth 2009, p. 232)

<sup>50</sup> Ralph Solecki, *Shanidar: The First Flower People*, New York: Alfred Knopf Inc. 1971, p. 253

<sup>51</sup> Ralph Solecki, "Shanidar IV, a Neanderthal Flower Burial in Northern Iraq", *Science*, vol. 190, iss. 4217, 1975, pp. 880-881



If we were to approach the phenomenon of intentional or “ceremonial” burial in this era as a rudimentary social institution, emphasizing the repetition and the collective character of the action, it follows that the emergence of these group behaviours most likely was associated with some form of agreement upon purpose, whether implicit or explicit. Durkheim’s basic definition of the institution as the *beliefs and modes of conduct instituted by the collectivity* towards an end, will serve to illustrate the point.<sup>52</sup> The actions seem to have been collective endeavours due to their reoccurrence over time, as well as the grave goods that decidedly were procured as a result of communal hunting, which imply a collective sanctioning of the burials, if not direct participation as such. Furthermore, the lack of an obvious practical benefit inherent in this type of handling of deceased, lends itself to the conclusion that these actions most likely were performed in accordance with *at least some basic idea* as to the function of this elaborate care which was bestowed on the dead. That is to say, the actions were likely not performed as a matter of course due to some obvious practical benefit, and since they were hardly random, they were probably considered meaningful in some sense.

At the point in time when what we understand as deliberate burial among early humans could be considered something more than just unreflective imitation (if it ever merely amounted to this), it thereby must be approached as an intentional group conduct. The fact that this form of burial practice endures and proliferates from this point onward (possibly even across the borders between sub-species) gives further credence to an understanding of the phenomenon as functionally institutional, as these kinds of social entities are far more permanent than instances of mere imitation, since institutions tend to be reproduced by other aspects of the social organization to which they are linked.<sup>53</sup> Thus, had the behaviour not been supported by a basic understanding of purpose which could be communicated and passed further on, it would have been significantly less likely to spread to its historical extent. Still, it is of course possible that deliberate burial developed independently in several groups of modern humans due to a basic inclination common to the species, which would imply that the propagation of the behaviour may have preceded its institutionalization in the proper sense. Also, a likely indicator of such proper institutionalization of some form of human burial rituals is the increasing usage of red ochre in grave sites seen during

---

<sup>52</sup> Emile Durkheim, *The Rules of Sociological Method*, The Free Press: New York 1982 p. 45

<sup>53</sup> Seumas Miller, “Social Institutions”, *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2012/entries/davidson/> 2013-04-09

the Middle Paleolithic, which is seen to spread with the migration of human populations.<sup>54</sup>

It seems an inevitable conclusion that at some point in human prehistory, intentional burial becomes the proto-institution of ceremonial burial. So what is revealed by us being able to tentatively define primitive human burial as a social institution, what does this imply for our purposes? One obvious answer, which seems entirely modest, is that it thereby becomes clearer that the early human beings who buried their dead in a ceremonial fashion, sooner or later did this in relation to what can be called collective beliefs regarding the significance of the act. However, since the concept of belief is generally related to propositional claims,<sup>55</sup> which presupposes a complex use of language that may or may not have had evolved prior to these elaborate burial practices, the concept of *intuition* seems less problematic.

### 3.3.1. The intuition of post-mortem continuity

The concept of intuition is conventionally taken to denote something like the direct, unmediated perception of a given truth, a basic insight anteceding any critical assessment of relevant data.<sup>56</sup> Some philosophers (among others, Peter van Inwagen) has instead sought to understand the concept as a form of propositional attitude to separate it from beliefs as such, arguing that an intuition can be understood as a disposition to believe a certain proposition (or class of propositions).<sup>57</sup> Here, we can approach the term along similar lines, and take the concept of intuition to mean something like an inclination or tendency to believe a certain proposition (or class thereof) if it were to be suggested; a basic insight foreshadowing or possibly forming the basis of subsequent propositional beliefs.

Given this, we have two obvious, viable perspectives upon the relation between this *collective intuition* and the burial practices with which it was associated. We can either theorize that these practices was formed around an existing intuition of such a sort, which functioned as the point of gravity around which intentional and subsequent ceremonial burial later crystallized, thereby pregnant with meaning from the very start. Or, we can surmise that the practices themselves spawned this intuition, the inclination toward a

---

<sup>54</sup> Peter Hiscock, *The Archaeology of Ancient Australia*, New York: Routledge 2008, p. 125

<sup>55</sup> Eric Schwitzgebel, "Belief", *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/belief/>, retrieved 2013-04-10

<sup>56</sup> "Intuition", *Random House Unabridged Dictionary*, Random House 2013

<sup>57</sup> Joel Pust, "Intuition", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/intuition/>, retrieved 2013-04-11

certain understanding of the meaningfulness of these communal actions – that human burial perhaps ultimately begins as a desperate expression of tenderness and unbearable loss, a hopeless desire to connect with and care for a deceased loved one, which later on via habituation results in an inclination towards more elaborate notions as to the purpose of burial.<sup>58</sup>

Still, notwithstanding the order in which the intuition and the practice appeared, the idea and the practice were at some point fused together into a complex and multi-faceted tool for meaning-making, arguably one of the earliest cultural *technics*: the ceremonial or ritual burial, today entirely universal in human cultures, and as previously stated, generally considered one of the criteria of behavioural modernity.

With this in mind, we shall put forth a claim which might be considered controversial due to what at first glance perhaps will seem a totalizing tendency, as well as the obvious difficulties of verification. This claim is that this basic intuition, which as have been argued lies at the very roots of human burial practices, thereby has exerted a significant influence upon a vast majority of any and all subsequent belief systems worldwide. This can be understood genealogically, such that certain tendencies or inclinations toward a given category of beliefs were inherited by human cultures which adopted ceremonial burial after interaction with groups who practiced it. The intuition may also be thought of as originating from a basic disposition common to all human animals, such as, perhaps, a proclivity towards a certain emotional experience given a specific set of challenging circumstances, which over time crystallized into more of a noetic tendency towards particular beliefs within a multitude of social contexts.

In spite of its *prima facie* contentious character, it seems that a general theory of this sort is in fact highly plausible. A basic, formative intuition common to human animals would seem a reasonable explanation for the great number of similarities found between disparate systems of religious thought.<sup>59</sup> It need not necessarily take on an essentialist character (even if it may tend towards such conclusions) as it may also possibly be construed as originating in the constraints surrounding our subjective, lived human condition, rather than any stable, identifiable “human nature”.

---

<sup>58</sup> Cf. the discussion on collective habits as a rudiment of the institution in Peter Berger, Thomas Luckmann, *Kunskapssociologi*, Stockholm: Wahlström & Widstrand 2007, pp. 71-81

<sup>59</sup> John Hick’s arguments for theological pluralism put forth in his classic, *An Interpretation of Religion* (New York: Palgrave MacMillan 2004), presupposes something similar in the affirmation of a universal human religious experience, but thereby tends towards a bolder claim than our intuition.

Nonetheless, even if one is skeptical towards the idea of such a formative intuition predating subsequent religious thought, archaeological remains dated to the Upper Paleolithic era (50-10,000 BP) indicate the presence of an increasingly elaborate symbolism in relation to ritual burial practices, which at this point can be clearly associated with explicit conceptual beliefs. The tendencies within these forms of symbolic practice point towards a handful of fundamental ideas, which by all accounts heavily influence subsequent religious behaviour. These ideas, extrapolated from the earliest remains of discernible symbolic thought in the Paleolithic, consist of a preoccupation with reproduction and fertility, totemic representations of animals, hunting, and the cycle of birth and death.<sup>60</sup> Of greater significance to this dissertation, we also see a general holistic approach towards the surrounding cosmos and its inhabitants, implying a reciprocal connectedness – particularly evident in the activities which have been interpreted as magical, wherein symbolic actions are likely to have been thought of as somehow influencing external phenomena.<sup>61</sup> With regard to burial rites and the associated conceptual framework, individual human persons were plausibly thought of as integrated with this holistic, interconnected reality, and are in widely different indigenous cultural locales explicitly understood as either returning to a previous, primordial level of reality beyond physical death; or thought of as transitioning to another stage or level of existence within this all-encompassing web of life.<sup>62</sup>

This view is attested in utterly disparate cultures. It features prominently within Australian aboriginal religion, where the eternal, mythological “dreamtime” to which one returns after death, houses the essence of all living beings.<sup>63</sup> This particular complex of myth has possibly persisted for an extremely long time with relatively few disturbances, as indicated by the fact that the oral traditions of the aborigines carrying these myths, also successfully and accurately have preserved historical data regarding features of their surroundings, features actually absent since the Neolithic era, as well as accurately describing geological events now determined to have taken place

---

<sup>60</sup> Andre Leroi-Gourhan and Annette Michelson, “The Religion of the Caves: Magic or Metaphysics?”, *October* vol. 37, The MIT Press, pp. 6-17

<sup>61</sup> Joseph Campbell, *The Masks of God: Primitive mythology*, New York: Viking Press 1959

<sup>62</sup> These conclusions are given additional support by a recent anthropological study which establishes the fundamental role of animism in primordial proto-religion, concluding a near total prevalence of this position. The authors further argue that belief in an afterlife generally tends to evolve from the more basic animistic position (Hervey C. Peoples, Pavel Duda and Frank W. Marlowe, “Hunter-Gatherers and the Origins of Religion” *Human Nature*, vol. 27, 2016).

<sup>63</sup> Daisy Bates, *Aboriginal Perth: Bibbulmun Biographies and Legends*, Carlisle: Hesperian Press 1992; “the Dreaming” *Encyclopædia Britannica*, *Encyclopædia Britannica Online Academic Edition*, Encyclopædia Britannica Inc. 2013, retrieved from <http://www.britannica.com/EBchecked/topic/171252/the-Dreaming> 2013-04-15

more than 10,000 years ago.<sup>64</sup> This gives us reason to possibly place the roots of aboriginal tradition at no later than the end of the Upper Paleolithic. An adherence to similar holistic views is also evident in indigenous North American cultures, separated from the Australian aboriginal populations by some 20,000 years of history.<sup>65</sup> The *Wakan Tanka*, the divine “great mystery” of the Siouan languages of North America, is a concept somewhat similar to the aboriginal dreamtime, interpenetrating all that exists with a profound, eternal “sacredness” that resides in everything.<sup>66</sup> Ancient Inuit mythology furthermore refers to the *anirniit* spirit world, wherein the eternal spiritual aspects of every living being exist. This worldview presents a three-part ontology of the human person, wherein a “personal spirit” and a “name-soul” is animated by a divine life-force which after death departs skywards, yet leaving the other two aspects to be reborn in relation to a different manifestation of the divine life-force.<sup>67</sup> Basic narratives of a similar kind are also seen among prehistoric Chinese tribes, for whom the ancestral cult was the central cohesive element of society, which during the late Neolithic era evidentially partook in ritual communication with deceased ancestors.<sup>68</sup> Curiously, subsequent Chinese mythology attests a soul structure with certain marked similarities to that of the Inuit.<sup>69</sup> Finally, the belief systems of many indigenous African tribes reflect these basic tendencies as well. The Khoi and San people of south western Africa, with direct lineages to the ancient Sangoan culture of 35-40,000 years ago, who have continuously occupied the originally Sangoan territories since, have preserved traditions regarding an eternal spirit world accessible via altered states of consciousness, wherein the dead reside and from which they continue to influence the living.<sup>70</sup> The

---

<sup>64</sup> Robert Malcolm Ward Dixon, *The Dyirbal Language of North Queensland*, Cambridge: Cambridge University Press 1972, p. 29; *Dyirbal song poetry, the oral literature of an Australian rainforest people*, (with Grace Koch), Brisbane: University of Queensland press 1996; Rainforest Conservation Society of Queensland (RCSQ), *Tropical Rainforests of North Queensland: Their Conservation Significance: A Report to the Australian Heritage Commission by the Rainforest Conservation Society of Queensland*, Special Australian Heritage Publication Series No. 3., Canberra: Australian Government Publishing Service 1986

<sup>65</sup> Alfred J. Andrea (ed.) *World History Encyclopedia*, Santa Barbara: ABC-CLIO 2011, p. 200; p. 98 states that Australia possibly was reached by human migrants as early as 50,000 years ago, while North America was settled around at least 16,500-13,000 years ago.

<sup>66</sup> Raymond J. Demallie, Douglas R. Parks (eds.) *Sioux Indian Religion*, Norman: University of Oklahoma Press 1987, p. 28-29

<sup>67</sup> Tom Lowenstein; Asatchaq (informant); Tukummiq (translator) *The Things That Were Said of Them: Shaman Stories and Oral Histories of the Tikiqaq People*, Oakland: University of California Press 1992, pp. xxxiii-xxxv

<sup>68</sup> David Keightly, “Art, Ancestors and the Origins of Writing in China”, *Representations*, No. 56, Special Issue: The New Erudition 1996, pp. 68-95

<sup>69</sup> Justus Doolittle, *Social Life of the Chinese*, New York: Harper & Brothers 1876, pp. 401-402

<sup>70</sup> W.H.I. Bleek, Lucy Lloyd, John Halstead (ed.) *Specimens of Bushman Folk-Lore*, Abela Publishing 2009, p. 41; p. 204

traditions pertaining to Khoi and San also admit the possibility of reincarnation.<sup>71</sup>

Something could also be said regarding the common nondescript category of what is normally defined as proto-religious beliefs and practices, namely *shamanism*, but due to its ambiguity and the contested nature of the term we will refrain from doing so.

Whether these decidedly relatable perspectives all developed in isolation, or may instead be traceable to a common mythological origin in the form of a full-fledged conceptual structure or narrative, inherited in a more contingent sense, it seems justifiable to root the emergence of these perspectives in some form of the basic intuition suggested above. As previously stated, multiple isolated origins would indicate a common human predisposition towards these forms of thought (not, for instance, unlike the hypotheses formulated by Barrett and others regarding the common global occurrence of theistic beliefs).<sup>72</sup> A single-source theory of cultural adaptation also seems reconcilable with the notion of a common predisposition, but is however at the same time open to a more contingent characterization of this type of intuition and its emergence – single-source theories would grant some plausibility to the suggestion that this type of human behaviour and associated belief structures are nothing but an evolutionary coincidence.

Now, having examined at length the viability of this hypothetical basic intuition without ever really approaching the issue of its actual content; what the inclination or tendency to believe actually moves the understanding towards, so to speak, we will look upon the possible clues divulging information in this respect. It would seem that the most useful indicators of what one may somewhat awkwardly call the “intuition’s cognitive content”, lie in the major themes discernible from such early forms of symbolic practice which directly succeeded the institutionalization of ceremonial burial, that have been very briefly discussed above.

It seems that three broad and general features of our hypothetical base-intuition may be extrapolated from the data pertaining to these practices, which if they perhaps do not form the very basis of them, nevertheless have been profoundly influential in relation to subsequent mythology and metaphysics of human persons, death and afterlife. These features are as follows:

---

<sup>71</sup> Arthur Flaggs Cotterell, *A Dictionary of World Mythology*, Oxford: Oxford University Press 1986, p. 242

<sup>72</sup> See for instance Justin Barrett, *Why Would Anyone Believe in God?*, Lanham: AltaMira Press 2004

*(1) a general holistic approach towards reality, according to which every living thing as well as “inanimate” matter is intimately (and often immediately) interconnected and interdependent.*

This, or similar types of holism is qualified in various ways among early human cultures, generally with reference to some “higher” or more fundamental level of the cosmos, whose inner, hidden workings weave reality together into a unified, coherent structure or entity. This evident in the aboriginal dreamtime-concept as well as native American religion, but is also arguably indicated by the earliest forms of cave art if we give at least some credence to something akin to the sympathetic magic theory which was espoused by the German ethnologist Leo Frobenius.<sup>73</sup>

*(2) a conceptualization of the person as in accordance to (1) interconnected with all of reality – that is to say, the subjective quality of personhood, the individual’s conscious experience of existence, is not considered separate from its counterpart in other sentient beings, nor “material reality”.*

Such a view is evident within all the traditions mentioned, although generally stated in somewhat different terms than those here employed to clarify its implications. The obvious indicator is such practices which via experience intend to relate the individual to this hidden web of basic reality, thus stating that (1) also holds for at least some part of a person’s self-consciousness. The best example is probably the aborigines’ complex philosophical concept of *dreaming*, which among other things denotes a practice relating the individual to the super-reality of primordial, eternal *dreamtime*.<sup>74</sup> Yet, this intentional relating is also fundamental to those mentioned, ostensibly proto-religious practices, which regrettably have been collectively labeled “shamanism”, wherein altered states of consciousness are commonly utilized to access deep reality, underscoring its fundamental entanglement with private consciousness.<sup>75</sup>

(2) is really more or less necessary given (1), but adds the implicit integration of the aspect of subjective experience into the totality of all interconnected aspects of reality of (1).

---

<sup>73</sup> Leo Frobenius, *Kulturgeschichte Afrikas: Prolegomena zu einer historischen gestaltlehre*, Paderborn: Salzwasser-Verlag 2012

<sup>74</sup> William Edward Hanley Stanner, *The Dreaming and Other Essays*, Black Inc. Agenda 2009, pp. 57-59

<sup>75</sup> Mariko Namba Walter, Eva Jane Neumann Friedman (eds.) *Shamanism – An Encyclopedia of World Beliefs, Practices and Culture*, ABC-CLIO 2004, pp. 250-252; David Riches, “Shamanism: The Key to Religion”, *Man*, New Series, Vol. 29, No. 2, 1994, pp. 381-405

*(3) an assumption that individual subjectivity, by virtue of (1) and (2), is to be understood as a possible feature of deep reality (which thereby implies that it does not necessarily cease upon “physical death”).*

(3) also seems to follow given (1) and (2), if (1) is understood as entailing some form of primordial, unifying basic reality of an essentially timeless or eternal nature, whereas this permanence is expanded to include something of individual subjective experience in the first person. This notion is in a rudimentary form arguably already attested by the earliest forms of ceremonial burial, evident in such things as grave goods and the intentional positioning of the body, which may well be intended as concrete preparations for an afterlife. It's a clear and stable feature of all the disparate indigenous traditions we've previously touched upon, as well as of the vast majority of the countless unmentioned, and thus likely to be a truly primitive inclination of human belief.

### 3.3.2. The intuition of post-mortem continuity and the first-person perspective

These features, coming together in (3) above, we term “the intuition of post-mortem continuity”, with an emphasis on (3) as it's more fundamental than the other two in relation to the issue of post-mortem survival. Still, such a distinction may of course be criticized, as the first two components may well be a necessary prerequisite for the assumption of (3), even implying it. Yet, in reality, (1) and (2) are for the purposes of this dissertation only relevant as hypothetical stepping-stones towards (3), so we won't dwell on them to any great extent – as long as (3) or an equivalent notion is produced by any means whatever, the intuition of continuity is present.

If we look closely upon this third feature, the core of our eminent intuition, we find that its chief aspect is the robust presence of subjective experience. The presence of this experience must be regarded not as a transitory illusion, but as something “real”, as an enduring part of the cosmos as such, if the notion that it can persist in the face of such radical change that physical death entails, is to be meaningful and intelligible. It seems plausible that this basic notion (3) can be found at the heart of the vast majority of advanced<sup>76</sup> human conceptualizations of post-mortem survival, due to the conceptual necessity of the existence of *something* carrying this potential of survival, and the need for an anchoring of this something within lived human experience if it is to be relevant and meaningful to actual living human persons.

---

<sup>76</sup> “Advanced” as in formulated as some form of propositional belief-systems rather than the more primitive inclination/intuition discussed previously



There are various options for defining this peculiar phenomenon which we've previously described as "the experience of existence" or in this chapter referred to as just "subjective experience", but one particularly fruitful way of denoting this phenomenon that has been embraced by many is "the first-person perspective". This is due to the latter variant's inclusion of the concept of personhood, emphasizing that this experience of subjective existence is of importance to what it is like to be a human person, as well as the fact that it refers to the point of view from an individual subject's mind, which arguably makes it easier to intuitively grasp than pregnant notions such as "the phenomenal subjective experience of existence".

In all, we shall summarize the contention of this chapter as two separate claims:

1. that a basic, primordial intuition of some form of post-mortem continuity could possibly be regarded as an early human cultural universal intimately connected to the emergence of the institution of ceremonial burial, and that it at the very least seems to have been prevalent among the groups of people who first exited Africa and later spread throughout Asia and beyond, due to the strong indications that it's attested to in many of the oldest available records of religious traditions connected to these populations, as well as arguably indicated by various archaeological findings.
2. that this intuition of continuity, which in various forms has played a significant role in a majority of religious traditions, can be summarized as the inclination or tendency to believe that the subjectivity of a human being is a primitive aspect of reality, and at least in some circumstances and in some sense may persist or return in spite of physical death.

## 4. Death and post-mortem survival

### 4.1. Introduction

Death, as perhaps the ultimate paradox, is an inevitable fact of life. It's generally considered a universal aspect of the human condition, as well as that of the community of life as such,<sup>77</sup> although humans are often assumed to possess a privileged perspective on the fact of mortality due to our comparably advanced sentience. The cessation of an individual's life has throughout history been defined according to various criteria. While the ancient accounts often focused upon the person's spirit evacuating the body, e.g. the departure of the *ka* as in Egyptian tradition,<sup>78</sup> or the event of the *psyche* leaving the corpse related by Homer,<sup>79</sup> death has in modern times generally been understood in a strictly biological sense. Traditionally, the onset of death according to modern accounts was considered the moment when cardiac circulation had ceased, yet contemporary clinical definitions of death mainly depend on the discontinuance of brain activity, chiefly due to the fact that cardiac arrest today more often than not is a reversible condition.<sup>80</sup>

#### 4.1.1. Definition of death and post-mortem survival

For clarity's sake, the notion of post-mortem continuity of experience needs to assume a thoroughly unambiguous definition of death. Since the purpose is to address the question of whether continuity of experience is conceivable, even after the utter destruction of the body and a complete scattering of its

---

<sup>77</sup> A partial exception could possibly be made with regard to such species whose members do not actually face an increased risk of mortality with an advanced chronological age, somewhat misleadingly being defined as "biologically immortal". Still, the life span of such creatures, while not being absolutely determined by senescence, is effectively limited by other environmental factors. Cf. Daniel E. Martinez, "Mortality patterns suggest lack of senescence in Hydra", *Experimental Gerontology* 33, 1998, pp. 217-225; J. W. Shay, W. E. Wright, "Hayflick, his limit, and cellular ageing", *Nature Reviews Molecular Cell Biology* 1, 2000, pp. 72-76

<sup>78</sup> Andrew H. Gordon, Calvin W. Schwabe, *The Quick and the Dead: Biomedical Theory in Ancient Egypt*, Brill/Styx, Leiden 2004, p. 18

<sup>79</sup> "death", in Nigel Wilson (ed.) *Encyclopedia of Ancient Greece*, Routledge, New York 2010; cf. Homer, *Iliad*, 23.54 ff

<sup>80</sup> Rajesh Bardale, *Principles of Forensic Medicine and Toxicology*, Jaypee Brothers Medical Publishers Ltd., New Delhi 2011, pp.132-133

material components, it would be misleading to employ possibly, or at least conceivably reversible corporeal states, such as the cessation of measurable electrical activity in the brain, as a basis for our definition. Rather, death will be taken to refer to *a state wherein a person's body is thoroughly disintegrated into separate parts of what is normally considered inanimate matter*. This understanding correlates well with the tentative definition which was presented in an earlier chapter, where corporeal death was considered equivalent to the complete disintegration of the human body in terms of a functioning biological system.

Such a definition will however possibly need to be clarified further. What level of disintegration and separation will be necessary to meet these criteria, for example? One may further argue that even maximal separation of the basic constituent parts of a human body, for all we know, isn't necessarily a completely irreversible physical state of the organism. A possible response to such remarks might be that any level of disintegration where the body is reduced to the basic chemical substrates of biological tissue is indeed sufficient for a claim that permanent physical death has occurred. This in turn does also imply that a successful reversal of such a state of complete decay, accompanied with the resurgent presence of the same subject as existed prior to the disintegration, would equate to an instance (one of several conceivable) of post-mortem continuity.

This understanding of death is a good bit stronger than other contemporary analyses. As an example, Cody Gilmore's definition of death in *The Oxford Handbook of Philosophy of Death* approaches it as a state of an organism, a "thing", the necessary and sufficient conditions of which are satisfied when said thing is dead and has since then not regained the capacity to live.<sup>81</sup> Gilmore more or less eschews the conceptual distinction between the presence or absence of a subject in a phenomenal sense, and the living or dead status of some organic entity. This position is based in an implied monist physicalist ontology, and underwrites Gilmore's analysis of "postmortem revitalization" in terms of exclusively physical and technological possibilities. This notwithstanding, Gilmore emphasizes that permanent irreversibility of physical death for various reasons is problematic as a necessary criterion of death, and is open to the position that an organism in suspended animation or cryostasis, as sci-fi terms it, is actually dead. We, on the other hand, must employ a stronger definition of death if we are to accommodate the traditional notion of an afterlife, which conventionally involves the problem of whether or not a person can survive the permanent and ostensibly irreversi-

---

<sup>81</sup> Cody Gilmore, "When Do Things Die?" in Ben Bradley, Fred Feldman, Jens Johansson (eds.), *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012

ble destruction of their body as a functioning physical organism. An interesting parallel to the definition we have suggested above is the concept of *final death* in contemporary vampire fiction and role-playing games. Final death is here equivalent to the irreversible destruction of a person's body, which then is clearly distinguished from both the process in which a person's body is killed, i.e. that whereby he or she attains *undeath*, and the utter, complete annihilation of the person and the phenomenal subject as such.<sup>82</sup> As an aside, questions pertaining to the actual annihilation of the phenomenal subject as such will not be relatable to our definition of death, as the dissertation's focus is upon survival beyond bodily death rather than the metaphysical viability of annihilationism or equivalent ideas.

It would nonetheless seem that our definition of death of a human person above, as *a state wherein a person's body is thoroughly disintegrated into separate parts of what is normally considered inanimate matter*, is sufficiently straightforward to circumscribe the hypothetical state of affairs this dissertation is addressing, namely the continuity of a subject beyond ostensibly irreversible corporeal death, while excluding instances of revival of persons where the threshold into actual physical death can be argued to never have been crossed, however close it may have been.

## 4.2. Legacy problems of death and the afterlife

With our explanation of the concept of death in place, we can now tentatively address the issues traditionally associated with the notion of an afterlife, or with our preferred terminology, a post-mortem existence of the subject which was present prior to corporeal death. For a positive response to the question of whether such post-mortem survival is at all possible, two basic things are needed, as has previously been discussed. We must provide a coherent model of the subject that without contradiction can be taken to persist or to be reconstituted in spite of the destruction of its particular body. We also need a basic metaphysical schema in place, according to which such a subject model is conceivable.

These issues traditionally associated with the notion of an afterlife, or the legacy problems of death, are certain difficulties that arise with regard to the provision of these two fundamentals. The legacy problems can be grouped under the problem(s) of personal identity; the related sameness or individuation problem; and the fundamental question of which metaphysical positions

---

<sup>82</sup> Justin Achilli et al, *Vampire – The Masquerade*, 20th anniversary ed., White Wolf Publishing 2011 p. 33, 161

are compatible with, or favorable with regard to the notion of an afterlife.<sup>83</sup> With regard to the latter, one may speak of a background metaphysics and an anthropological metaphysics, i.e. if the person is thought of along reductive materialist terms but the assumed global metaphysics is traditionally theistic, the background metaphysics would be theistic and non-materialistic, while the anthropology is materialistic.

We will here refrain from at length discussing problems which are mainly of an axiological relevance, such as the question of whether a post-mortem existence is at all desirable, due to our focus which is basically metaphysical.

The problems will first be discussed in relation to our suggested response, the phenomenological criterion of personal identity, which will be presented in connection with the problems. Following this, other types of responses to the problems will be discussed separately, in terms of whether the responses can produce criteria for survival which can overcome the legacy problems.

#### 4.2.1. The problems of personal identity

The problems associated with personal identity are essential to the discussion, and among the first to emerge in any philosophical exploration of death and any type of afterlife. Moreover, they're particularly pressing with regard to the widespread eschatological notion of the resurrection of the dead. We will here briefly discuss the most prominent of the problems of personal identity, and the manner in which they may conceivably be addressed utilizing what we call the phenomenological criterion of personal identity, from which our suggested ontology of the (human) self which is then developed and argued to be coherent in chapter 5. The necessary basics of the self-model are thus established in this chapter, in relation to the particular problems they're intended to address, while these necessary basics are then portrayed in terms of a possible model of the person in the following chapter.

#### 4.2.2. The synchronic and diachronic problem in view of the phenomenological criterion

The more basic aspects of the problem have often been labeled the *synchronic problem of personal identity*, i.e. questions regarding how a person can persist in a particular moment. The synchronic problem addresses issues of defining personhood, what attributes are necessary or sufficient for something to count as a person, or what a person ontologically speaking consists

---

<sup>83</sup> Baker 1997

of. The *diachronic problem*, on the other hand, regards how, if at all possible, a person may persist over time, from one moment to the next.<sup>84</sup>

As stated, our suggested response to the synchronic problem will be based upon the conceptualization of the human self, presented in more detail in chapter 5. This position, which is of a phenomenological character, essentially entails that the subjective, phenomenal first-person presence, or the first-person perspective, is an entity, substance, mode or attribute which is irreducible to objective material reality, and therefore arguably ontologically primitive and to some extent independent.<sup>85</sup> Given this conceptualization, the synchronic persistence of a given self will be equated to the *internal subsistence of a first-person perspective from the point of view of a certain subject*. This is to say that person X persists synchronically, or in the moment, iff she actually has a first-person perspective, or in the somewhat more occult phrasing of Nagel and Sprigge, iff *there is something which it is like to be* person X at the particular moment in question.<sup>86</sup> This criterion builds on the understanding of survival as necessarily involving subjectivity, as was discussed in the beginning, and simply entails that whatever survival actually means, phenomenal subjectivity is a fundamental ingredient.

Thus, the more or less ineffable persistence of one's own subjective phenomenal experience is taken to function as the basic criterion for oneself to exist as a person in the present moment. By extension, the response to the problem of diachronic persistence from the same position is to simply state that if the subjective, qualitative phenomenal first-person experience of person X happens to subsist in different moments, said person (or the self-aspect of the composite person, in case one prefers such an ontology) will also be taken to have persisted over time. For reasons of simplicity, we will call this the *phenomenological criterion of personal identity*.

This criterion is accompanied by two obvious caveats. First and foremost, the phenomenological criterion is an internal criterion. It is ostensibly useless as an epistemic tool for attaining knowledge regarding other persons and their supposed synchronic or diachronic persistence (if we exclude the interesting possibility of the direct experience of others' first-person perspectives,

---

<sup>84</sup> Harold Noonan, *Personal Identity*, New York: Routledge 2003, p. 84-86

<sup>85</sup> Similar positions are held by many idealist and dualist philosophers, yet the particular version espoused in this dissertation which is presented in chapter 3 is mainly indebted to Klawonn's and Baker's respective work on the topic, cf. Erich Klawonn, *Mind and Death: A Metaphysical Investigation*, Odense: University Press of Southern Denmark 2009; Lynne R. Baker, *Naturalism and the First-Person Perspective*, Oxford: Oxford University Press 2013

<sup>86</sup> Timothy Sprigge, Leemon B. McHenry (ed.) *The Importance of Subjectivity*, New York: Oxford University Press 2011 p. 41

the merging or interaction of such perspectives or similar phenomena), and only functions as a metaphysical criterion, stating what *makes it the case* that a certain person really exists in the moment or persists over time, disregarding whether this is epistemically accessible from the point of view of external subjects. Moreover, it implies that unbroken temporal continuity of the first-person perspective is not necessary for the identity of person A at time X and person B at time Y as such, who may nonetheless persist diachronically. So, for instance, if we when referring to the Norse explorer Erik the Red (*Eiríkr hinn rauði*) identify a certain person who existed sometime in the late 10<sup>th</sup> century, he may as well be identical with some future person existing at a given moment iff this future person exists as such *due to the presence of the very same internal first-person perspective*. Also, this is independent of whether the particular first-person perspective is instantiated in the time span separating the two occasions or not (it could conceivably lie dormant).

The necessary status of “particularity” of a certain first-person perspective can be expressed in a number of ways, but is basically a way to describe the fact that your subjective phenomenal reality is irreducibly your own, or primarily given with regard to your own consciousness, and thereby in a fundamental sense private. This ownness or mineness of the subjective phenomenal reality may also be considered a haecceity of the particular first-person perspective, with the concept of haecceity denoting a “thisness” with regard to particular entities.

There need not be any spatial or temporal continuity between the two instances of Erik the Red, just like Pink Floyd’s “Wish You Were Here” would actually be the same song, even if the song was not continuously instantiated in any sense from now until a thousand years into the future, when it for some reason then reemerged, exactly the same down to every note and syllable.

At this point, it could be countered that the causal origins of an entity factors into its identity conditions, so that the second instance of the Pink Floyd song cannot be identical with the first if they’re causally unrelated. They’re just very similar. This argument, however, entails the denial of any sense of the song’s existence as an abstract object, implying the assumption of some form of nominalism, or at least conceptualism. Notwithstanding all of its other difficulties, nominalism entails such problematic conclusions as that the Babylonian concept of multiplication is actually different from the Aztec concept of multiplication due to their dissimilar origins, even though they’re completely indistinguishable in propositional terms. Conceptualism faces similar issues.

However, even if we completely disregard the discussion of the metaphysics of abstract objects, this kind of critique can nonetheless not be employed with regard to the phenomenal criterion, for two reasons. Firstly, the phenomenal criterion turns on the presence of a particular phenomenal, non-abstract, subjective mineness as the sole grounds of identity of the persevering subject. Accordingly, iff this particular mineness is present, the subject in question is taken to exist. If it's not present, the subject does not exist, and no other subject can be indistinguishable from it without actually possessing the very same mineness, which then would entail it actually being identical with the subject in question.

Furthermore, the conceptually necessary indivisibility and induplicability of this mineness, which will be discussed in detail in the following chapter, implies that there can only be one such mineness in existence at a particular given time. All that any conceivable division of mineness can possibly accomplish is the emergence of two separate subjects, or a bicameral situation with a unifying singular meta-subject.

If we assume that two instances of qualitatively identical, phenomenal, subjective mineness were actually not identical (e.g. due to particular causal relations), they should obviously have been able to co-exist at a particular time, yet this is impossible due to the indivisibility of the very quality of phenomenal mineness. Thus, two instances of the same mineness cannot be distinguished with regard to causal relations – if they can at all be distinguished, they're not the same subjective mineness, and two coexisting, separate instances of the very same subjective mineness are inconceivable due to its absolute indivisibility.

Additionally, if ostensibly separable instances of the phenomenally identical and unique subjective mineness or ownness are considered to be properties rather than thing-like substances as such, they can neither be effectively distinguished. Approaching subjective mineness a property or an attribute in the way Lynne Baker does in her recent work *Naturalism and the First-Person Perspective* wherein she argues for the irreducibility of the phenomenal self, subjective mineness can simply be considered a uniquely instantiated attribute. Examples of such would be the world's tallest man or the fastest man alive, yet the mineness as an attribute would be one which is essentially phenomenally subjective, and not thing-like – a phenomenal thisness of an immaterial substance, for instance, or a phenomenal thisness-attribute emergent upon a material substrate. In accordance with this position, then, whenever the unique attribute which is identical with the subjectivity of Erik the Red



happens to be instantiated, the person to whom this name refers is also taken to exist.<sup>87</sup>

The diachronic problem can, as previously stated, be addressed in an equivalent fashion. Subjective persistence of the first-person perspective from one moment to the next will on our favoured conceptualization of the human self, function as the effective criterion for diachronic personal identity.

To sum up, the criteria for synchronic personal identity is thus satisfied, paraphrasing Nagel and Sprigge, iff *there is something which it is like to be* the subject in question at a particular given moment. The criteria for diachronic personal identity is similarly satisfied iff the *internal subsistence of a first-person perspective from the point of view of a certain subject* persists in one moment as well as in a subsequent one. Regarding the metaphysical implications of the notion of a post-mortem experiential continuity, we can at this point simply acknowledge that the conceptual viability of this notion presupposes an adequate response to the twin problems of diachronic and synchronic continuity. Simply put, if a person cannot even persist synchronically within the moment, the possibility of a post-mortem continuity seems rather slim. And likewise, if a person cannot exist diachronically, from one moment to the next, post-mortem continuity is also undermined. However, it seems that a response to the synchronic problem is the most essential, since the possibility of a person's synchronic existence might on its own imply the possibility of some form of death-transcending continuity, even if we exclude the particular form of diachronic persistence within time.

Non-phenomenalist responses to the synchronic and diachronic problems of personal identity will either need to employ the criterion of spatio-temporal continuity, or possibly make use of some form of metaphysical principle to the effect that each and every person is substantially unique, e.g. assume that every human person for some reason is the only possible member of his or her unique *species*. Spatio-temporal continuity as a criterion for personal identity over time is, however, beset by a certain set of problems that the phenomenological criterion does not face, which is addressed in some detail in relation to the individuation problem. Criteria appealing to the unique species-nature of every person are neither without issues. They will only be plausible if explicitly formulated within some form of immaterialist or phenomenalist framework, and will even in that circumstance necessitate a further explanation (such as divine intention or laws of nature) to ground such essential, individuating differences.

---

<sup>87</sup> Baker 2013 pp. 32-34

### 4.2.3. The problem of individuation

The problem of individuation basically regards how to distinguish separate persons or subjects without reference to particular bodies, which is necessarily actualized by the notion of a post-mortem continuity implying a person's continued existence past the destruction of her physical body. This is to say, that if person A persists diachronically during two separate moments in between which the physical death of his or her body occurred, how can person A be individuated or distinguished from other persons?<sup>88</sup>

The problem of individuation is generally expressed in relation to the assumption of a non-reductionist, non-physicalist ontology of mind, i.e. if that which is considered to survive is a person- or subject-entity with no spatial designation. Arguably, any notion of the post-mortem existence of persons requires a response to the problem from such a non-reductionist viewpoint, since there needs to be an actual diachronic continuity (although not necessarily temporally unbroken) or equivalent preservation of the subject during a period when it's obviously disembodied. In other words, since the body of Erik the Red ceased to exist at a certain point, his continued non-spatial and disembodied existence needs to be individuated as his own rather than that of another person, if his particular personhood is to be taken to persist (whether or not he is embodied anew later on).

Such an individuation is generally taken to be extremely difficult to perform, since the person in question ostensibly lacks a spatial designation, and as a distinction merely with reference to mental content at least in principle could fail due to the assumed possibility of such content being identical.<sup>89</sup> To illustrate the problem, consider the aforementioned individual, once a prominent Viking explorer, yet now reduced to the disembodied person X in a hypothetical afterlife populated only by incorporeal entities. Let us furthermore assume that person X harbours no memories whatsoever of his previous embodied existence, and thus is indistinguishable from other such entities on the grounds of particular experienced mental content. In this situation, how is person X to be individuated? How shall we go about in distinguishing this person from other disembodied entities of a similar kind, with recourse neither to any spatial location nor mental content?

A response along lines similar to our treatment of the synchronic and diachronic problems (anchored in our phenomenologically informed understanding of the human person), will also enable our response to the problem

---

<sup>88</sup> Baker 1997

<sup>89</sup> William Lycan, "Is property dualism better off than substance dualism?", *Philosophical Studies*, 164:533-542, 2013

of individuation of human persons in such a subjective or non-spatialized sphere. The incorporeal, non-spatialized person X, lacking memories and having no unique mental content, is accordingly ontologically distinguished from the spatially and physically indistinguishable person Y due to the basic, primitive fact that *there is something which it is like to be person X* which is *not equivalent to what it is like to be person Y*. In other words, person X and person Y, respectively, denotes particular subjects, which can be individuated by the fact that both has a unique and ineffable phenomenal first-person presence or *mineness* which is not reducible to that of the other. This state of affairs, verbally somewhat elusive, can be effectively illustrated with several thought experiments, one of the most poignant of which is the story of the failed teleportation. In this thought experiment, we imagine a situation wherein a technology of teleportation exists and is implemented, which allows persons or objects to be almost instantly (although perhaps not faster than the speed of light) transferred from point A to point B.<sup>90</sup> This teleportation is in practice actualized by the disintegration of the object at point A, and the reconstitution of an objectively identical object at point B. Notwithstanding the issues regarding whether such a process could actually allow oneself as a conscious subject (as essentially an embodied first-person perspective) to reappear at the other end, we shall assume that this is reliably possible. We then ask you, the reader, a veteran traveler by teleportation, to assume that you are to embark on a journey from Tokyo to the international colony on Mars. In the teleportation process, however, some sort of error occurs, and you find yourself still at the platform in Tokyo. However, the reconstitution seems to have taken place as planned, as there's a perfect physical copy of yourself at the receiving end on Mars, which unfortunately is claiming your identity, with a set of memories seemingly identical to your own in support for this. So the problem is how you are to meaningfully distinguish between yourself and this bothersome tele-clone.

From one's own perspective, disregarding the issue of authenticity (which of the two is to be considered the original), this can hardly be more straightforward. The person in Tokyo is you, the one at the other end is not. Whatever happens to the person at the platform in Tokyo is immediately and inescapably present to yourself, while this does not hold with regard the person at the other end, from your own point of view. Likewise, were the person in Tokyo to be annihilated together with your own unique first-person presence, *you* would not continue to exist merely due to the fact that a person with an iden-

---

<sup>90</sup> This particular type of the thought experiment is to my knowledge first explored by Erich Klawonn in his *Mind and Death*, regarding a related matter

tical body and identical memories was preserved on Mars.<sup>91</sup> Thus, you can be metaphysically individuated from the clone with regard to your own, unique first-person presence which happens to be associated with your body rather than the one on Mars.

Importantly, the phenomenological criterion connects with the metaphysical fact of one's subjective, first-person presence and does one again not effectively function as an epistemic criterion. Hence, it does not necessarily allow a limited observer to distinguish between particular external subjects none of which are him- or herself, but instead establishes the metaphysical fact of their distinctiveness which lies in their unique phenomenal selves. Therefore, it can be employed as a response to the *metaphysical* problem of individuation.

#### **4.2.3.1. The phenomenological criterion and spatio-temporal individuation**

Conventionally, spatio-temporal continuity has been put forth as grounds for the identity of entities over time.<sup>92</sup> This position is beset by certain difficulties, such as problems of definitional circularity and conceptual incompleteness which do not as such plague the phenomenological criterion.<sup>93</sup> Moreover, the phenomenological criterion operates even in the absence of conditions of spatio-temporal continuity – that is, the first-person presence effectively functions as grounds for individuation even when we cannot distinguish two or more persons with recourse to the spatial locations of their bodies. Even under normal circumstances, however, spatio-temporal designation is a limited tool for the effective individuation of persons.

This follows on the assumption of either of the two basic metaphysical positions on the nature of space. Assuming some form of *relationalism*, which essentially is the view that space is constituted by the relations between particular objects (such as the distance between objects x & y), they seem to be impossible to individuate in certain situations.<sup>94</sup>

---

<sup>91</sup> Arguably, mental content in other than a completely propositional sense is coloured by the same haecceity as the particular first-person perspective, and is therefore equally unique (by virtue of the ineffable phenomenal first-person presence).

<sup>92</sup> Cf. David S. Oderberg, *The Metaphysics of Identity Over Time*, London: Macmillan Press 1993

<sup>93</sup> Ibid. p. 2-15

<sup>94</sup> Cf. Oliver Pooley, "Substantialist and Relationalist Approaches to Spacetime", in Robert Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press 2013

For example, if we put our two clones above in the situation of an otherwise completely empty space when the only point of reference is the body of the other, it becomes obvious that spatial designation alone ultimately cannot support a meaningful individuation. To be precise, this situation can be visualized as an infinite space populated with nothing whatsoever except the two identical bodies in question. With no objective frame of spatial reference other than the bodies themselves, the positions of these poles are identical (either is placed at distance  $X$  from the only other object of reference). A similar, if somewhat narrower counter-example can also be constructed in relation to an empty finite space, where the edges of the finite space function as a frame of reference common to the two clones. In this case, the same result is obtained as long as the objects are at the same distance from each other, as well as the edges of said space. A common response to these difficulties is to invoke the notion of absolute space.

Assuming e.g. *substantivalism* or other forms of absolutism, which entail that space is an independent entity capable of functioning as a framework within which particular entities can be objectively positioned, brings other significant problems.<sup>95</sup> The foremost of these seem to be the risk of a vicious regress of reference due to the problem of individuating particular spatial regions, but it's neither entirely clear that absolutism can actually provide effective individuation. On the position of absolutism, even if objectively existing regions of space could serve as a basis for the individuation of particular objects therein, it seems that these regions of space as such, cannot straightforwardly be distinguished in the same manner. Obviously, the regions themselves cannot be distinguished with regard to the physical bodies occupying them without begging the question, since these bodies are supposedly individuated by virtue of their spatial locations.

There are then two basic options. Either one assumes that the position of each region of space can be individuated with regard to another such spatial region. This would imply the necessity of postulating a vicious, infinite regress of meta-spaces, which compounds rather than solves the problem, since no identity statement would ever actually be complete. Another option would be to postulate that space is finite, and that sub-regions thereof could be objectively identified with regard to the hypothetical boundaries of the total space. However, then the same problem of individuation as in the previous example occurs, in that one can without contradiction posit spatial regions with identical individuating relations to the hypothetical boundaries. Since the boundaries cannot themselves be directionally individuated in turn, more than one spatial region can possess the same individuating attributes,

---

<sup>95</sup> Dean Rickles, *Symmetry, Structure and Spacetime*, Amsterdam: Elsevier 2008, p. 73-75

i.e. both objects can be at distance X from one “side” of the spatial region, and distance Y from the other, yet the two sides cannot be further distinguished since there are no directional referents. Thus, even if we were to concede the objective existence of spatial regions as a framework for individuating reference with regard to the objects occupying them, the identification of the spatial regions as such are subject to similar problems of individuation as the identification of our ideal objects on a relativistic spatial framework.

Spatial designation thereby cannot by itself effectively and consistently individuate even non-conscious physical bodies, devoid of intentionality, and with regard to persons, it seems to leave out the arguably most important aspect, namely the subjective sphere.

In summary, there seems to be no major obstacles in successfully applying the phenomenological criterion of individuation to the situation of the disembodied persons. It’s a satisfying response to the problem of the failed teleportation, where memory, physical constitution and mental content cannot be relied upon, and there’s no obvious reason it cannot work in a situation where no reference to spatial designation can at all be made.

### 4.3. Basic possible responses to the legacy problems from the major metaphysical positions

In this section, we will explore some of the most common possible responses to the problems discussed above, advanced from the viewpoints of the general philosophical anthropologies available. In regard to these, there can be distinguished four major basic metaphysical positions relevant to how we conceive of the human person, each, however, with a number of sub-categories. Having discussed these metaphysical positions, it will be discussed which basic strategies of addressing the legacy problems are available given the respective positions.

First of all, there are the various interpretations of *reductionist physicalism*, which essentially claim that human persons are material objects, however we more exactly understand the metaphysics of the latter. *Dualism* of several types has been a staple of Western metaphysics since the Ancient era, today perhaps most common in the variants of *substance dualism*, *property dualism* (generally considered a non-reductive variant of physicalism), *hylomorphic dualism*, as well as certain other forms of non-reductive materialism that also sort under dualism. *Idealism* comes in several flavours as

well, and finally we have the position of *neutral monism*, according to which only one basic substance, neither mental nor physical, exists.

#### 4.3.1. Reductionist physicalism: mnemonic and bodily continuity

Reductionist physicalism, assuming that every fundamental property is physical, places obvious restrictions upon possible responses to the synchronic and diachronic problems. From the point of view of reductionist physicalism, criteria for synchronic persistence or diachronic continuity of the person will exclusively refer to memory, to mental content, to actual physical continuity, or some combination thereof.<sup>96</sup> Since reductionist physicalism equates mental events with the activity of the physical organism, the human person is on this position commonly likened to a body incorporating an extremely complex computer. On this account, consciousness is generally identified with neural processes which are taken to be physical in nature.

The criterion of physical continuity essentially entails that for a person to exist in the present moment, they need to be a living body, and for them to persist diachronically, there needs to be a concrete physical continuity between the organism which was them in the previous moment, and the organism they supposedly are in the next.

The memory criterion is on the surface a relatively straightforward concept, intended to identify a person by referring to the particular set of interconnected information, physically present in his or her brain (or an analogue thereof) which makes up an individual's memories of his or her past. Thus, the synchronic problem can be addressed with the aid of the memory criterion if we merely equate the presence of this set of interconnected information as an aspect of a particular living physical body, with the synchronic subsistence of a human person.

The diachronic problem can be approached similarly, granted that the persistence of a human person over time is the same thing as the presence of this unique information-history as an aspect of a living body, from one moment to the next. It should also be noted that the memory criterion is on the reductionist's account functionally identical to the criterion of physical continuity, since mental states are reducible to physical states.

---

<sup>96</sup> Bruce Reichenbach, "Monism and the Possibility of Life After Death", *Religious Studies*, Vol. 14, No. 1 (Mar., 1978), pp. 27-34

Nevertheless, utilizing the notion of a unique pattern of neurally encoded information in a response to the diachronic problem faces several difficulties.

First of all, our memories aren't static over time, but change in response to the environment, to our choices, experiences and thoughts, which could be taken to imply the impossibility of diachronic persistence given this criterion. Still, it's possible to employ the notion of a unique yet changing set of information in structuring such a criterion which would evade this difficulty, unique by virtue of its particular history rather than its composition. This could be likened to considering our persons a pile of Tetris blocks in an ongoing game, which however changing, nonetheless can be said to retain a particular identity with regard to a unique historical continuity. Still, this type of uniqueness with regard to history nonetheless seems to provide rather weak support for the notion that a person can persist diachronically. For instance, there seems to be no matter of fact regarding person X at time T1 which is necessarily retained by person X at time T2, other than a rather vague continuity of informational history. Lacking such matters of fact, what makes it necessarily the case that person X is still the same person, rather than the new person Y? Moreover, an equivalent continuity of informational history could in principle be shared by other entities besides person X at time T2, making them possible contenders to be identified as having been person X at time T1 (e.g. in relation to counter-examples pertaining to tele-cloning or replication via cell-division).

Instead of the perhaps somewhat cumbersome "continuity of informational history", Lynne Baker suggests a stronger definition of the memory criterion, and presents it as a *particular type of causal connections between mental states*.<sup>97</sup> Here a somewhat arbitrary "particular type" is added, to remove other possible contenders of having been person X at time T1. If we for instance consider my memory of feeding my cat in May 2003 having been related to my friend Francis twelve years later, causing in his imaginative self a relevantly similar mental state, this would then be a faulty type of causal connection. Also, if I after having told Francis this story came to suffer from retrograde amnesia, forgot the event, yet were later reminded of it by him, and thus had access to a relevantly similar mental state, this connection would also seemingly be unacceptable, since the criteria for me being myself would then for a period of time have been inaccessible to me, yet carried by another person.

---

<sup>97</sup> Baker 1997



Another significant difficulty with the memory criterion is related to the thought experiment regarding the failed teleportation we've previously discussed. However we understand the admissible types of causal connection between the mental states of the memory criterion, it's nevertheless logically possible that two identical such sets of mental states can be produced "in the right way", via the acceptable types of causal connections. An example of such a situation would arguably be the unintended duplication of the failed teleportation above, another would be the hypothetical amoeba-like division of a human person where the same memories were retained by each resulting individual. In effect, it's possible that the memories of Erik the Red are transferred via acceptable causal connections to two future individuals, person X and person Y, which on the assumption of the memory criterion would imply that Erik is simultaneously both person X and person Y. This is contradictory on the assumption that personal identity is a unique one-to-one relationship, which has the consequence that two different, separate persons simultaneously existing at time T1 cannot be one single person, identical with itself. Thus, the exclusive use of the memory criterion implies the contradiction that our Erik the Red would be identical to X as well as Y, while X and Y at the same time are non-identical.<sup>98</sup>

The responses to (analogues of) the problem of individuation available to the physicalist can be criticized for similar reasons. Obviously, the situation of the disembodied, non-spatialized persons described above never arises on a physicalist ontology, however, situations equivalent to the duplication of the failed teleportation are logically possible. Assuming this type of duplication, wherein two physically identical individuals are produced, both sharing the same mental states pertaining to memory, they can at this point in time only be individuated with regard to spatial designation, which is problematic for the reasons stated earlier. The only way around this, utilizing the memory criterion, would be to acknowledge some form of non-reducible haecceity of particular mental content, or sets thereof, unrelated to their accidental material constitution. This would however seem to violate the assumption of reductive physicalism.

Finally, two rather significant problems with the physicalist position seem to arise from related considerations. Given the immediate fact of subjective realities, of the phenomenal first-person presence and its relation to animals and the human individual, two issues emerge on the assumption that the person is reducible to its physical constituents, whether this is expressed with the above criteria or not.

---

<sup>98</sup> Bernard Williams, "Bodily Continuity and Personal Identity", in *Problems of the Self*, Cambridge: Cambridge University Press 1973, p. 19-25

Assuming physicalism, it seems difficult to explain why one's own phenomenal first-person presence in particular is associated with the specific body that one seems to inhabit, rather than another. There seems to be no objective state of affairs on the purely material level which implies that *my* subjective experiences, the mineness of my first-person presence, must be related to this particular body, rather than that of my 88-year old neighbour in his apartment a few feet away.

Secondly, given the physicalist account, it would in principle be possible to create any number of identical copies of oneself, each generating *one's very own* unique phenomenal presence. For instance, the copies resulting from the failed teleportation would subjectively speaking both be yourself at the same time, which aside from being a decidedly counter-intuitive notion, indirectly runs into a number of problems. The most important of these seems to be the fundamental unity of the first person: either both clones would have to function as something like separate sensory nodes of your own subjectivity, as if you alone were watching a CCTV feed where two cameras projected images onto a split screen, or both of them will function as separate, independent subjects (no matter how intimately connected they may be) precisely as the assumed outcome of the failed teleportation thought experiment.

The split-screen alternative would then imply that your meta-subjectivity isn't reducible to any one particular body in existence, whereas the only other available option, assuming the two independent subjects of the teleportation example, similarly implies that you and your phenomenal first-person presence isn't identical to any particular physical body.

#### 4.3.2. Dualism, idealism and non-reductive materialism: immaterial continuity, composite continuity and phenomenal presence

The immaterialist-type ontologies can with no difficulty utilize the criteria of mnemonic or physical continuity discussed above, yet are also capable of accommodating at least three basic responses to the legacy problems of an entirely different character. Most obviously, we have the criterion of continuity of the immaterial mind or the soul, but also, the criterion of continuity of the composite entity of the immaterial and physical aspects of the person, and of course, the phenomenological criterion of personal identity which we've previously discussed.

##### 4.3.2.1. Immaterial continuity

The notion of a concrete continuity of the immaterial part of the person is conceptually similar to the criteria of mnemonic or bodily continuity, and

simply entails that something like an immaterial soul-substance (or another form of immaterial substrate, such as an immaterial property), independent of the physical body, is what fundamentally constitutes the human person. As a response to the synchronic problem, this position would contend that I exist as a particular person in a given moment, if and only if the immaterial foundations of my own self are “present”, active, or in existence. I am a person here and now, for example by virtue of the existence of my objective immaterial substance.

A response to the diachronic problem on the assumption of this position, would state that for it to be the case that I existed as a person at time T1, as well as subsequently at time T2, the very same immaterial substrate at the very least needs to be present, in existence, or active on both occasions.

The criterion of immaterial continuity avoids some of the issues which plague the materialist criteria, yet faces certain problems of its own. With regard to its advantages, it doesn't seem to run into something equivalent to the mnemonic criterion's issue of the impermanent nature of memory, as it seems perfectly possible for the constituent immaterial substance to remain essentially unchanged, unscathed by the passage of time and physical processes. But is the criterion of immaterial continuity also safe from the counter-argument based on the hypothetical duplication of the person-bearing substance? Obviously, since the body isn't the substrate of the person on this account, a physical duplication of the person's body isn't problematic in the same sense. If the body of our old friend Erik happens to be copied by an event similar to the teleportation accident, it's not obvious that two identical sets of person-bearing substances has been produced, which implies that the criterion of immaterial continuity is at least not rendered contradictory by this particular counter-example alone.

However, a duplication of the immaterial person-bearing substance in question is arguably still logically possible, at least on certain accounts of said immaterial entity. If we were to assume that this non-physical substrate of the human person could possibly be copied in a manner similar to that of the physical body, we suddenly find ourselves facing the more severe form of the problem of individuation mentioned previously. That is, if the non-physical, essential substrate of Erik the Red happens to be duplicated, how do we individuate his own person, with recourse to neither his physical body nor any spatial designation?

There seems to be two possible, related responses based on the criterion of the continuity of an immaterial substance.

An irreducible haecceity or *thisness* of the unique immaterial substrate of each person might be assumed to evade the problem (i.e. each soul is ineffably and irreplicably unique, whereas a duplication of the immaterial substance, if at all possible, simply would result in two distinct persons). However, the most obvious ways of anchoring such a haecceity is to either argue that it consists in the irreducible first-person perspective as such, or that it is somehow engendered by divine fiat. Given the former response, immaterial continuity collapses into the phenomenological criterion. The assumption of divine fiat as a basis for haecceity, while perfectly metaphysically workable, will naturally indebt the position to certain types of theism, and indirectly presupposes some equivalent to or analogue of the intentionality which the phenomenological criterion is intended to capture.

Of course, the immaterialist could also without contradiction simply qualify his or her understanding of immaterial continuity in terms of the *thisness* of the phenomenal presence we've previously discussed, and simply utilize the advantages of this criterion.

It could also be contended that that a duplication is not at all possible, supporting this position with a definition of the immaterial substrate as something immune to the particular causal and spatio-temporal laws governing physical matter (if it has no spatial designation nor physical extension, the notion of duplication as such can possibly be argued to be incoherent).

#### **4.3.2.2. Composite continuity**

The criterion of continuity of the composite entity of the immaterial and physical aspects is an interesting additional possible response to the legacy problems with Aristotelian roots. Basically, it assumes a metaphysics of the human person defining it as a composite entity of immaterial as well as material aspects, where both need to be present to constitute the complete person. This position is most closely associated with the philosophy of St. Thomas Aquinas, and particularly his reconciliation of Aristotelian metaphysics with the notion of resurrection in Christian eschatology. Aquinas' position has also been called hylemorphic dualism, and consists in a kind of compound dualism, according to which a human person cannot be equated with a disembodied, non-material self or soul without recourse to a body. Here, however central, the soul is a part of the person rather than fully equivalent to it.<sup>99</sup> However, as should be noted, it doesn't seem like Aquinas really advocated a criterion of composite continuity to be employed in relation to the conceptual problems of the post-mortem existence of persons.

---

<sup>99</sup> Jeffrey E. Brower, *Aquinas's Ontology of the Material World*, Oxford: Oxford University Press 2014, p. 284-285

Rather, it seems that Aquinas would place the haecceity of the specific individual in a disembodied soul's potentiality to become actualized in a certain way. Since Aquinas expressly distinguishes souls by virtue of specifically the intellect's potentiality to become actualized in a particular manner, using such a criterion of individuation and identification becomes functionally equivalent to our phenomenological criterion.<sup>100</sup>

On the composite continuity-response to the synchronic and diachronic problems, there are no particular surprises over and above the approaches already described starting from the positions of physical, mnemonic or immaterial-substrate continuity. If we assume that a human person is a composite entity, the response to the synchronic problem will entail that both aspects need to be somehow present or active in a given moment, enabling the synchronic persistence of the person. Likewise, the diachronic problem will be addressed in arguing for the possibility of the presence or activity of the very same enabling substrata at time T1 as well as at the subsequent time T2.

The problems and difficulties discussed in relation to the similar solutions suggested by both the pure materialist and non-materialist outlooks apply to the composite-continuity response as well, although with some minor variations. Just as we've discussed in relation to the criterion of physical or mnemonic continuity, the notion that the human person demands *a particular body* to be fully manifest is open to criticism on the grounds that particular bodies are impermanent, as well as in principle subject to the possibility of duplication.

Assuming the impermanence of physical bodies, we seem forced to seek the basis of the human person in its imperishable immaterial aspects, unless, perhaps, some irreducible haecceity of the matter in question is granted. As we recall, even the historical continuity of sets of physical entities do not render them immune to possible duplication, given the logical possibility of the replication of the correct type of causal basis as has previously been discussed.

Regarding said haecceity, if the particular set of contiguous substrata which constitutes my person, e.g. by virtue of divine fiat or the irreducible self-

---

<sup>100</sup> In *Summa Contra Gentiles*, (St. Thomas Aquinas, *On the Truth of the Catholic Faith: Summa Contra Gentiles*, New York: Image Books 1956 p. 216) arguing that there is not one common intellect among all human persons due to the fact that each soul differs in its potentiality to be actualized in a certain way, St. Thomas essentially maintains that the haecceity of the incomplete, disembodied soul lies in its intellect's potentiality. One could perhaps here speak of a relational haecceity.

consciousness of matter itself, can be considered unique and non-interchangeable, it seems possible to argue that my person is effectively constituted by a composite of particular and unique mental and material aspects. However, the panpsychist response seems to undermine the importance of a physical basis of the person which the hylemorphic dualist emphasizes – if the uniqueness of the material substrate is based in that of some non-material lower-level consciousness associated with it, we’ve indirectly assumed that the human person most fundamentally is based in an immaterial substrate. The position that the haecceity of the material substrate is anchored in the divine will isn’t obviously problematic for the same reason, yet would indebt the position to certain formulations of theism, and presupposes something similar to the phenomenal criterion. Such an assumption would also arguably seem somewhat superfluous, as the identity of the person with no contradiction *could* be exclusively based in its phenomenal first person-aspect in perfect compatibility with the hylemorphic position, but then the criterion of composite continuity obviously collapses into something similar to the phenomenal criterion of personal identity.

In relation to the problems of individuation, the solution on the composite-continuity account of hylemorphic dualism is to distinguish two separate immaterial person-entities (i.e. souls) by virtue of their affinity with particular bodies. This implies that when person X is disembodied, s/he can still be distinguished from person Y by a particular tendency to be reunited with a certain body.<sup>101</sup> However, this solution isn’t entirely satisfying, unless it is formulated as some type of immaterial haecceity. Individuating person X with regard to the particular affinity between his or her immaterial aspect and a non-haecceitic, interchangeable material substrate, seems contradictory if it’s not uniquely based in the immaterial aspect. It may, in parallel with Aquinas’ views on the uniqueness of souls, be argued that the immaterial aspect of the person has a certain tendency to become manifest as a certain type of body, to incarnate in a certain manner or the like, which would locate the basis for individuation in what might be likened to the intention of the immaterial substrate, rendering the material one functionally irrelevant. Assuming the haecceity of the material substrate for any reason or other solves the problem of individuating two ostensibly identical embodied individuals, yet, importantly, it doesn’t help us at all in setting apart two or more disembodied persons where no references to particular locations can be made, unless some direct or tacit appeal to the phenomenological criterion is made, such as the haecceity of the person’s unique soul.

---

<sup>101</sup> Baker 1997

#### 4.3.3. Brief assessment of the remaining metaphysical positions

We still haven't explicitly addressed several of the basic positions which were previously mentioned: property dualism or non-reductive physicalism, idealism, and neutral monism. Considering these positions in the above order, beginning with property dualism, this perspective seems quite incapable of uniquely providing any useful solutions to the legacy problems of death and the afterlife over and above those already addressed in relation to reductive physicalism and dualism.

Most obviously, this is due to the ontological primacy most forms of property dualism ascribes to the material basis, which on the non-interactionist accounts of property dualism renders the non-reductive properties causally inert. The interactionist variant, on the other hand, cannot obviously supplement the substance dualist's solutions. To briefly summarize the position, property dualism basically entails that there are non-reductive properties ontologically distinct from the material substrate from which they spring forth. The most common accounts of property dualism are epiphenomenalist, meaning that the emergent properties, although irreducible, still cannot influence the material substrate, while other accounts are interactionist, either affirming some variant of downward causation, or merely accepting that, for example, mental properties may cause changes in other emergent non-material properties.<sup>102</sup>

However, neither of the two variants of property dualism seems able to present any new criteria in addition to those already mentioned in connection with general dualism. Epiphenomenalism, while affirming the ontological distinctiveness of the non-physical properties, seems unable to add anything beyond the criteria of physical continuity and continuity of memory since the non-physical properties are taken to be causally inert. This assumption implies that any hypothetical continuity of an immaterial property necessitates a continuity of the material substrate which enables the existence of the former.

Similarly, interactionist property dualism in this context adds nothing in terms of solutions to the legacy problems, beyond what the immaterialist positions substance dualism or idealism can bring to the table. Interactionist property dualism is just like standard substance dualism capable of utilizing the physicalist solutions of physical and mnemonic continuity, and could also accommodate some variant of the criterion of immaterial continuity. However, this seems to be dependent on whether or not the irreducible men-

---

<sup>102</sup> Jack Crumley, *A Brief Introduction to the Philosophy of Mind*, Oxford: Rowman & Littlefield Publishers Inc. 2006, p. 23-26

tal properties are ascribed a reasonable causal independence from their material substrate, i.e. if the causative effects of mental properties are not merely considered to be indirect effects of material properties or phenomena.

Looking to neutral monism and associated positions, the situation is somewhat different. If we assume there is only one fundamental basic substance, variously manifesting as either mental or physical phenomena and entities (and possibly other types as well), any basis for continuity of the person must reside in this neutral substrate alone. There seems to be little room to argue for a meaningful continuity based in the particularly mental or material aspect of the neutral base substance, as separate from this substance itself, since the base is necessarily implied whenever either its material or mental manifestation comes into play.<sup>103</sup> Unlike interactionist property dualism, where the mental property and the material substrate are two different things ontologically, neutral monism admits of only one substance. However, neutral monism could obviously be formulated so as to mirror property dualism and provide the same methods of attacking the legacy problems.

Thus, the neutral monist can address the twin problems of diachronic and synchronic continuity assuming an analogue of the materialist's physical continuity, as well as the memory criterion. It also seems that the neutral monist may employ a criterion on par with that of objective immaterial continuity, previously discussed, if the neutral substrate is taken to be non-spatialized and physically incorruptible (i.e. not subjected to the constraints of time and space).

However, this possibly invites certain contradictions if the physical manifestation of the neutral substrate is simultaneously to be considered equally real. Assuming e.g. the fundamental non-spatiality of the basic substrate, it seems that we cannot without contradiction also grant it being definitely localized. Of course, a way around this issue is to consider time and space ephemeral aspects of the neutral basis, which then as such would be free of the limitations of physical reality. Nonetheless, if the physical aspect of the neutral substrate is secondary, if the substrate is most fundamentally a non-physical substance, the position becomes open to critique on the grounds that it's functionally indistinguishable from idealism, a problem further compounded

---

<sup>103</sup> William Jaworsky, *Philosophy of Mind*, Malden: Blackwell Publishing 2011, p. 246, 258-262



if we regard time and space as ephemeral or non-essential aspects of the neutral basis.<sup>104</sup>

#### 4.4. Prominent arguments for post-mortem survival: physicalist-friendly arguments

At this point, we will specifically present and assess some important complete contemporary arguments for the position that human persons may possibly, or will probably, survive corporeal death. The arguments will be examined in terms of plausible soundness, their metaphysical assumptions as to whether or not they can provide the necessities for survival, and to what extent they can adequately address the legacy problems discussed in the beginning of the chapter. Again, the basic necessities for the affirmation of the possibility of survival was stated to be a coherent model of the experiencing subject that without contradiction can be taken to persist or to be reconstituted in spite of the destruction of its particular body, as well as a metaphysical schema according to which such a subject model is possible. The legacy problems are fundamentally the issues of personal identity; the related sameness or individuation problem; and the question of which metaphysical positions are compatible with, or most favorable with regard to the notion of an afterlife. We will begin with such arguments and cases that are compatible with a reductive physicalist anthropology, and then move on to specifically non-materialist arguments.

As it happens, the materialist-friendly arguments here presented are advanced by Christian theists who presume a non-materialist background metaphysics. The arguments as such, however, are in principle compatible with a physicalist anthropology.

##### 4.4.1. Zimmerman's case

To begin with, we will examine Dean Zimmerman's case for survival, a metaphysically broad argumentation intended to be widely applicable, presented in his "Personal Identity and the Survival of Death", in *The Oxford Handbook of Philosophy of Death*.<sup>105</sup> In this text, he intends to give both the

---

<sup>104</sup> This critique seems to first have been suggested by V.I. Lenin in his 1909 work *Materialism and Empirio-Criticism* (*Материализм и эмпириокритицизм*), Moscow: Foreign Languages Publishing House 1959

<sup>105</sup> Dean Zimmerman, "Personal Identity and the Survival of Death", in Ben Bradley, Fred Feldman, Jens Johansson (eds.), *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012

reductive physicalist and the dualist viable grounds for assuming the possibility of survival. Before introducing his arguments, Zimmerman emphasizes the importance of the role played by the persistence conditions in the assumed metaphysics of the human person, which determines the possible responses to the problems of continuity. Zimmerman particularly focuses on the affirmation or negation of what he terms “the doctrine of temporal parts”. This doctrine can essentially be described as the position that an object, such as the human person, actually consists of parts on different positions on a timeline. Any existing entity is according to this position both spatially and temporally extended, and can be considered a distinct set of spatial and temporal parts:

Taking up space is a matter of occupying many different locations; and the normal way to do that is by having different parts located *just* in those different locations. Similarly, according to the friends of temporal parts, things “take up time”—that is, things exist at more than one time—by having different parts located *just* at the different times at which they exist. So, if I exist at noon and also at midnight, there is an instantaneous part of me that exists just at noon, and another one that exists just at midnight—each with exactly the shape and size I have at the moment it exists.<sup>106</sup>

Assuming the doctrine of temporal parts, Zimmerman contends that human persons may without contradiction possibly survive physical death, both on a dualist and materialist ontology of the self, if only theism as a background metaphysics is granted. Zimmerman examines two variants of third-person criteria of continuity, biological continuity and psychological continuity. He maintains that both types of criteria can plausibly be satisfied by variants of theistic resurrection, most obviously utilizing the doctrine of temporal parts, but also without it.

From this doctrine, Zimmerman argues towards what he calls a “protean” account of the human person, according to which the persistence conditions of a person is dependent upon his or her intentional, conceptual self-reference. In other words, exactly *how* we conceive of ourselves is taken to determine whether or not we may survive physical death. More exactly, the “person-stage”, a minimal candidate for the set of temporal parts that is taken to make up a person, is assumed to be a kind of referring subject, whose referential act then generates the entity that possibly may survive. If a certain set of person-stages refers to itself by means of a psychological continuer, this intention determines their ontological status as psychologically continuous, and vice versa.

---

<sup>106</sup> Ibid.

This interesting account is supplanted with a few additional conditions intended to disarm certain objections. The most important of these are the *intended constancy of the first-person reference* and the *local determination of local persistence conditions*. The constancy-condition entails that when one person-stage refers to some larger set of person-stages which themselves also self-refer in a similar manner, they should all, if possible, be interpreted as referring to the same set of person-stages. The local determination-condition entails that if a series of person-stages organize themselves around a particular referent during a certain period of time, these stages constitute a persisting person iff there is no branching of, or discontinuity in the referent during that period. The second condition is intended to preserve the person, if, for instance, a change occurs so that the referent-set suddenly begins to understand itself in terms of psychological continuity instead of, e.g. exclusively biological continuity. On these assumptions, Zimmerman argues that the doctrine of temporal parts can support an ontology of the person according to which his or her self-referentially generated persistence conditions may change, while the person as such still remains.

Given this, survival is possible on both dualism and physicalism, as long as theism and the doctrine of temporal parts is accepted. In case a set of person-stages refers to itself as in terms of a psychologically continuous entity, Zimmerman claims that survival is trivial, since all that is required is that God

... initiates a new series of person-stages, at some unspecified time and place, with mental states that are the natural continuation of the original person's psychology. The cause is a reliable one, so long as God resolves to pick up the pieces of our mental lives more or less where we left off.<sup>107</sup>

This would arguably be a way of satisfying the memory criterion-variant of psychological continuity as previously discussed. Unless it's formulated along explicitly phenomenalist lines, however, it faces the same objections as were mentioned in relation to the memory criterion, such as the problems of acceptable and non-acceptable causal connections (and the issue of saving the former without excessively ad-hoc assumptions), as well as the possibility of duplication and the resulting difficulty of individuation.

Zimmerman also specifically argues that the doctrine of temporal parts makes a post-mortem survival possible on a materialist ontology, albeit on the assumption of theism. Assuming a series of person-stages that stubbornly refuses to refer to itself in terms of psychological continuity, God could possibly simply force the resurrected set of person-stages to refer to itself in

---

<sup>107</sup> Ibid.

terms of a psychological continuer, with the relevant causal relationship between the former, non-psychological self-reference and the subsequent psychological one, so that continuity presumably is established.

When we get to the issue of an afterlife on the rejection of the doctrine of temporal parts, Zimmerman finds the situation a bit more difficult for the materialist. The dualist need not worry very much, since the unique piece of mental substance that is a person can simply be preserved by God after the person undergoes bodily death. On materialism, and the position of exclusively biological continuity, the situation is different. Zimmerman proposes three necessary conditions for the persistence of persons as material objects, which seem to preclude any form of postmortem existence of the person. These conditions are as follows. First we have the principle of gradual replacement, which entails that the material constituents of a physical entity cannot be replaced all at once, while the object remains identical with itself. Secondly, there is the position that living biological entities cannot continue to exist as dead matter (the rejection of the quaint idea that you somehow can survive death since “you” continue to exist as a corpse). Finally, we have the no-causal-gaps criterion, which states that all stages in the temporal history of a material object must be causally related to prior such stages.

To allow survival in spite of these three criteria, Zimmerman proposes what he calls his “falling elevator model”, which illustrates an argument to the effect that survival in terms of a divinely mandated reassembly is possible, if the criterion of gradual replacement is rejected. The argument is essentially that the person can be rescued from annihilative death, if God miraculously imbues every piece of matter (e.g. every atom) in the dying person’s body to cause a duplicate of itself at a later time and another place, at the very moment that the person’s body dies. The model can without problems satisfy two out of three criteria for physical persistence, and the rejection of the principle of gradual replacement is defended on the grounds that no such principle necessarily seems to hold with regard to elementary particles, and by extension neither should be presumed to determine the ontology of the objects that are constituted by such particles.

All in all, Zimmerman’s case must be regarded as an ingenious response to the issue of the possibility of post-mortem survival. One caveat, however, is the peculiar importance of intentional, conceptual self-reference for the pro-*tean* account, and the question of how it may be the case that a person-stage’s self-conception actually can determine its own persistence conditions, and how the implied self-causation can be avoided. Another problem is the seeming presumption of an actualized phenomenal subject performing

this generative act of intentional self-reference, and how its own persistence conditions then would be satisfied in turn.

Neither is it entirely clear how the case is able to adequately address all the legacy problems. There seems to be an implicit reductionist assumption throughout the case, as the human person is always regarded in terms of third-person metaphysical criteria. If this was interpreted as excluding what we've called the phenomenological criterion, Zimmerman's case would, even on the psychological continuity-path, face the same issues of possible duplication and disembodied individuation as we discussed in relation to mnemonic continuity. His materialist reconstitution-account seems to fare better with regard to the legacy problems due to its arguable absence of disembodiment, even though it's in some sense hampered by the addition of ad hoc hypotheses of miraculous acts of copying. However, any reductive materialist account of the person will always be faced with certain conceptual issues of synchronic persistence, as is emphasized in the possibility of duplication. These issues will arguably remain unless the materialist account actually defers to some phenomenal reality in terms of a first-person presence, or by otherwise establishing the relevant haecceity, such as in terms of unique, divine intentions which would be functionally similar to phenomenal deference. A no-branching clause can always be used to avert the possibility of duplication, but such assumptions are themselves not without problems, and would arguably add to the ad hoc-load of the model. If a relevant haecceity can be established with reference to divine action, the issue can obviously be avoided, but on this position, there seems to be at least a *prima facie* problem in identifying the basis of such a haecceity as applied to a material object consisting of interchangeable parts, or an objective pattern of information.

#### 4.4.2. Van Inwagen's argument

Peter van Inwagen's materialist argument for post-mortem-survival presented in *The Possibility of Resurrection and Other Essays in Christian Apologetics*, is somewhat akin to Zimmerman's position on the matter.<sup>108</sup> Van Inwagen specifically focuses on the problem of causal gaps in the materialist accounts. He rejects Aristotelian and Thomist accounts on the assumption that they necessarily imply such a causal gap, as well as entail absurdities. Instead, van Inwagen opts to argue for the metaphysical possibility of theistic resurrection on the grounds that God could conceivably preserve the corpse of a dead person, and replace it with a simulacrum, while the former is later resurrected with causal continuity ostensibly preserved. Since the

---

<sup>108</sup> Peter van Inwagen, "The Possibility of Resurrection", in Peter van Inwagen, *The Possibility of Resurrection and Other Essays in Christian Apologetics*, Boulder: Westview Press 1998

newly dead, non-annihilated corpse can be brought back to life without any issues of continuity, it's thus argued that physically annihilative death which would make resurrection impossible conceivably is an illusion, perpetrated by God. Therefore, survival of what we think of as annihilative death is possible, since such death actually never occurs. This type of response, while arguably even more ad hoc than e.g. the "falling elevator model" does indeed satisfy the necessary conditions for the persistence of persons as material objects, and arguably shows that "survival" of what we believe to be death is at least logically possible on materialism. However, it presumes a decidedly counter-intuitive event which in its portrayal of a deceptive deity conflicts with many traditional theistic accounts to an unacceptable extent. Moreover, it's also vulnerable to the issue of duplication without positing a non-branching clause, in the same way as Zimmerman's suggestion of how a materialist survival may possibly come about.

#### 4.4.3. Trenton Merricks' account

Another position, intended to be compatible with materialist ontology is the one suggested by Trenton Merricks, exemplified by his account in *The Oxford Handbook of Philosophical Theology*.<sup>109</sup> This position is also explicitly theistic with regard to its background metaphysics, but approaches persons entirely in terms of physical objects. Quite similar to van Inwagen's position, Merricks identifies several problems with the possibility of post-mortem survival on a materialist metaphysics, but contends that they're not decisive. Like Dean Zimmerman, Merricks' focus is on the problem of continuity, more precisely on the problem of spatio-temporal gaps and the issue of reassembly. As an important basis, it is established that a variant of no-causal-gaps criterion is to be preferred over the criterion of strong spatio-temporal continuity.<sup>110</sup> Merricks' interpretation of this criterion is somewhat weaker than Zimmerman's equivalent, since it admits of indirect causal relations by proxy. Necessary causal relations can on Merricks' assumption thus possibly bridge a spatio-temporal gap, e.g. if a time machine acts as a proxy:

Suppose, once more, that the time machine sends you to the future. You arrive in the future with a familiar tattoo on your leg. That tattoo's being on your leg was caused not only by a youthful lapse of judgment, but also—and more importantly—by your having that very tattoo on your leg before entering the time machine. This implies that causation can occur across a temporal

---

<sup>109</sup> Trenton Merricks, "The Resurrection of the Body", in Thomas P. Flint, Michael C. Rea (eds.) *The Oxford Handbook of Philosophical Theology*, Oxford: Oxford University Press 2009

<sup>110</sup> Cf. van Inwagen, Peter, "The Possibility of Resurrection," *International Journal for Philosophy of Religion* 9: 1978, pp. 114-121.

gap, since the way your leg was before time travel causes it to be a certain way after time travel.<sup>111</sup>

Merricks then addresses the objections to an admission of the possibility of spatio-temporal gaps in a body's "career" by two arguments. The first and arguably most contentious maintains that it's entirely possible that some instances of identity of a material object over time actually have no ground. This would imply that identity over time can be maintained without it having any particular foundations, and thus that the no-gap criterion need not be violated by instances of post-mortem resurrection in a bodily sense. The second argument concedes that *something* must ground identity, but instead maintains that the theist can simply assume that this criterion will be satisfied in the event of actual resurrection, even if we currently have no means of discovering what the grounds for identity actually is. Merricks concludes that he can thus see no reason to accept that there both is a condition necessary for the preservation of bodily identity, and that this condition cannot be satisfied across a spatio-temporal gap.

Merricks' account invokes the background metaphysics of theism to the least degree of the three arguments or cases which have so far been addressed, but is also not entirely convincing. An argument to the effect that some instances of bodily identity over time actually has no grounds at all has a lot of weight to carry, as it needs to convincingly address the *prima facie* contradiction it entails. It seems to tacitly presuppose a denial of the principle of sufficient reason, which should be enough to give any theist pause. Conceding that identity actually needs to be grounded, the position that the theist can simply assume that something inherent to the physical object in question will satisfy the criterion of bodily continuity, even in a situation when the body as such is utterly destroyed, is neither very plausible (if the body is destroyed, what factor inherent to the body could maintain its continuity?). In solving this issue, the persistence-condition of a particular person can very well be considered to be nothing but a unique divine intention, but then the person is of course something more than a physical object.

## 4.5. Immaterialist arguments

Finally, we will recount a few examples of arguments for post-mortem survival anchored in explicitly immaterialist anthropologies, such as idealism or dualism. The classical arguments towards what is often phrased as "the immortality of the soul" all share a similar structure, and are all compatible with our phenomenological criterion, whether or not they clearly identify

---

<sup>111</sup> Ibid.

some equivalent of the phenomenal self with the unique, immaterial part of humans that possibly survive corporeal death.

#### 4.5.1. Swinburne's argument

Swinburne's position with regard to the metaphysical possibility of an after-life is based in an explicitly dualist metaphysics, and mirrors what we've called the phenomenal criterion well. In his recent work *Mind, Brain and Free Will*, it is argued that the substantial, immaterial human self faces no inherent limitation as to the particular kind and length of its possible mental life, due to its *prima facie* independence of a corruptible body. Moreover, Swinburne maintains that the mental substance which is identical with a person, possesses a certain "thisness which makes that person what they are", which is independent of any possible thisness inherent in physical matter.<sup>112</sup>

Swinburne's response to both the synchronic and diachronic problems are very similar to that which we have presented in terms of the phenomenal criterion. In terms of the synchronic problem, he argues that the synchronic presence of subjective phenomenal consciousness is sufficient to ground identity within the particular moment. With regard to the diachronic issue, Swinburne argues that no continuity of any sort is either sufficient or necessary to ground personal identity, but that the particular, unique thisness which inheres in the immaterial substance which is a person effectively grounds personal identity in a similar manner as we've discussed in relation to the phenomenal criterion, although Swinburne specifically maintains that this "thisness" must inhere in a particular mental substance.<sup>113</sup> Also, Swinburne never explicitly identifies the haecceity of the mental substance with the phenomenal self, the mineness of the first-person perspective, or the what-it-is-like to be a particular person in the subjective sense, yet his position is obviously compatible with such an identification, and arguably makes the identification necessary due to the problems discussed in relation to objective immaterial continuity above, and Swinburne's rejection of any kind of continuity as a prerequisite of personal identity.

Swinburne's identification of the human person with an immaterial substance, identified with regard to a particular thisness, thus seems to allow for the metaphysical possibility of post-mortem survival, either in terms of

---

<sup>112</sup> Richard Swinburne, *Mind, Brain and Free Will*, Oxford: Oxford University Press 2013, pp. 164-165

<sup>113</sup> *Ibid.* p. 147-148; 151; 162-163



reembodiment or a continued, disembodied subjective experience, insofar as substance dualism is tenable.

#### 4.5.2. Hasker's argument

William Hasker's line of reasoning, first presented in his *Metaphysics* from 1983, is similar to Swinburne's. Hasker maintains that it's logically impossible to actually recreate a deceased person on a reductive physicalist anthropology. His argument is that since there can be no unique identity conditions for an individual considered as a physical object, on such an anthropology it's incoherent to claim that a recreated or resurrected body is actually the very same person as that which was disintegrated.<sup>114</sup> He gives an example to this effect in the form of a thought experiment wherein a person is killed, then replicated and brought back to life while the corpse remains in place, ruling out any possibility of identifying the living body with the dead one.

This is roughly equivalent to saying that on a materialist anthropology, there are no solutions to the problems of identity or the problem of individuation, which renders post-mortem survival an incoherent notion within such an ontological framework, and is clearly in line with our discussion in relation to these very problems. Therefore, Hasker argues, the only possible way to affirm post-mortem survival is to maintain that the "core person" by which he indicates the mind or the soul, somehow survives bodily death, the possibility of which he contends is granted by dualism.<sup>115</sup> Like Swinburne, however, Hasker employs no phenomenological terminology in discussing this solution, still, his position is fully compatible with our phenomenological criterion.

#### 4.6. Conclusions

Having reviewed the major basic perspectives on the nature of the human person that are of clear relevance to the three problems in focus, we find that they offer us a plethora of possible responses. The synchronic and diachronic problems receive similar kinds of solutions from all the positions discussed here. The synchronic persistence of the person is quite simply taken to possibly hold iff whichever substance is assumed as the basis for the person persists in a given moment.

---

<sup>114</sup> William Hasker, *Metaphysics*, Downers Grove: InterVarsity Press 1983, pp. 76-80

<sup>115</sup> Ibid. p. 80

Diachronic persistence can be argued for in a similar manner, although this problem is somewhat more delicate, since diachronic persistence necessitates that the state of synchronic persistence in one given moment can possibly function as the grounds for an equivalent state in the next. Assuming physicalism, this can, for instance, be achieved by identifying the person with a particular physical structure in one moment, and assuming that the presence of said physical structure can efficiently cause the presence of a relevantly similar physical structure in another succeeding moment.

However, this solution has two major drawbacks:

1. As we have seen, continuity on the grounds of informational history provides no matter of fact regarding person X at time T1 which is necessarily retained by person X at time T2, other than a rather vague continuity of informational history. Lacking such matters of fact, nothing seems to make it necessarily the case that person X is still the same person, rather than person Y – one could then, objectively speaking, without contradiction be a new person in each moment.

2. Also, an equivalent continuity of informational history could in principle be shared by other entities besides person X at time T2, making them possible contenders to be identified as having been person X at time T1.

The qualified memory criterion, employing the elusive notion of a relevant type of causal connection, does seem to solve the second problem, but cannot quite satisfyingly deal with the first one. Even if we speak of “proper causal connections”, there still is no specific objective matter of fact regarding person X at time T1, which is also necessarily present in person X at time T2 due to merely objective historical continuity.

The non-reductionist basing personhood in objective aspects of e.g. an immaterial substance, can sidestep these problems, yet faces severe difficulties of her own. 1. or an equivalent thereof does not arise for the dualist, since the immaterial core of the person could possibly remain essentially changeless, carrying all relevant aspects of person X at time T1 into the next moment, time T2. The duplication issue of 2. can be avoided by the dualist on the assumption of a haecceity of said immaterial core of the person, precluding any possibility of multiplying the person.

In regard to the third problem, however, that of the individuation of the person, responses exclusively based on the notion of objective substances or similar substrata nevertheless seem ineffectual. The physicalist position can-

not enable a meaningful response to the analogue of the problem of individuation in the situation of a hypothetical physical duplication, mainly since there's no room for an irreducible haecceity.

On the other hand, the immaterialist position, responding on the assumption of an objective immaterial substance, *can* make room for an irreducible haecceity of the immaterial core of the person. Still, in expressing and defining this haecceity, it seems that the non-reductionist is best served by entertaining something like the phenomenological criterion we discussed in the introduction. First of all, it's not entirely obvious in exactly what this "thisness" or haecceity would reside when we're discussing an objective immaterial substance. Moreover, if we accept this relatively problematic objective attribute as a criterion for thisness, the position suddenly becomes open to some equivalent of the duplication problem, demanding a convincing response as to why an objective immaterial entity is supposedly immune to duplication.

Insofar as the physicalist-friendly arguments presented under section 4.4. work, it's either in the denial of annihilative death as such (van Inwagen), by tacitly assuming phenomenal subjectivity (Zimmerman), or due to the reliance upon God's intentionality as a foundation given which the problems of identity and individuation can be addressed (Merricks). If one thus bases the solution in the ineffable intentions of the transcendent deity, in the final analysis, one implicitly assumes an equivalent to the phenomenal criterion – a person's uniqueness is then an active divine intention. A further issue is that denial of the actuality of death cannot by itself adequately respond to the problem of individuation.

It seems that the assumption of the phenomenological criterion of personal identity effectively provides the proponent of any ontology besides strictly reductive physicalism with a workable metaphysical solution to the synchronic and diachronic problems, as well as the issue of individuation, without necessitating the specious assumption of objectively available thing-like haecceities. This criterion is clearly available with regard to the strong immaterialist positions such as idealism or dualism, but it can with little difficulty also be employed by non-reductive materialists or property dualists. As long as the non-reductive materialist accepts the presence of some aspect, attribute, mode or substance that is not exhaustively reducible to the physical substrate, this can be used to anchor the ontological independence of the phenomenal, which is all that the phenomenological criterion of post-mortem survival needs to potentially be fulfilled.

It's not relevant if this aspect, attribute, mode or substance anchoring the phenomenal is determined by the character of the physical substrate, such as according to certain emergence theories, for as long as irreducibility of the property in question is granted, the internally unique, possibly subsisting phenomenal self can be metaphysically grounded.

All in all, the only kind of metaphysical position that necessarily cannot accommodate the criterion is reductive physicalism. Such an anthropology would however allow for the metaphysical possibility of post-mortem survival, utilizing something like van Inwagen's argument, if it could ground an adequate response the problems of identity and the problem of individuation. The prospects of doing so, however, do not seem promising at this point. Indeed, as Shelley Kagan argues, it's difficult to support even synchronic identity on a reductive physicalist anthropology.<sup>116</sup> Also, importantly, van Inwagen's and Merricks' arguments presuppose a theistic background metaphysics – they're not globally physicalist.

Thus, a global reductive physicalist ontology seems to prohibit the metaphysical possibility of post-mortem survival. This category of metaphysics will be criticized in detail in chapter 8.

As a final remark, many dualists or idealists would likely consider our criterion too inclusive at this point, and prefer that e.g. dual-aspect theory or non-reductive physicalism were excluded from the positions possibly satisfying the phenomenal criterion, for instance due to the often purported incoherence of positing immaterial properties on a physical substrate. Our response to these critics will simply be that if it is true that the existence of any form of immaterial properties necessarily implies some variant of dualism or idealism, the criticism of reductive physicalism which is presented in chapter 8 must then be taken to support the truth of these metaphysical positions exclusively.

---

<sup>116</sup> Kagan 2010

## 5. A basic metaphysics of mind: the ownness model of the minimal self

### 5.1. Introduction

This chapter will provide a model of the self, compatible with the phenomenological criterion discussed in chapter 4. Our model will be compared with other important self-models related to the affirmation of post-mortem survival, such as Swinburne's and Baker's. This self-model is based in the notion of the first-person perspective, and will integrate the phenomenological criterion, while remaining as ontologically neutral as possible. This is to say that the self-model is intended to be compatible with all ontological positions that can be reconciled with the phenomenal, subjective mineness of the last chapter's criterion.

How such a movement from a first-person experiential phenomenon to something more akin to a delimited subject may play out, is not entirely obvious.<sup>117</sup> Descartes familiar suggestion, for one, was that we assume the basic, private, subjective experience of our own presence to ourselves, implying that consciousness as such represents *prima facie* evidence for the existence of the self as an experiencing subject.<sup>118</sup> At this point, however, he identified this subject as essentially “a thinking thing” (*res cogitans*), and thereby proceeds by arguably reifying subjective experience as the very type of ontic object which phenomenology insists can never fully accommodate the realities of *being's* irreducible presence.<sup>119</sup> Descartes' transition from a given first-person experiential phenomenon to his *res cogitans* seems to rest upon the assumption that any and all aspects of reality must necessarily be analyzed according to something like a substance metaphysics, possibly

---

<sup>117</sup> The definition of phenomenon we assume is that of the phenomenon as “an immediate object of awareness in experience”, generally distinguished from the Kantian “thing-in-itself” or “noumenon”, cf. Martin Heidegger, *Wegmarken*, Frankfurt am Main: Vittorio Klostermann 2004, p. 367

<sup>118</sup> René Descartes, *Meditations on First Philosophy*, Oxford: Oxford University Press 2013, pp. 18-20

<sup>119</sup> *Ibid.* p. 21-24

genealogically related to the Aristotelian position that predication always must involve some kind of substance.<sup>120</sup>

## 5.2. The basics of the model

We may however conceive of a similar movement which lands in a perhaps less ontologically entrenched position, yet assumes premises akin to those of the Cartesian argument, based in our phenomenal criterion.

Such attempts have previously been made from a variety of perspectives, especially in the sphere of late-modern 20<sup>th</sup> century philosophy where insights from analytic as well as continental thought have often been intertwined. Constructing the model in this way possibly provides something of a compromise between the purported extremes of analytic philosophy's onto-theological reductionism and the continental traditions' bemoaned obscurantism,<sup>121</sup> and if so, it will be compatible with a wide variety of metaphysical positions.

Our suggested way of conceiving a delimited human subject likewise assumes the first-personal experience of reality as a premise for further analysis, but rather than necessarily going from this notion to a particular metaphysical position such as dualist substance-metaphysics, it settles for a more ambiguous description of the subjective self, focusing on the ownness of the phenomenal first-person perspective. It may of course be argued that the reality of the phenomenal implies the existence of non-physical substances in the sense that e.g. Richard Swinburne contends, yet we will not assume a strong position in this particular debate within this dissertation.<sup>122</sup>

On ownness, the idea with this perhaps peculiar term is that the ostensibly private, subjective experience of one's own presence to oneself, where Descartes starts off, can most obviously be distinguished from other such experiences by something which may be referred to as its particular quality of ownness; by *for whom* they are primarily present from a first-person perspective. This presumes that what we may describe as the *primary givenness* of the phenomenal first-person perspective, the fact that experience as such is immediately present in the subjective mode. This primary givenness in turn necessitates a minimal *for whom* to which experience is present, which

---

<sup>120</sup> Ibid. XXX; Aristotle, *Metaphysics*, V.viii.1 (1017b), VII.iii.3 (1029a)

<sup>121</sup> Steven Postrel, Edward Feser, "Reality Principles: An Interview With John R. Searle", *Reason*, Feb. 2000

<sup>122</sup> Cf. Swinburne 2013, ch. 6

can be described in a more or less ontologically neutral manner, for instance as a “center of consciousness”; or, to further emphasize the seemingly irreducible *mineness* or ownness of the subjective phenomena; an “own-sphere”.<sup>123</sup> This is strongly related to Nagel’s and Sprigge’s conception of the self, exemplified in the bat-argument.

The two basic assumptions of this ostensibly neutral notion of the minimal “subject-entity” associated with the first-person-perspective are relatively straightforward: A. First, it’s taken to be true that *there is consciousness*; that the immediate, intentionally couched first-person experience of *something* is a given aspect of reality. B. Secondly, *the irreducible unity* of the first-person-aspect of this phenomenal experience is a necessary premise if one is to posit a singular, coherent subject-entity of an ontologically primitive character as the bearer of said conscious experience. We will now elaborate on the two assumptions, A and B, in order.

### 5.2.1. On the presence and character of subjective consciousness

The primary assumption A, that there is subjective consciousness is generally taken to be self-evident, at least from the individual’s own perspective, as it’s for the most part considered incontestable that *I do experience something* from a first-person perspective, since this phenomenal consciousness of *something* from the point of view of the “I”, is the default mode of presentation of my awareness of anything at all. Still, it must be said that certain philosophers have endeavoured to argue against this and similar assumptions, thus denying the existence of even any form of “mental sphere” whatsoever; denying the presence of phenomenal consciousness. B. F. Skinner, for instance, approached this position with his reductive radical behaviourism, contending that subjective mental experiences are to be regarded as quasi-real “explanatory fictions”<sup>124</sup> without any actual causal function or independent existence (this was however mainly in polemic against ideas related to interactionist dualism). Similar positions in the category of reductive physicalism do not explicitly deny the first-person perspective per se, but by extension imply the conclusion, since existence is only attributed to objective, material reality, which cannot accommodate the presence of a subjectively phenomenal one. This is roughly the essence of David Chalmers’ “hard problem of consciousness”; the seeming inability of reductionist physicalism to satisfyingly explain the presence of phenomenal consciousness, or “qualia”, that which is *primarily given* from the point of view

---

<sup>123</sup> Klawonn 2009, p. 58

<sup>124</sup> B. F. Skinner, “The Operational Analysis of Psychological Terms”, *Psychological Review* 52, 4, 1945

of a subject.<sup>125</sup> A common type of follow-up argument on the part of the eliminative materialist is then to make the case for a physicalist understanding of the first-person-perspective as such, of which there have been many interesting suggestions during recent years which we will later examine, but the outright rejection of the phenomenal first-person experience as such comes off as counter-intuitive, arguably denying what is immediately and inevitably present to experience. This critique and others like it will be engaged with in detail in chapter 8.

Setting such problems aside for the moment, we contend that we can without any glaringly obvious contradictions entertain the primary assumption, designated as A, as a basis for the suggested ownness- or givenness-model of the minimal self.

### 5.2.2. On the indivisibility and unity of subjective consciousness

Assumption B, the indivisibility of subjective consciousness is essentially the idea that the first-person experience as such cannot be reduced to any potential constituents, and thereby at least potentially may be considered ontologically basic. Our discussion of the indivisibility or irreducibility of first-person consciousness has been briefly touched upon in chapter 4.3.1, where the conceptual indivisibility of the first-person perspective was maintained. This was related to the two problems of personal identity; the problem of explaining how any particular self can remain the same (identical with itself) through time and different circumstances, as well as the necessary and sufficient conditions for identifying a particular self at a given moment in time.

This position we endorsed in response to these problem, again, contends that what makes a person identical with herself over time and in different situations is the persistence of a certain fundamental feature or entity. Commonly suggested candidates for such an entity have been *consciousness*, the “*I*”, *selfhood*, the *soul*, etc. It’s worth mentioning that the non-reductionist position on phenomenal consciousness doesn’t necessarily nor normally construe the “*I*” or the self as permanent in its complete form, and is perfectly capable of accommodating the idea that our complex personal identities in the psychological sense, are composite structures subjected to change. The non-reductionist minimally must affirm the existence of at least *some* basic feature of the person, which is of a permanent and irreducible nature. Neither does B beg the question in terms of definitionally ruling out such reductive

---

<sup>125</sup> David Chalmers, “Facing up to the Problem of Consciousness”, in *Journal of Consciousness Studies* 2: 200-19, 1995 p. 202



physicalist models of consciousness as self-representationalism, since these models may perfectly well approach the phenomenon of subjectivity as an either-or affair (either a particular system implements self-representation or not).

In elaborating on the issue, if we look to the aspect of primary presence, givenness, or the “phenomenality” of phenomena from the first-person point of view, we find that this basic primary presence is the only conceivable basis of the “ownership-aspect” of all experienced mental phenomena. The experience of “mineness” inherent to any and all conscious experience simply necessitates the existence of a first-person perspective.

This experience of mineness is generally entangled in a complex web of experiences, emotions and memories pertaining to ourselves as complete persons; the elaborate selves of ours which we e.g. attain through physical development and interaction with the environment as well as significant others already in the stage of infancy.<sup>126</sup> The fact that this complete person potentially can be reduced to its more primitive levels, and drastically altered by destruction of nervous tissue, may lead one to surmise that this is also the case regarding the basic experience of mineness resulting from the primary presence of phenomenal experience. But the two are, in actuality, ontologically distinct. The complete person is, again, an embodied complex system of experiences, of habits, memories and emotions which taken together form a somewhat coherent self-description for the specific consciousness to which this system of experiences is primarily present. Yet, the mineness-aspect of all phenomenal experience is merely a consequence of the fact that it’s present in the first-person perspective – a consequence of the basic phenomenality of said experiences and emotions as such. Thus, while the complete person as a complex system of experiences apprehended objectively may in principle be reduced to its constituents, the more basic mineness-aspect of phenomenal experience which is primarily present cannot as obviously be reduced to any more fundamental components in this manner.

The following thought-experiment may perhaps help to illustrate this ontological distinction. Consider a neurology patient with a severe degenerative brain disease which gradually reduces him or her to the cognitive level of a “lower” animal (certain disorders which are characterized by cerebrocortical lesions may be described in this manner, especially the later stages of those

---

<sup>126</sup> Marian Radke-Yarrow, Barbara Belmont et.al., “Young Children’s Self-Conceptions”, in Dante Cicchetti, Marjorie Beeghly (eds.) *The Self In Transition: Infancy to Childhood*, Chicago: University of Chicago Press 1990, pp. 345-363

characterized by degeneration of the frontal and temporal lobes).<sup>127</sup> The patient's cognitive functions are progressively diminished until he or she reaches a level of consciousness which in comparison to the healthy state is quite primitive, wherein the experience of him- or herself as a complete individual is no longer present as a framework of phenomenal perceptions.

Yet the perceptions as such are nevertheless either primarily present in the subjective sense, or they are absent; they are given in the first-person perspective, which either can be equated with, or considered producing the basic "mineness" of experience. If we attempt to conceive of a division, we must either end up with two separate, independent subjects, or something akin to the bicameral situation mentioned in the preceding chapter, since it's impossible that the division of subject S would generate two separate, non-identical subjects that nonetheless were the very same subject.

Indeed, primary presence in this basically indivisible, subjective sense is something which all experiential content must have in order to at all exist. No matter what other specific attributes any particular mental phenomenon may exhibit, whether it takes the form of a certain type of physical pain, or the complex quality of an interconnected emotional pleasure derived from revisiting the familiar surroundings of one's childhood, these phenomena share a similar inevitable *quality of presence*; they are present to the individual mind in and of themselves. This primary quality of presence is necessarily permanent in relation to the experience (it cannot be suspended while the experience persists), no matter how much the character of the particular mental phenomena changes, and is always actualized in the first person.

This is to say that the quality of presence of mental phenomena is inescapably manifest in the first-person perspective; the presence of any particular phenomenal content is necessarily accompanied by the consciousness's awareness of itself experiencing something. In other words, the system of experiences (or however we describe the phenomenal contents of a conscious mind) is primarily present to a subject experiencing them – the recipient of the experience without which there is no experience – and this subject is then the basic constituent of the first-person perspective. Even if we conceive of an exquisitely primitive mind only capable of harbouring one single instance of experiential content – let's say, for instance, the hypothetical blissful, total satisfaction of a feeding amoeba – this content is necessarily circumscribed by the first-person perspective in which it is manifest, which

---

<sup>127</sup> Alexandre Henri-Bhargava, Morris Freedman, "The frontotemporal dementias: an overview", in Michael J. Strong (ed.), *Amyotrophic Lateral Sclerosis and the Frontotemporal Dementias*, Oxford: Oxford University Press 2012, pp. 55-78

implies the ownness of the phenomenal experience as such. To state this feature of experiential content in somewhat simpler terms: it is impossible to experience phenomena (the feeling of cool water on one's skin, for instance) *without simultaneously also being conscious of one's experiencing something*. Experience as such, thereby implies a primal self-presence or self-givenness of any particular system of experiences; a presence of the phenomenal sphere to itself. This self-givenness has been described in a variety of ways; the idealist Timothy Sprigge subscribes to it by defining the "being of consciousness" as "its own knowing of itself", elaborating that "...experience is the only thing which has this kind of being in which being and knowing are one. [...] A state of [consciousness] is one with its own knowing of itself."<sup>128</sup> Merleau-Ponty expressed this quality or mode by stating that the thing perceived partakes in the being of the consciousness perceiving it.<sup>129</sup> In brief, phenomenal experience isn't possible without a form of meta-level awareness of the act of experience as such, given qualia's character of primary presence. We cannot speak of mental phenomena without this attribute of primary givenness, and in turn, this primary givenness is inconceivable without also *being present to itself*. This can be expressed as a kind of pre-reflective, immediate self-illumination inherent to consciousness as such – conscious experience is always, to an extent, reflected back onto itself as the convergence point of the particular phenomenal experience in question and the perceiving consciousness, a movement synonymous with the phenomenal first-person perspective, with "mineness" or "ownness".

Returning again to the thought experiment with the brain-damaged patient, another dimension may be added to the example which emphasizes this mineness and separates it from higher forms of complex personhood, by assuming that oneself is the patient in question, and asking whether it makes any difference whether this situation has happened to oneself or someone else. By affirming that it does indeed make a difference whether I or another would suffer and ultimately perish from this type of condition, it's implied that concepts like "I" or "mine" very well can be used to refer to more rudimentary forms of consciousness devoid of the higher mental life which characterizes full personhood. By extension, this indicates that there are no conceptual hindrances in considering that mineness is present even with the most rudimentary forms of life capable of experience as such (the slug and the shrimp is also an "I" from within, in the very same basic manner as a complex human person), and accordingly, mineness is present on an either-or-basis: if phenomenal experience is given, mineness is thereby implied. Thus, basic mineness may be described as an ontologically primitive, irre-

---

<sup>128</sup> Sprigge 2011, pp. 41-42

<sup>129</sup> Hayim Gordon, Shlomit Tamari, *Maurice Merleau-Ponty's Phenomenology of Perception*, Westport: Praeger 2004, p. 61

ducible phenomenon without contradiction. On the *prima facie*-level, it seems to be necessarily irreducible in that there are no conceivable shades of gray in this matter; mineness is either or not present, as phenomenal experience is either or not at hand. There is no partial or incomplete phenomenal experience which does not fully count as such, due to the fact that even the faintest, most basic sensation simply either is experienced or not. Further arguments for the ontological distinctiveness of the first-person mineness will be explored in chapter 6, as part of a more general examination of the viability of reductive physicalism.

What has here been stated regarding primary presence as the foundation of conscious mental content, implies that consciousness as such is indivisible and irreducible. The first-person character, or mineness of mental contents, is simply equivalent to the fact that they are exclusively present to consciousness. This can also be expressed in terms of the ontological subjectivity of mental phenomena. The first-person character of qualia is due to the subjective character of the “self-knowledge” of a given system of experiences; the fact that these experiences are only accessible “from within” a system renders them a *subjective existence* separate from the objective existence of phenomena at the third-person level of reality.

One may also approach mineness or primary presence utilizing the concept of *intentionality*, and conceive of the mineness in terms of a certain relation between these particular contents and the mind that perceives or contains them, which cannot be accessed at the objective level of the neural activity which correlates with mental content. This relation is then essentially binary – it’s either present or absent – and cannot be further subdivided.

Intentionality, a concept central to phenomenology, mirrors this relationality particularly as defined in Husserl’s sense. All three parts in the Husserlian description; the intentional act, the intentional object and the intentional content, all take place within the field of primary presence, in the subjective space of an “own-sphere”.<sup>130</sup> Intentionality can thus be thought of as a particular type of relation between a system of phenomenal experiences and the intentional objects (any type of entity of any ontological status), that is completely unintelligible without the assumption of primary presence or first-person subjectivity.

---

<sup>130</sup> Edmund Husserl, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, The Hague: Martinus Nijhoff Publishers 1983, p. 74-77

### 5.3. The ownness model: the first-person perspective as the minimal self

The two basic features discussed in part 5.2., A and B, are the foundation what we describe as the ownness-model of the minimal self. This model essentially represents the minimal self as synonymous with the first-person aspect of the presence of subjective experience. The phenomenon of a strictly subjective dimension of reality, the quality of being phenomenally present which is inextricably fused together with the existence of experiential content as such, is taken to necessitate some self, *for whom* experience exists; *for whom it is like something to be*, to paraphrase Nagel. This self can within this framework be conceptualized in various ways, but is basically characterized by its ability of actualizing experience, while it's generally thought that it cannot as such be directly experienced as a part of reality external to another subject. Incidentally, we do not wish to endorse this latter assumption.<sup>131</sup>

As this self is difficult to adequately describe in normal object language, a variety of metaphors are usually employed, some of which have already been mentioned. Prominent examples of such like “stream” or “center” of consciousness and “the phenomenal” or “own-sphere” further underscore the idea of the self as something capable of actualizing experience while it simultaneously in some sense of the word contains it. The concept of “consciousness” as such is also an example of this, stemming from the combination of *scio/scīre*, latin for “knowing, to know”, and the prefix *con-/com-*, approximately indicating “togetherness” or “joint action”.<sup>132</sup> This analysis of consciousness as a concept is compatible with the relational analysis of ownness or primary presence alluded to above, but also works well with the idea of the self as some form of vessel making experiential content possible by bringing it together. In ownness-type models the self is thus interpreted as that which enables subjective, experiential content and its coherence, or the experience and its “togetherness”.

In our ownness model, intended to satisfy the phenomenal or phenomenological criterion of personal identity, the minimal self is equated to the basic mineness or first-person perspective of subjective, phenomenal conscious-

---

<sup>131</sup> It must be noted that other minds may possibly be directly experienced by another, through something like an intersubjective participation of one in another's particular “own-sphere” or stream of consciousness. This possibility seems to be indicated by certain reports of consubjective experiences, and is reflected in the conceptions of many philosophers, such as Henri Bergson's with regard to his usage and understanding of the concept of intuition, as well as Vladimir Solovyov's understanding of intentionality as necessarily a form of communion.

<sup>132</sup> P.G.W. Glare, W. M. Edwards et al (eds.), *Oxford Latin Dictionary*, Oxford: Oxford University Press 1968, pp. 403, 1705

ness. This self is argued to be in principle irreducible, and is therefore a viable candidate for having an ontologically basic or primitive character. The self of this model is *minimal* in that it does not, for instance, presuppose the complete, elaborate and embodied self of a human person, and can therefore accommodate many types of ontologies (it does not in principle rule out embodiment as a necessary criterion of survival, in conjunction with the phenomenological one), and does not arbitrarily restrict or circumscribe the metaphysical possibility of survival.

## 5.4. The ownness model in relation to alternative conceptualizations

The ownness-model is compatible with most non-materialist conceptualizations of the human self, yet would be considered too narrow from the perspective of such dualists and emergentists which emphasize embodiment, and in a sense too wide with regard to certain other perspectives such as Lynne Baker's.

If we look to Baker's model of the self, which our minimal-self model is closely related to with regard to the phenomenological criterion Baker also anchors in her self-model, the latter makes a crucial distinction which is not reflected in our model. Baker's self-model of the human person explicitly involves a self-reflective capability over and above the mineness of the first-person perspective as such. According to Baker, the human person is the bearer of a particular type of first-person perspective which includes an explicit self-concept and a reflective self-understanding,<sup>133</sup> and it's this particular type of the first-person perspective which can anchor the phenomenological criterion, and this robust first-person perspective is the primary candidate for post-mortem survival. This distinction is clearly parallel to the one made in classical Thomism, according to which only the human animal possesses the aspects which make post-mortem survival possible, namely the intellect and the will, whereas non-human animals merely possess the sense, a more rudimentary form of phenomenal consciousness.<sup>134</sup>

Taliaferro's integrative dualist account of the human person, which stresses integral embodiment yet maintains the non-identity of the human person and the body, is for our purposes functionally equivalent to the ownness-model, with the single difference that embodiment would be added to our phenome-

---

<sup>133</sup> Baker 2013, p. 42

<sup>134</sup> Aquinas, *Summa Theologiae*, Q 75, Art. 3

nal criterion as a further necessary aspect of post-mortem survival.<sup>135</sup> Taliaferro's model is significantly thicker and more detailed than ours, but in terms of the "basic subject" that possibly survives corporeal death, they are effectively similar.

Such strictly immaterialist accounts like Swinburne's dualist model of the human person or John Foster's dualist-towards-idealist model of the self will accommodate our ownness-model without any conflicts or caveats.<sup>136</sup> Foster specifically maintains that his model of selves as basic subjects of phenomenal experience cannot be species-restricted, with the implication that it can be thought of as equivalent to the rudimentary first-person perspective of Baker's, or the basic sense-consciousness discussed by Aquinas.<sup>137</sup>

---

<sup>135</sup> Charles Taliaferro, *Consciousness and the Mind of God*, Cambridge: Cambridge University Press 1994, p. 161-173

<sup>136</sup> John Foster, *The Immaterial Self*, New York: Routledge 1991, p. 221-240; Swinburne 2013, p. 163-170

<sup>137</sup> Foster 1991, p. 236

## 6. A critique of reductive physicalism

In the last analysis, most common things will be found to be highly complicated. Some men of science do indeed get over the difficulty by dealing only with the easy part of it: thus, they call first love the instinct of sex, and the awe of death the instinct of self-preservation. But this is only getting over the difficulty of describing peacock green by calling it blue. There is blue in it. That there is a strong physical element in both romance and the Memento Mori makes them if possible more baffling than if they had been wholly intellectual. No man could say exactly how much his sexuality was coloured by a clean love of beauty, or by the mere boyish itch for irrevocable adventures, like running away to sea. No man could say how far his animal dread of the end was mixed up by mystical traditions touching morals and religion. It is exactly because these things are animal, but not quite animal, that the dance of all the difficulties begins. The materialists analyze the easy part, deny the hard part and go home to their tea.<sup>138</sup>

G. K. Chesterton

In this chapter, a critique of reductive physicalism is presented, with the purpose of ruling out the set of ontological positions which in accordance with chapter 4 are the most inimical to the viability of afterlife beliefs. By extension, the remaining ontological positions, which are compatible with the suggested self-model of chapter 5, and thus capable of navigating the legacy problems of death and the afterlife, are thereby implicitly defended. It will be argued that reductive physicalism is untenable, in that it either collapses into the incoherent position of eliminative materialism, or succumbs to phenomenism (at the very least to one of the weaker forms, such as non-reductive materialism).

### 6.1. Introductory remarks

As shown in relation to preceding chapters, the ontological positions best suited as a foundation of an effective response to the major problems associated with the notion of an afterlife, are those which provide an ontological basis of the human person outside of the strictly physical realm as it is conceived within the framework of the Cartesian-Newtonian paradigm. This is,

---

<sup>138</sup> G. K. Chesterton, *What's Wrong With the World*, London: Cassell and Company, Limited 1910, p. 47



again, to say that certain non-reductive versions of materialism can *prima facie* function as a basis for a response to these problems, together with the conventionally immaterialist positions such as dualism or idealism.

The positions with the worst prospects of adequately addressing the legacy problems of death and the afterlife are those which cannot anchor the subject properly in some substrate, however defined, that permits of a continuous or permanent identity and individuation. We argued that the preferable way to attain this was to employ the *phenomenological criterion* according to which the persistence of a particular subject obtains iff this subject's own first-person perspective or phenomenal, subjective mineness actually remains or reemerges as such. The ontological positions most clearly suited to support survival thus construed was argued to be the classical immaterialist positions of idealism or dualism, including Aristotelian-Thomist hylemorphic dualism. However, variants of so-called non-reductive materialism which retain the relevant immaterialist features also provide a framework within which the phenomenological criterion can possibly function.

As the category of reductive physicalism effectively rules out all forms of post-mortem survival as we have conceived of it, and specifically blocks what we have considered to be the most promising criterion of personal identity, a successful criticism of reductive physicalism will by extension support the rational permissibility of afterlife beliefs.

## 6.2. What is reductive physicalism?

Reductive physicalism is the ontological position according to which nothing exists but matter, conceived in accordance with how the physical is described on the Cartesian-Newtonian or post-Cartesian paradigm, i.e. where the physical matter is described as devoid of all intentionality and subjective qualities, as fully quantifiable and entirely interchangeable. The objective, abstract description of the quantifiable material constituents are taken to be exhaustive, which implies that no reliance on intentionality is necessary in conceptualizing or describing matter. It also implies the lack of any unique, distinguishing attributes of the fundamental constituents, whereby they are in principle totally interchangeable.<sup>139</sup> Reductive physicalism thus considers reality to consist only in physical properties in this Cartesian-Newtonian sense (an expanded definition of physical properties within the framework of reductive physicalism such that it includes intentionality, for instance, must be rejected from the outset, as such an expansion would entail compatibility

---

<sup>139</sup> Cf. "Materialism", in Craig 2000

with such positions as panpsychism or Aristotelian-Thomist hylemorphism or the contemporary non-reductive physicalisms). Reductive physicalism is generally also taken to contend that everything that can be explained, can be explained on the basis of physical laws and conditions.<sup>140</sup>

The set of positions which we here refer to as “reductive physicalism” is variously known as naturalism, scientific naturalism, scientific realism, materialism, reductive materialism, and by certain other names. Reductive physicalism is also associated with certain interpretations of scientism that entail reductive physicalism, yet the inverse relationship does not necessarily hold, i.e. scientism does not necessarily imply reductive physicalism.<sup>141</sup>

Reductive physicalism can thus be taken to refer a number of related positions, depending on the particular field of inquiry. Until quite recently, physicalism in general was popularly formulated in terms of supervenience, often in relation to Jaegwon Kim’s work.<sup>142</sup> The most basic formulation of supervenience physicalism would imply that it’s not possible for the world to differ with regard to such things as its phenomenal attributes and aspects, while it remains physically identical.<sup>143</sup> Supervenience models thus formulated, however, do not contradict immaterialist positions such as dualism, emergentism or panpsychism, since these, as well as many other positions incompatible with reductive physicalism, in principle allow for such strict psychophysical laws.<sup>144</sup> Neither do such supervenience models guarantee the presumed ontological primacy of the physical and the dependency of the mental, since if it is conceded that the mental exists as such in an intimate relation to the qualitatively different physical level, there’s no obvious reason that the inverse supervenience relationship and/or causal relations does not hold in reality. The mental could just as well be ontologically primary, unless the supervenience theorist can specifically show how it reduces to the physical, which would be a much stronger claim.

Thus, a stronger formulation is necessary to properly distinguish fully reductive physicalism from alternative metaphysics. There are two basic options. The first alternative is the position anchored in radical empiricism and behaviourism, today popularized by Daniel Dennett, Stephen Stich and the Churchlands, variously known as eliminative materialism, mental denialism

---

<sup>140</sup> “Materialism”, in Borchert 2006

<sup>141</sup> Mikael Stenmark, “What is scientism”, *Religious Studies*, v. 33, no. 1, 1997

<sup>142</sup> Kim 1995

<sup>143</sup> David Lewis, *On the Plurality of Worlds*, Oxford: Blackwell 1986, p. 14

<sup>144</sup> Jaegwon Kim, “‘Strong’ and ‘Global’ Supervenience Revisited”, in Kim 1995

or nihilism.<sup>145</sup> The second alternative is the purportedly non-eliminative yet reductive identity theory in its different forms, which can be representatively considered in terms of type-identity and token-identity theory.

### 6.3. Why does reductive physicalism emerge?

Reductive physicalism is a distinctively modern phenomenon, with complex historical roots. However, it's clear that the Enlightenment, post-Reformation rejection of mainly the Aristotelian metaphysical framework, in favor of a mathematically oriented empiricism ideologically wedded to the institutional role of natural science within the framework of early capitalism, ushered in the general reductionist programme of modernity.<sup>146</sup> This reductionist programme finally resulted in the explicit reductive physicalism of the early 20<sup>th</sup> century, which was particularly exemplified by the logical positivist movement.

All of the post-Cartesian ontologies we have inherited, including the immaterialist ones, reflect this transition in some sense. This is particularly evident in their conception of matter, which has been vacated of such notions as formal and final causes and inherent intentionality that have just recently begun to reintroduce themselves.<sup>147</sup> However, the reductionist influence can also, for example, be seen with regard to the Cartesian dualist conception of the mental as essentially a quasi-material object that can be objectively apprehended, the *res cogitans*,

This fundamental ontological shift with regard to how we conceived the nature of the matter of physical reality and the relationship between matter and the perceiving subject also provides the most important arguments for reductive physicalism, i.e. those related to the problem of interaction:

The shift from hylomorphism to atomism and substance dualism created what is now seen by many to be an insoluble problem: mind–body interaction. Whereas for Aristotle and his followers the soul was but one instance of form, in modern thought the mind becomes an anomaly in an otherwise purely material world

---

<sup>145</sup> Ramsey, William, "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2016/entries/materialism-eliminative/> 2016-10-11

<sup>146</sup> Cf. Geoffrey Mure, *Retreat from Truth*, Oxford: Oxford Basil Blackwell 1958 for an early ideological analysis of British empiricism, its conceptual roots and its contemporary entailments.

<sup>147</sup> James F. Ross, "The Fate of the Analysts: Aristotle's Revenge\*: Software Everywhere" *Proceedings of American Catholic Philosophical Association* 1990

of nature. Furthermore, the very conception of matter has changed. Before the atomist revolution, matter and form had been correlative concepts – matter was that which had the potential to be activated by form.<sup>148</sup>

The transition towards reductive physicalism was, in terms of a paradigm shift, by all accounts an expression of the changes in the basic conceptual structure of the dominant Western worldview, which in turn were brought about by the rapid and thorough institutional transformations of the period in question. This can be seen in the fact that the development towards the reductive physicalism of the 20<sup>th</sup> century is generally not accompanied by any particular arguments attempting to show the preceding metaphysical paradigm to be insufficient with regard to shared premises, but rather rests on the presumption of naturalism in some form or another.<sup>149</sup> This is also a presumption which is clearly reproduced by the general ideological environment and the cultural milieu of industrial capitalism.<sup>150</sup> Not even Russell and Moore, whose revolt against British idealism has generally been thought of as the definitive break with the last remaining vestiges of immaterialism, in spite of all their important, constructive work erected upon the conceptual foundations of the reductionist programme, actually provided any decisive arguments for abandoning immaterialist ontology, pre- or post-Cartesian.

Moore, for instance, mounts a devastating critique of a particular variant of British idealism known as absolute idealism, but his work on the topic during the first half of the 20<sup>th</sup> century actually gives us no compelling reason to abandon immaterialist ontologies in general.<sup>151</sup> Likewise, Russell, while

---

<sup>148</sup> Nancey Murphy, *Bodies and Souls, or Spirited Bodies?* Cambridge: Cambridge University Press 2006, p. 45

<sup>149</sup> This tendency can be seen with most of the important proto-physicalist or early physicalist thinkers of the early to late modern period, from early Enlightenment thinkers such as the experimental anti-Cartesian empiricist Pierre Gassendi, early naturalists like Jean Meslier, materialist thinkers such as La Mettrie (his *L'homme machine* introduces the basic intuitions of modern reductive physicalism as early as 1747), and Encyclopédistes such as Paul-Henri Thiry or Denis Diderot. Notwithstanding all interesting speculation and constructive contributions, the basic rationale for assuming naturalism is in these cases mostly ideological or based in the proposed implausibility (rather than ostensibly proven incoherence) of specifically *post-Cartesian* immaterialism. Not even Marx' and Engels' synthesis of Feuerbach's materialist metaphysics actually contributes any real arguments for the truth of the position as such (cf. Friedrich Engels, Karl Marx, *Das Kapital* vol. I, Hamburg: Otto Meisner Verlag 1867, ch. 1)

<sup>150</sup> Cf. Herbert Marcuse, *One-Dimensional Man: Studies in the ideology of advanced industrial society*, New York: Routledge 2002

<sup>151</sup> In "The Refutation of Idealism" (*Mind*, 12, 1903, pp. 433-453), Moore argues against the plausibility of Berkeley's conceivability principle as well as the coherence of the absolute species of idealism. The latter line of argument is very compelling, while the former does not even intend to obligate us to reject Berkeley. Moore's "A defense of common sense" (in J. H. Muirhead (ed.) *Contemporary British Philosophy*, London: Allen & Unwin 1925, pp. 193-223) fails to effectively criticise most species of idealism, including the Berkeleian variant,

effectively targeting certain problematic aspects of the British idealism of his predecessors and producing a constructive ontological interpretation of the reductionist position, does not provide any binding reasons for such a wholesale rejection of immaterialism.<sup>152</sup>

The ideological reproduction of post-Cartesian naturalism within the framework of industrial capitalism, especially in relation to its emphasis on abstractions and interchangeable use-value, engenders the reduction of language to representation which comes to full fruition in relation to the positivist movement. As language also must be reduced to the physical, it comes to be understood as simply a *re*-presentation of some particular objective state of affairs, suppressing language's relation to intentionality or subjectivity.<sup>153</sup> In the hypothetical, actual absence of intentionality, however, the success of this representation becomes untenable (how can a meaningless pattern of ink on paper be said to successfully refer to e.g. the city of New York as an actual object or entity?), and the postmodern period's representational monism ensues. Objective representation is retained as the basic function of language, since coherent communication is inconceivable without it, but the object that it cannot really represent (without an effective intentionality connecting subject and object) is metaphysically "bracketed", and the possible truth of any meta-narratives is as a consequence called into question.<sup>154</sup> This preoccupation with representation is paralleled by the functionalist models of consciousness developed in relation to early behaviourism, which later are elaborated into eliminativist models of functional self-representation.<sup>155</sup> They are beset by conceptual difficulties we will soon examine, which are very

---

and the 1939 piece, "Proof of an external world" (*Proceedings of the British Academy* 25, 1939, pp. 273-300) is once again, only effective as a refutation of the absolute species of British idealism with specific regard to its anti-realist entailments.

<sup>152</sup> Russell was also strongly motivated to preserve realism in opposition to absolute idealism. In his 1903 work, *Principles of Mathematics* (New York: Routledge, 2010, p. 4, 379), Russell builds a case against absolute idealism with regard to the incoherence of its non-realist entailments. Otherwise, Russell's work is mainly constructive, in elaborating the positive outlines of reductionist ontology, particularly in the 1921 *Analysis of Mind* (New York: Cosimo Inc., 2004). This work provides the conceptual foundations of his *neutral monism* on the presumption of physicalism. These foundations are later developed into an early variant of eliminative materialism by Gilbert Ryle in his 1949 work *The Concept of Mind* (Chicago: University of Chicago Press 2000), yet in the mid-nineties, David Chalmers interprets the foundations in the direction of panpsychism (*The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press, 1996).

<sup>153</sup> See Ryle, 1949

<sup>154</sup> John Zerzan, "The Catastrophe of Postmodernism", in *The Anarchist Library*, (1990s, exact date unknown) retrieved from <https://theanarchistlibrary.org/library/john-zerzan-the-catastrophe-of-postmodernism.pdf> 2016-11-05

<sup>155</sup> Cf. Jerry Fodor, *Psychological Explanation*, New York: Random House 1968; Thomas Metzinger, *Subjekt Und Selbstmodell: Die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*, Paderborn: Mentis 1993

similar to those affecting post-modern representationalist monism, and likewise rest upon modernity's foundational presumption of naturalism.<sup>156</sup>

In summary, the understanding of matter as devoid of intentionality (formal and final causes) that emerged during the early modern period, and which was finally established by the Cartesian-Newtonian paradigm, entails a strong ontological separation of the mental and the physical. This separation is pregnant with such issues as the mind-body problem and the interaction problem. The ephemeral nature of the mental as compared to the physical, together with the lack of any empirically verifiable point of interaction between them, then prefigures the ontological position of reductive physicalism. This tendency is given further support by a reductively inclined institution of natural science, and its prominent role in Western culture and economy. Following is a summary of the preceding genealogical sketch:

*1. Contemporary reductive physicalism is erected upon an ontology of matter which strips matter of intentionality, and describes it in abstract terms as interchangeable and reducible to strictly objective descriptions*

*2. This view of matter is introduced during the early-modern era, and is consolidated particularly with Descartes and Newton, who apply a mechanistic, reductive and abstract model<sup>157</sup>*

*3. Reductive physicalism is conceptually contingent upon this particular Cartesian-Newtonian ontology of matter, whereas the Aristotelian predecessor could not in principle have underwritten reductive physicalism in its modern form due to its inherent intentionality, and its final and formal causes which infuse non-physical aspects into the material objects themselves*

*4. The arguments for modern reductive physicalism are chiefly generated by the Cartesian-Newtonian ontology of matter, especially with regard to the problem of interaction<sup>158</sup>*

---

<sup>156</sup> Foster 1991, p. 235

<sup>157</sup> The Cartesian-Newtonian ontology of matter can certainly be traced to earlier developments within philosophy, particularly Scotus' and Ockham's critique of Aristotelianism, but the influence of the former two suggests the Cartesian-Newtonian paradigm as a more proper label.

<sup>158</sup> If the form and final cause are non-physical and an integrated part of material reality as on the Aristotelian model, there's nothing conceptually problematic in that the non-physical then can interact with material entities. A body as such is inconceivable without its formal cause, and necessarily has a non-physical teleology in terms of its causal potential. There is therefore

*5. The natural science of the industrial era which presumes the Cartesian-Newtonian ontology, also spreads and reinforces it, due to its focus upon utility and increased control over nature, rather than final causes*

Moreover, there's arguably a problem in that the most important positive arguments for reductive physicalism conceptually presume an ontology of matter which is not self-evident, and rarely is given much in the way of independent philosophical support, yet whose prominence to a certain extent can be explained with regard to a set of institutional arrangements.

## 6.4. Which arguments effectively undermine reductive physicalism?

There are many arguments advanced against reductive physicalism in its different versions. Some of these strike broadly against any and all forms of reductive physicalism, while others are specific to certain models or interpretations thereof. They involve two basic strategies. It is either endeavoured to show how the exclusion of phenomenal subjectivity of reductionist physicalism cannot possibly produce a complete ontology, or that the various forms of the position are incoherent, either necessarily so, or with regard to the models or interpretations that are available. Most comprehensive cases combine the two strategies. We will first look at the important criticisms specific to the major interpretations of reductive physicalism, and then move on to the broad-scope arguments directed against reductive physicalism as such.

### 6.4.1. On eliminative materialism

Beginning with the most radical form of reductive physicalism, eliminative materialism, the position is generally criticized for its purported incoherence. Eliminative materialism maintains that phenomenal consciousness, intentionality or mental states do not actually exist as such in any form whatsoever.<sup>159</sup> This includes beliefs, desires, thoughts, or language-meaning – that the presumed, inherent “aboutness” of a concept or a proposition is entirely illusory.<sup>160</sup> The position is generally arrived at via some form of functionalism

---

nothing *prima facie* problematic with the proposition that the non-physical may interact with material reality, *if* bodies as such already have non-physical aspects *as* material objects.

<sup>159</sup> “Eliminativism”, Craig 2000; William Lycan, (ed.) *Mind and Cognition*, Oxford: Blackwell 1990, p. 236ff

<sup>160</sup> Alex Rosenberg, *The Atheist's Guide to Reality: Enjoying Life Without Illusions*, New York: W. W. Norton & Company 2011

or computationalism that is item-eliminative,<sup>161</sup> as were the behaviourist models which originated with Ryle.

There are two basic types of counter-argument, disregarding those relating to the ethical consequences of the position, however important. One is to the effect that the position can never actually be endorsed without contradiction, while the other maintains that the position can in no way be true according to any coherent epistemology.

In detail, the problem of assertion or endorsement entails that in even raising the issue of whether there is a mental realm or not, we are presupposing our capacity to think, and thereby our own mentality. That is, if one is in a position to actually *consider* eliminative materialism, one must necessarily be conscious of one's own mental states, inferences and beliefs, which directly refutes the thesis as such.<sup>162</sup> If one were to doubt one's own subjectivity, one would as a prerequisite have to be conscious of said doubt, ultimately undermining it as such.

The contention is thus that it's incoherent to actually endorse eliminative materialism, since one can never assert the position without simultaneously implying that it is something one believes to be true, or believes to be plausible, which necessarily involves some form of first-person intentionality. The heart of this particular issue is then whether or not the eliminativist can coherently state his or her position while entirely avoiding any implicit or explicit commitment to the notion of phenomenal intentionality.

The proposed response is to employ some variant of the functionalist understanding of concepts, with the implication that language is nothing over and above neural phenomena, the behaviour of human bodies and their effects on the environment. Yet since successful conceptual reference presupposes some form of intentionality, such a behaviourist analysis of language-meaning must be ruled out. A purely causal understanding of a concept renders it unable to uniquely and meaningfully refer. If a concept, reduced to a set of behavioural causal functions, is to be analyzed in terms of what caused it, and what effects it brings about, there's no reason that it should refer to any specific link in the causal chain or any particular effect, unless intentionality is invoked.

---

<sup>161</sup> Gualtiero Piccinini, "The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism", *Philosophy and Phenomenological Research*, vol. 81, no. 2, September 2010

<sup>162</sup> Foster 1991, pp. 18-21



The basic functionalist response – that successful “reference” using e.g. the concept of *dog* can be assessed in terms of how the language-behaviour involved relates to actual canine bodies in a causal context, but that there’s no actual *reference* being made – is thus entirely incoherent, since then there can be no actual grounds for privileging that particular aspect of the causal network we identify as canine bodies.

Another variant of the problem of assertion involves the necessity of successful reference in terms of actually negating mental phenomena. Thus, if eliminativism were true, there would be no actual mental objects or contexts for mental terms to apply to, and thereby, reference cannot be thought of with regard to how a certain type of language-behaviour relates to some particular entity, simply because there are no such entities to relate to.<sup>163</sup> Therefore, even if we disregard the problems of positive functionalist reference, reference in an eliminativist framework is arguably incapable of actually ascribing any kind of ontological status to the mental. I.e. you can’t coherently refer to the *idea* of mental phenomena and claim that this idea is false or misleading.

With regard to the problem of the inherent incoherence of eliminativism, it is argued that many of the fundamental notions necessarily employed by the eliminativist presuppose intentionality in several ways, often quite subtle ones. For one, as Edward Feser argues in *The Last Superstition*, eliminativism with regard to the mind and mental phenomena, maintaining that there are no purposes or meanings of any sort whatever, will inevitably encounter a certain kind of self-referential incoherence. This incoherence can be seen insofar as eliminativism entails a position asserting the truth of a particular ontology, and generally also the falsehood or illusory character of contrary positions.<sup>164</sup> Yet truth, falsehood or illusion, as Feser contends, are normative concepts which presuppose meaning and representation. Falsehood presupposes the *meaning* of a statement or proposition that has failed to accurately *represent* some state of affairs – but on the assumption that there are no purposes or meanings, an inevitable consequence of eliminativism, then neither can there be any illusions or falsehoods. The implication is then, of course, in which sense eliminativism as a position can be said to give us a true description of reality in any coherent sense, i.e. on any theory of truth that is in principle capable of conceding eliminativism some measure of normativity.

---

<sup>163</sup> Ibid. p. 21

<sup>164</sup> Edward Feser, *The Last Superstition*, South Bend: St. Augustine Press 2008, ch. 6

The wholesale elimination of meaning inherent to eliminativism arguably also strikes against the entailment relations of basic logic, since these presuppose the successful reference and meaning of symbols geared towards expressing and attaining truth, without which any and all arguments that ostensibly would support eliminativism as a position cannot possibly be valid.

At this point, an eliminativist defense strategy has sometimes been found in reference to an appropriate deflationary account of truth, such as the disquotational or redundancy theories, since these arguably afford the eliminativist the possibility of explicating truth and reference without at all deferring to psychology.<sup>165</sup> However, even on the particular interpretation of the deflationary accounts often attributed to Ayer, such as the disquotational or redundancy theories of truth, that regards the deflationary accounts as “disappearance” theories, our intuitive understanding of truth as in some basic sense related to meaning and reference cannot coherently be eliminated. This is due to the very same reason as related above – the normativity inherent to truth-claims, or affirmations of truths, is still indispensable. If truth, on the contrary, is not in some sense a normative property that our statements can fail or succeed in having, however weakly defined, e.g. in a culturally relativist, constructionist sense, this very epistemological position as such cannot be stated and endorsed. Such a radically nihilist interpretation of the deflationary theory entails the self-refuting position that no statements are true in any sense, including the very statement asserting the position as such. This is entirely different from even the most radical forms of epistemic relativism, which even on the denial of any stable, objective truths of any kind whatever, never actually mean to deny that propositions can be true in the context-dependent sense, and thus succeed in referring.

Moreover, if the eliminativist, disregarding this incoherence, nonetheless opts for some variant of the “disappearance-theory” of truth, the entire eliminativist programme based in scientific realism seems to be imperilled as well – how could possibly realism be asserted when no statements can actually affirm anything? In other words, there is a significant dilemma here. If the eliminativist in wanting to save realism instead chooses to retain something like the classical correspondence theory of truth, it will be argued that intentionality has not at all been eliminated, given the paradigmatic role of reference as an intentional notion, and its indispensability for non-disappearance-theories.

---

<sup>165</sup> Hilary Putnam, *Representation and Reality*, MIT Press 1991, p. 65-70

At this point, one may perhaps take a few steps back, and ask the fundamental question of whether eliminativism really entails this apparently contradictory denial of all linguistic meaning as such. Unfortunately for the eliminativist, this entailment seems to be inevitable. The eliminativist of the Churchland-Stitch type, assuming scientific realism, will for instance regard such things as beliefs or phenomenal experiences as illusory or non-existent since they arguably have no objectively available material attributes between them that enable a generalizing definition. Thus, since for instance, one's phenomenally present experience of pain is said to have no exact material properties that are scientifically describable (e.g. with regard to neurobiology) which are held in common with other such purported instances of pain, the notion of pain as such is simply a fiction (or an instance of folk-mythology).<sup>166</sup>

If we then move on to consider a sentence written on a paper, the eliminativist cannot, along similar lines, accept it as expressing a proposition carrying some meaning or other, since no meaning can be found in the material objects as such, in the ink on the paper and their concrete physical constituents. The sentence on the paper is taken to be only a link in a causal chain which manifests in the behaviour of human bodies, neither more nor less.<sup>167</sup> Thus on the eliminativist account, assuming also the negation of subjectivity and agency, there can be no actual intentionality, and thereby neither any real desires or even purposes expressed by such a "sentence", and hence, no linguistic meaning. The sentence "The cat sat on the mat" thus means absolutely nothing – it merely causes certain behaviours within the framework of a particular context.

As a further debilitating problem of eliminativism, Hilary Putnam here presses the question of what commonality then remains between separate phenomena that depend upon intentionality to be classified as belonging to the same category.<sup>168</sup> This issue seems to arise for abstract as well as concrete objects – if there actually are no purposes, the definitions of everyday objects such as cutlery or drinking utensils become incoherent, and the general categories as such collapse. What's worse for the eliminativist, the commonality between purported instances of, for the eliminativist, extremely important categories such as *causes*, also disappears. In the case of causes as actual events, there are namely no scientifically identifiable attributes that can provide the basis for a complete definitional generalization and identifi-

---

<sup>166</sup> Ramsey, William, "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2016/entries/materialism-eliminative/> 2016-10-11

<sup>167</sup> Rosenberg 2011, p. 179

<sup>168</sup> Putnam 1991, p. 58-64

cation, unless one, contra Hume, brings on board actual purposes, intentions or the teleology of e.g. an Aristotelian metaphysics. See chapter 7 for a more extensive discussion of theories of causation.

### 6.4.2. On non-eliminative reductive physicalism

Since eliminativism for the preceding reasons has generally been considered an untenable position, proponents of reductive physicalism have often preferred the item-conservative route, such as functionalist models within the framework of non-eliminative identity theory.

The fundamental conceptual problem of non-eliminative reductive physicalism regards whether it can be meaningfully and consistently distinguished from both the eliminativist position, as well as from those positions acknowledging the reality and ontological distinctiveness of the phenomenal qua phenomenal. Three basic strategies of navigating this dilemma have historically been employed. These strategies all entail some positing of a hidden nature or ontological bracketing of the mental, insofar as they do not either collapse into phenomenalism (e.g. property dualism) or eliminativism. We will discuss these strategies in order.

#### 6.4.2.1. Conceptual reductionism

The first strategy attempts a *conceptual reductionism*, meaning that all our talk of the mental can be translated to statements, generally causal-functional, regarding objective physical reality without any loss of information. At the same time, judgment is suspended regarding the ontological status of the mental, so as to avoid eliminativism and the categorical negation of any truth-claims involving mental attributes. A successful reduction along these lines would entail that all of our mental language is entirely redundant, and thus that any ontologies dependent thereupon would be suspect. Thus, there *seem* to be subjective phenomena, but all we can say of these ultimately reduces to talk of physical things, yet they nonetheless have a kind of secret nature that we cannot access. One prominent example of this position is Putnam's early development of classical behaviourism known as analytical functionalism.<sup>169</sup>

The relation to Russelian neutral monism with regard to the hidden nature-strategy is obvious. An intrinsic nature of the "mental object" which is ontologically neutral (i.e. neither mental nor physical) is posited or implied to avoid the concession to phenomenalism as well as the explicit collapse into

---

<sup>169</sup> Hilary Putnam, *Mind, Language and Reality. Philosophical Papers, vol. 2*. Cambridge: Cambridge University Press, 1975

eliminativism. We will later discuss this bracketing-strategy in its pure form, but at this point, two basic responses which undermine it can be stated:

1. if the *inherent nature* of the “mental object” can play the basic role of the phenomenal as we normally experience and conceive of it, the position is for all intents and purposes on par with phenomenalism (including at least the weak variants such as non-reductive physicalism or property dualism). If not, it collapses into an equivalent of eliminative materialism, whatever the inherent metaphysics.

2. to avoid eliminativism, it must be conceded that we seem to experience something, and that the nature of this *seeming* is qualitatively different from that of physical objects (defined in the post-Cartesian sense). If this is conceded, however, it's not possible to doubt that there ever really was a *seeming*, since every such representation is in itself a manifestation of intentionality. Then we have a qualitative, metaphysical gap between the intentional seeming-events and the objective physical states of affairs which precludes exhaustive reduction. This gap, which will be explored more extensively in relation to the broad-scope arguments, is also generally the rationale behind the assumption of positions such as property dualism or non-reductive materialism, as with regard to Jaegwon Kim's relatively recent concession to the former position.<sup>170</sup>

#### **6.4.2.2. The identification-problem of analytical reductionism**

Aside from these basic reservations, there are inherent difficulties with analytical reductionism.

The most advanced forms of analytical reductionism are formulated in relation to the type-identity theory. This position entails that all mental *types* such as pain, are identical with a particular physical type. This is in terms of pain canonically suggested to be group C nerve fiber activity, but the particular example is not very relevant. Type-identity theory is generally taken to be insufficient as a reductionist strategy with regard to the *multiple realizability thesis*, which incidentally simultaneously undergirds the bracketing-strategy previously mentioned.<sup>171</sup> The multiple realizability thesis, which will be discussed in detail later, maintains that mental events can possibly be realized in relation to a multitude of physical substrates. This precludes exclusive type-identification. In other words, the possibility that different neu-

---

<sup>170</sup> Jaegwon Kim, *Physicalism, or Something Near Enough*, New York: Princeton University Press 2005, p. 29

<sup>171</sup> R. Casati, "Hillary Putnam" in Gianni Vattimo (ed.) *Enciclopedia Garzanti della Filosofia*, Milan: Garzanti Editori 2004

ral states play the same functional roles in different individuals or different species, implies that there can be no intrinsic physical type that *uniquely* plays the particular functional role that is identified with a particular mental type.

A suggested response to this problem of identification is to broaden the functional type into more generic, encompassing types such that the concepts of our folk-psychology identifies a generic type which includes *all* intrinsic types that actually fulfill the specific functional role. This, however, removes the internal unity of the intrinsic type that is necessary for satisfactory and complete identification – i.e. in a situation where different intrinsic types are assigned to a relevant functional role, we cannot identify a particular and unique type-state with our ordinary mental concepts. What's more significant, if the same functional states can be assigned to different types, it necessarily follows that different functional states can also be assigned to the same types without contradiction. Thus, if e.g. *pain*, identifying a type, does not exclusively indicate a particular function, then neither can we consistently successfully distinguish types with regard to such a function.<sup>172</sup>

#### 6.4.2.3. Non-eliminative metaphysical reductionism

The second strategy attempts a *metaphysical identification* that nonetheless preserves the mental items, and is generally exemplified by certain forms of token-identity theory that allow for the distinctiveness of the mental. Token-identity theory entails that for every actual particular (object, event or process)  $x$ , there is some physical particular  $y$  such that  $x = y$ .<sup>173</sup> However, there is even less space for the distinctiveness of the mental on this account, and the agnostic response is not available insofar as the position entails an explicit metaphysical identification.

In detail: precluding analytically reductive yet ostensibly item-conservative analyses such as type-functionalism which would allow us to deduce mental facts from purely objective descriptions and indirectly imply token-identity, there's little obvious support for the reductionist token-constitution theory as such. If we thus reject the deductive route from physical to mental truths and accept the problems of analytical reductionism as decisive, then there's little basis to argue that our knowledge of physical facts can exhaustively reveal all the mental facts which they purportedly sustain.

---

<sup>172</sup> Foster 1991, pp. 112-120

<sup>173</sup> Stoljar, Daniel, "Physicalism", *The Stanford Encyclopedia of Philosophy*, Spring 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/spr2016/entries/physicalism/> 2016-11-08

Yet in simultaneously rejecting eliminative materialism, the metaphysical reductionist must still accept propositions involving mental claims as true, and claims of mental events as factual. This implies that the mental is metaphysically independent, unless physicalist reductionism is assumed and taken to entail the derivative and fully reductive nature of the mental. But then, importantly, because physicalism is taken to be valid, the reductionist could never coherently maintain that there would be an objective mistake in rejecting the supposed mental facts altogether. The implication is thus that metaphysical reductionism must presuppose the *possibility* of eliminative materialism, and merely opts for a softer interpretation of the situation in terms of ostensibly conceding some space for subjective phenomena. However, due to the necessity of this presupposition, all the arguments which rendered eliminative materialism untenable, also becomes effective against metaphysical reductionism.<sup>174</sup>

Any strategy of metaphysical identification such as the token-identity theory will therefore have an exceedingly difficult time differentiating itself from eliminative materialism, unless, precisely, it actually acknowledges either non-reductive materialism, or some form of dualism like Donald Davidson's model. Explicit metaphysical identification, again, also precludes the path of obfuscation arguably available to the type-identity theorist as previously discussed, i.e. the bracketing of the ontology of the mental in parallel with Russelian neutral monism.

One prominent formulation of the token-identity theory is Davidson's *anomalous monism*, which combines token-identity theory with non-elimination and as a consequence maintains that mental events cannot be predicted by the laws that govern basic physics, yet that they are ontologically (substantially) physical.<sup>175</sup> The end result of this kind of strategy is always, however, some equivalent of property dualism or non-reductive materialism. It can accordingly never be exhaustively reductive, as long as it concedes that there are genuine mental events which cannot be fully described or predicted with the same conceptual and analytical tools used to describe and predict physical, third-person phenomena.

#### **6.4.2.4. Pure agnostic functionalism**

Finally, it may be argued that one could formulate a version of agnostic functionalism that is non-eliminative and fully reductive, yet not analytically reductive in the mode of type-token identity, nor anchored in token identity

---

<sup>174</sup> Foster 1991. pp. 145-153

<sup>175</sup> Cf. "The Material Mind", in Donald Davidson, *Essays on Actions and Events*, Oxford: Clarendon Press 1980

theory. Thus it would not be sensitive to the problem of identification, nor have to concede anything to dualism in the manner of Davidson's theory. It could then without issue accommodate the multiple realizability thesis, and maintain it as a basis for the bracketing strategy. Positions equivalent to agnostic functionalism has been a popular response among contemporary proponents of physicalism, since they arguably afford a kind of back-door defense of reductive physicalism.<sup>176</sup>

More exactly, functionalism, whether or not analytically reductive, entails that the mental is to be defined in terms of its functional causal relations, just like a spoon is intentionally defined in terms of its function rather any particular, exact configuration of any material substrate, such as the physical element of gold. This means that the mind can be realized in relation to a multitude of substrates, which does not in principle rule out either e.g. Cartesian dualism or reductive physicalism. This is the multiple realizability thesis previously mentioned.<sup>177</sup> However, with the background assumption of naturalism as a fundamental aspect of the ideological conditions of modernity as discussed in the introduction, as well as the premise that Cartesian dualism together with non-reductive materialism fails due to Ockham-related principles, and that idealism is untenable, the truth of physicalism is generally argued to be the reasonable assumption.

Even if we were to concede all of this, it's not at all clear how such a version of agnostic functionalism could avoid the dilemma of collapsing into either eliminativism or phenomenism as discussed under 6.4.2.1. Moreover, it seems impossible to define mental items in terms of *functions* without considering them as instances of types in some sense, and still succeeding in maintaining the metaphysical commonality between different mental items of e.g. the pain-kind – unless, of course, the *phenomenal* distinctions are presumed.

Any form of agnostic functionalism will also, together with the two other strategies, face certain important difficulties with regard to the more comprehensive arguments against reductive physicalism we will now address.

### 6.4.3. Broad-scope arguments

The broad-scope arguments all defer to the phenomenal or the intentional as present to the subject. The most well-known is probably the qualia-type, popularized by Jackson's knowledge argument as well as Chalmers' zom-

---

<sup>176</sup> Edward Feser, *Philosophy of Mind*, Oxford: Oneworld 2006, pp. 69 ff

<sup>177</sup> Casati 2004



bies, and variants of these.<sup>178</sup> We will explore two general, main lines of reasoning in terms of broad-scope arguments.

The first one argues towards the indispensability of the phenomenal in our understanding of the physical, and is related to Berkeley's main argument for idealism.<sup>179</sup> Arguments to the effect that e.g. functional definitions necessarily presuppose intentions are taken to fall into this category.

The second line of reasoning, which we will categorize as incommensurability arguments, attempts to establish the metaphysical irreducibility of the phenomenal with regard to the introspective immediacy of subjective states of affairs, which are in principle impossible to even describe in terms of third-person language regarding objective physical reality. In other words, the qualitative distinction, immediately available from the subject's point of view, is argued to entail the incommensurability of the physical and the mental. Here we find the undead and the night creatures – conceivability-arguments such as Chalmers's zombies fall into this category, together with Nagel's and Sprigge's "what is it like to be a bat?"-argument.<sup>180</sup>

#### 6.4.4. Indispensability arguments

These arguments attempt to establish the indispensability of the phenomenal when it comes to defining, accessing or understanding other aspects of reality (generally the physical level). Given this, there are two entailments one may possibly argue towards: 1. That the physical as such is irreducibly mental or phenomenal, either entirely (e.g. as on idealism) or in some part (as, for instance, on the non-reductive materialisms; hylemorphic dualism; dual-aspect theories, or integrative dualism). 2. That the mental is necessarily distinct from, and non-identical with the physical, which opens for all forms of dualism and non-reductive materialism.

The basic indispensability argument can be formulated on Berkeley's terms; that any definition of the physical in line with post-Cartesian materialism nonetheless necessarily rests upon an assumption of subjective phenomenal presence. Berkeley argued that (1) we have no grounds to assume that what we take to be physical objects actually are anything over and above what we

---

<sup>178</sup> Chalmers, 1996; Frank Jackson, "Epiphenomenal Qualia", *Philosophical Quarterly* 32, 1982

<sup>179</sup> This position is first outlined in Berkeley's *An Essay Towards a New Theory of Vision*, published in 1709

<sup>180</sup> The bat-argument was popularized by Thomas Nagel (Thomas Nagel, "What is it like to be a bat", *The Philosophical Review*, vol. 83, no. 4, 1974), yet was independently suggested by Timothy Sprigge some years prior (Sprigge 1971)

“perceive by sense”; that (2) we cannot perceive anything except what is phenomenally present to ourselves, and finally and most controversially, that (3) we cannot assume that what is phenomenally present to us can exist as such without actually being perceived.<sup>181</sup> Berkeley, like many later proponents of the position, then made a case for phenomenalist idealism upon these grounds.

#### 6.4.5. The physical’s phenomenal dependency

A much more modest version of Berkeley’s kind of argument employs nothing but a modified variant of (1) in the last paragraph, maintaining that the physical cannot be understood or accessed without deferring to the phenomenal, which in turn can be argued to indicate that any purportedly existing objective physical realm cannot be identical to the phenomenal.

Charles Taliaferro provides an example of this very kind of argument in his *Consciousness and the Mind of God*. In establishing the argument, Taliaferro maintains that in characterizing the material world, one must make use of ostensive definitions, that is, definitions which “identifies what is manifested or disclosed in experience”.<sup>182</sup>

Taliaferro combines this position with a criterion of warrant for non-identity inspired by Leibniz’ principle of the identity of indiscernibles. Taliaferro’s criterion states, in an inverse manner, that if one has warrant to believe that one can *conceive* of A without at the same time conceiving of B, one also has warrant to suppose that the two are not identical.<sup>183</sup> This seems uncontroversial – if one actually had warrant to believe that one could conceive of the actual person referred to by the name of Voltaire without at the same time conceiving of the person named François-Marie Arouet, then one would also have warrant to believe that these two are not identical.

If this necessity of ostension which Taliaferro maintains holds true, it then follows that any and all characterizations of the objective physical level of reality must presuppose the phenomenal and its independent epistemic relevance, which arguably entails a minimal phenomenal realism. Thus assuming the necessity of ostension would mean that the phenomenal realm can independently and uniquely supply true propositions regarding ontological facts on the objective physical level of reality. In effect, one would therefore

---

<sup>181</sup> George Berkeley, *Principles of Human Knowledge and Three Dialogues*, Oxford University Press, New York: 1996, p. 25-30

<sup>182</sup> Taliaferro 1994, p. 91

<sup>183</sup> Ibid. p. 49-52

have to concede that one needs to assume the existence of A, that is, the phenomenal realm, in order to actually characterize, that is, to accurately, with true propositions, describe B, objective physical reality. The issue pivots around exactly *what* one, if ostension is necessary, would need to concede regarding the phenomenal realm.

So if our knowledge of B, that is, if true propositions regarding objective physical reality are only accessible through A, this entails that A is a unique source of actual data, of actual knowledge. If A is such a unique source, it necessarily follows that there are propositions which are uniquely true of A, and thus do not accurately describe anything else, including B. This would satisfy Taliaferro's criterion, entailing the non-identity of A and B, on the relatively unproblematic assumption that the true propositions in question are necessarily linked to actual properties. However, it would also suffice to establish non-identity according to the direct inverse entailment of Leibniz' principle of the identity of indiscernibles, even in its weakest formulation, which denies that distinct *states of affairs* ever exactly resemble.<sup>184</sup>

How and where does then this problem arise for the physicalist? On the Cartesian-Newtonian view of matter, it appears quite instantly. If one for instance defines what is physical in terms of those seemingly material objects which are phenomenally present to one's senses, the phenomenal and intentional is obviously presupposed. Such a definition is thus incoherent as it in this presupposition is in direct conflict with the modern view of matter as devoid of intentionality, as well as open to the Berkeleyan critique.

The physicalist must therefore provide a more occult definition of the physical, a definition of a more abstract and objective character which is not as readily vulnerable to these kinds of contradictions. Accordingly, performative and more flexible definitions of the physical such that "whatever physics claims exists", has been favoured, especially in the light of the revolutions in the field during the past century, which indeed demolished much of the Newtonian paradigm and fundamentally changed our understanding of matter as such.<sup>185</sup> Such a definition is, however, possibly too flexible to ground a post-Cartesian reductive physicalism, and is arguably circular. Moreover, in the final analysis, it also defers to the phenomenal or the intentional, as we shall soon see.

---

<sup>184</sup> Peter Forrest, "The Identity of Indiscernibles", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/fall2008/entries/identity-indiscernible/> 2016-11-14

<sup>185</sup> This definition was explicitly suggested by J. J. C. Smart in his article "Materialism", in *Journal of Philosophy*, 1963

We can actually support this central premise analytically, i.e. the premise that our access to the physical presupposes a deference to the intentional or the phenomenal, by simply examining the post-Cartesian understanding of the material. Since it is by definition devoid of intentionality, there's a qualitative difference between the thing in itself and the phenomenal experiences associated with or related to it (mirrored in the Kantian distinction between the phenomenon and the noumenon), which implies that the physical can never by itself be phenomenally manifest or present *as physical*. This inability is inherent to the definition.

Looking back to Taliaferro's case, he accordingly examines all conceivable types of definition of the physical realm, and finds that they are either circular and non-informative, or in some sense presuppose ostension, that is, a deference to the intentional or the phenomenal.<sup>186</sup> This contention is thus related to, but independent of Hempel's dilemma, which states that a definition of the physical in terms of what current physics considers to be real will most probably turn out to be false, yet a deference to future physics is open to the possibility that such a physics would actually consider the mental *qua* mental to be a part of the ontology.<sup>187</sup>

However, Taliaferro seems to omit the fact that physics as a discipline must also be said to imply ostension, to the extent that the very discipline as such, notwithstanding the nature of the entities it posits, can be argued to presuppose intentionality in exactly the same sense. Physics as a practical tradition of knowledge is namely inconceivable without intentions, without language-meaning, without empirical knowledge and so on, and such concepts as "empirical" and "knowledge", are inseparable from the intentional and the phenomenal in turn, which can be used to ground an indispensability argument on the meta-level.

The circular definitions (those that define the physical in terms of what physics describes) are thus not only subject to Hempel's dilemma, but also implicitly yet inevitably defer to ostension since they depend upon reference to, for instance, the discipline of physics as such, which presupposes the epistemic relevance of intentional states.

Taliaferro further argues that the only way out of Hempel's dilemma from the point of view of the circular definitions is to actually beg the question in

---

<sup>186</sup> Taliaferro 1994, pp. 90-105

<sup>187</sup> Carl Gustav Hempel, "Reduction: Ontological and linguistic facets", in S. Morgenbesser, P. Suppes, and M. White, *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, New York: St. Martins Press 1969 pp. 179-199

favour of physicalism (i.e. to presume that the only entities a future physics will be concerned with must be material in the post-Cartesian sense), yet this would still not address the ostension implicit to physics as a discipline, in terms of a practised tradition of knowledge.

Andrew Melnyk has in defense of physicalism reasoned in exactly this manner, and maintained that physicalists ought to navigate Hempel's dilemma precisely by formulating as a criterion on any purportedly satisfactory definition of physicalism, that it must "possess content that is at this current time determinable by us", meaning that the definition cannot refer to any entities beyond what current science can describe, so as to avoid any future physics that admits ghosts into its machine.<sup>188</sup> The implication is then that such a definition of physicalism is unlikely to actually be true, since natural science will probably come to refer to entities beyond what current science can describe. Melnyk, however, defends the stance by maintaining that one does not need to actually believe that physicalism is even likely to be true, to be a proponent of physicalism, so long as one has warrant that physicalism is *closest* to the truth when compared to its alternatives.<sup>189</sup>

This position is unfortunately incoherent, since reductive physicalism is a binary proposition – if reductive physicalism is false, then actual non-physical entities, properties or substances necessarily exist. Reductive physicalism could thus not be "closer to truth" than the negation of reductive physicalism, yet nevertheless false, since in that case the negation of reductive physicalism would actually hold. So if physicalism is unlikely to be true, and you positively know this, you cannot in principle at the same time have warrant that physicalism is closer to the truth than its negation. In other words, the law of excluded middle is here actualized. If the proposition instead were one of e.g. the utilitarian value of a political system, rather than one of metaphysical truths, Melnyk's defense could be reasonably put forth, whereby one for instance could argue that a certain system is closest to realize a relevant set of goods compared to all others, although not yet perfected.

The outcome is thus that the circular definitions of physicalism of such formulations not only beg the question - they beg the question in favour of a position that is not even likely to be true. What's even worse, they also exhibit a vicious circularity. As Melnyk further observes, some equivalent of the criterion that the definition of physicalism cannot refer to any entities beyond what contemporary science can describe, is actually necessary for us to be able to infer the truth of physicalism from the findings of contemporary

---

<sup>188</sup> Andrew Melnyk, *A Physicalist Manifesto*, New York: Cambridge 2003, p. 11-12

<sup>189</sup> Ibid. 225-226

science.<sup>190</sup> So we have a vicious circle, which one arrives at by begging the question in favour of a position which is probably false, and that due to its binary, either-or character cannot possibly be warranted by virtue of it being somehow “closer to truth” than the alternatives.

The possible non-circular definitions Taliaferro explores all explicitly defer to intentionality and the phenomenal, since they presuppose such states of affairs or events as “knowledge” or “observation”. This is, again, really inevitable, in that any definition of the physical aligned with post-Cartesian materialism, yet that does not refer to the physical sciences, must necessarily be anchored in knowledge attained by some form of empirical observation, insofar as the truth of post-Cartesian physicalist monism cannot be shown to be an analytical truth derived from premises independent of such empirical observation. And on the other hand, if the definition does refer to the physical sciences, they in turn presuppose the very same kinds of intentional states of affairs, in addition to the problems previously discussed.

#### 6.4.6. Functions and intentionality

Another broad-scope critique of reductive physicalism can be based in the indispensability of intentions for any equivalent of functionalism or computationalism, that is, any analysis of the mental that attempts to reductively define it in terms of the activity of something akin to a computer program or as a rule-governed function. The classical computationalist theory of the mind, first introduced by Putnam in the late 60s, is a development of classical behaviourist functionalism, but can be distinguished from this position mainly due to the behaviourist version’s focus on strict causative relationships.<sup>191</sup> Computationalism is a variant of functionalism with a conceptual structure similar to the latter’s, yet arguably not with all of the metaphysical commitments of the behaviourist predecessor.<sup>192</sup>

Since functionalism in its different forms is the go-to method for contemporary reductive physicalism (or the neutral monist backdoor defense), and as there really is no other way to conceive of consciousness outside of eliminativism or phenomenalism (the mind must outside of these positions then in some sense *be* what the brain *does*), a critique of the conceptual foundations of functionalism or computationalism strikes at the heart of non-eliminative reductive physicalism as such. This type of critique is related to the family of

---

<sup>190</sup> Ibid. p. 12

<sup>191</sup> Hilary Putnam, “Psychophysical Predicates”, in W. Capitan and D. Merrill (eds.), *Art, Mind, and Religion*, Pittsburgh: University of Pittsburgh Press 1967

<sup>192</sup> Piccinini 2010

arguments against materialism generally known as “the argument from reason”.<sup>193</sup>

The basic argument is relatively simple, and intends to show that anything akin to a function on computationalist or functionalist theories of the mind, must in the final analysis presuppose or tacitly introduce intentionality. An early form of this critique is first found with Karl Popper, who targeted both causal functionalism and computationalism,<sup>194</sup> but variants of it have been proposed by many philosophers, such as John Searle.<sup>195</sup>

In *The Self and Its Brain*, Popper argues that the *descriptive* and the *argumentative* forms of language cannot be reduced to causal functions, since the descriptive form makes a truth-claim, and the argumentative form makes an inference from truth-claims. Both inferences and truth-claims are namely rooted in intentionality (cf. 6.4.1. above) since no objective, physical state of affairs can unambiguously refer to anything.<sup>196</sup> In Popper’s discussion of the argumentative role of language, there’s also an implicit objection to computational theories as such. Popper argues that the reason that a computer operates according to logical principles is that those have been intentionally imparted to it, that the properties which makes it work *as* a computer cannot be physical properties.<sup>197</sup> Searle argues along similar lines in his later development of the critique: computation is equivalent to the manipulation of symbols according to rules of syntax. But syntax and symbols cannot be defined in terms of the physics of a system, and therefore, computation cannot be intrinsic to the physics of a system, and must thus be intentionally imparted to it if it at all takes place. Therefore, the physical brain cannot coherently be said to be intrinsically an equivalent of a digital computer.<sup>198</sup>

We will examine a more fleshed out version of this line of reasoning which targets reductive materialism as such. This argument is originally Saul Krip-

---

<sup>193</sup> Victor Reppert, “The Argument from Reason” in William Lane Craig and J. P. Moreland (eds.), *The Blackwell Companion to Natural Theology*, Oxford: Blackwell Publishing Ltd. 2009

<sup>194</sup> For Popper’s critique of causal functionalism, see Edward Feser, “Hayek, Popper, and the Causal Theory of the Mind”, in Leslie Marsh (ed.) *Hayek in Mind: Hayek’s Philosophical Psychology (Advances in Austrian Economics, Volume 15)* Emerald Group Publishing Limited 2011, pp.73 – 102. The outline of Popper’s similar argument against computationalism is found in Karl Popper, John Eccles, *The Self and Its Brain*, New York: Springer Verlag 1977, p. 78 ff

<sup>195</sup> John Searle, “Minds, Brains and Programs”, *Behavioral and Brain Sciences*, 3 1980

<sup>196</sup> Popper & Eccles, 1977, pp. 57-60

<sup>197</sup> Ibid. pp. 75-81

<sup>198</sup> John Searle, “Is the Brain a Digital Computer?”, *Proceedings and Addresses of the American Philosophical Association*, vol. 64, no. 3 1990

ke's, and an outline is found in his *Wittgenstein on Rules and Private Language*.<sup>199</sup> The argument turns on the contention that any definition of the function of something like a computer, the question of what kind of program it's running, necessarily involves idealization. If the software for instance would malfunction, and fail to perform the intended operations, there's nothing whatsoever in the physics or the operations of the machine as such that can definitely establish that is actually *not* doing what it's supposed to do, unless intentions are presupposed.

An exposition of Kripke's argument provided by Jeff Buechner in the article "Not Even Computing Machines Can Follow Rules" illustrates Kripke's point with the use of the latter's "quus"-paradox. This paradox entails that the logical operation you perform when you think that you perform addition, could, if we completely disregard your intentionality, possibly be an entirely different operation in reality. This is due to the fact that the idealized functions of an abstract mathematical automaton (or the human brain considered as such) are always only imperfectly realized in physical computing machines, since these are not actually infinite, and can possibly malfunction. Thus, an external observer can justifiably infer that the same machine is instantiating entirely different idealized functions or programs with regard to the same finite empirical data regarding its physical state.

This different function or program could for instance work in such a way that it produces the exact same results as addition - up until one of the addends supercede a certain given number, above which it instead produces "5" as a result. Kripke names this different type of operation "quaddition", and since the limit of the addend beyond which quaddition is actualized can be arbitrarily high, the inductive evidence will always be consistent with you having *actually* performed quaddition and not addition, unless, precisely, some intention or other is taken to define the program or the function.<sup>200</sup>

In the same manner, there's no way to determine whether a physical computer is actually performing one or the other of these different operations with reference to its physical features. No matter what output the computer produces, we cannot determine whether it's actually running a program of addition rather than quaddition with regard to nothing but the physical features of the machine itself. Nothing in the behaviour of the computer, considered as such, can tell us whether it's been running a particular idealized

---

<sup>199</sup> Saul Kripke, *Wittgenstein on Rules and Private Language*, Cambridge: Harvard University Press 1984

<sup>200</sup> Jeff Buechner, "Not Even Computing Machines Can Follow Rules", in Alan Berger (ed.), *Saul Kripke*, New York: Cambridge University Press 2011



program for addition, or if it has instead been running a program for quaddition that has malfunctioned.

We can even assume a physically infinite computer running a certain program indefinitely. When the machine finally does cease to operate, if for no other reason than the heat death of the universe, and thus has not fully realized the potentially infinite abstract state diagram of whatever function it's supposed to compute, it is still fully possible to attribute to the machine any of an infinite number of functions entirely compatible with its complete history of operations.

The problem cannot be circumvented by appealing to counterfactuals, claiming that the program has indeed malfunctioned if, say, the result of a particular calculation is "5" and the addends are higher than this number, since this would simply presuppose that the computer was running an idealized program of addition. It can obviously neither be avoided by appealing to the programmer's intention to construct a program of addition in accordance with the appropriate axioms.

But cannot one simply inspect the computer's structure and determine what output it *would* produce in performing certain future calculations, and simply check if there was an upper limit for the value of the addends included in the programming, above which every output will be "5"? In other words, is it not possible to inspect the "causal grooves" in the machine, ascertain that input A will yield result X, input B result Y and so forth, up to the limits of the program, and thus determine that the machine definitely does not perform quaddition?

Here, one must again recall that the machine is finite, while the abstract function is potentially infinite. Kripke's argument turns on precisely this gap, and maintains that you could never in principle read off what the machine is supposed to be doing from the limited, strictly physical facts regarding its operations, without any reference to intentions whatsoever. Absent such intentions, there is an indefinite number of intended, compossible functions that are compatible with the operations of the machine at any particular moment, among which we cannot pick out the right one unless we presuppose design. The identification of the function is underdetermined by the available data.

Thus, one surely could inspect and interpret the machine, and infer that it's not supposed to perform quaddition, but this assumes particular intentions of the programmers or designers. How could one know that the program was

not really supposed to perform idealized quaddition, but has now malfunctioned, or is simply shoddily programmed? You might of course *infer* that this is unlikely given the probable *intentions* of its designers, but it could never be verified in relation to the physics of the computer due to the discrepancy between the potentially infinite idealized function, and the actually finite physical computing machine.

A perhaps more accessible example might be found if we turn to computer games in terms of programs. A canonical example of a bug in a computer game is found in the 1985 NES release Super Mario Bros. At the end of most levels of this game, the player jumps up on a flag pole to capture the flag, which stops the time, finishes the level and initiates the transition to the next one. However, on level 3 in World 3, it is actually possible to jump over the flag without triggering the transition, which enables Mario or Luigi to keep on running alongside a brick wall until the time eventually ends and they die. Now, is this a bug or a feature? We have no idea, unless we make certain presumptions as to the intentions of the programmer. For all we know, this bug might really be a political statement of the programmer Toshihiko Nakago's, implying the futility of the socialist struggle against the state, epitomized by one's royal nemesis Bowser, unless we first dismantle the institution of nationalism (taking down the flag). Whether or not clearing the flag (quaddition) is actually a bug, as opposed to lowering the flag and initiating the transition to the next level (addition), is impossible to determine without presupposing design and intentions. Moreover, if it were actually impossible to clear the flag (or if the program only could perform addition with regard to its limited framework), we could never ascertain whether or not *this* was in itself a bug, and that the program's purpose was not really another.

Appeals to Ockham's Razor, tempting as they may be at this point, is of no help either, since there's no basis to assume that the interpretation of the machine's function in terms of level transition or addition is actually the most parsimonious in comparison to *all* other compossible functions, such as quaddition or the clearing of the flag. The only reason that the assumption of e.g. idealized flag-clearing would be less parsimonious than the assumption of idealized level transition in relation to the physical basis, would be to assume a particular purpose of the machine and/or the programming as the "normal" one.

The only conceivable way to get around Kripke's argument would be if we had a computing machine that could never malfunction and lasted forever, since if it had a limited life-span, the problem of the indefinite number of compossible functions in relation to its history of operations remains. But again, the assumption of normal function as opposed to malfunction still

entails idealization. Even if we were to encounter an actually infinite physical computer that lasted forever and perfectly instantiated a particular, potentially infinite function, it could still not be coherently maintained that this was necessarily and actually the machine's normal function, without any deference to the intentional whatsoever. *The normal function must first be prescribed.* Even if we were to assume the simplest possible computer, performing the simplest possible finite function, the functional ascription as such must presume *this* particular function rather than one of the indefinite set of alternative compossible functions.<sup>201</sup>

This argument targets all forms of functionalism, since all variants of the position necessarily must make the problematic distinction between the states of an abstract causal network and the purely physical states of the brain realizing that causal network, whether or not they can be modeled on a computational structure.

So how does this undermine functionalism? The point is, again, that given the post-Cartesian view of matter, there is nothing in the physical properties of the machine that can tell us what its particular function is. Therefore, non-physical intentions are indispensable for any functional ascriptions as such, and thus, the latter cannot in principle be used to naturalize intentionality on the schema of reductive physicalism.

Yet, one might at this point argue that even if we can never in principle be warranted to say that the brain at a given moment realizes a particular function, we may nevertheless be warranted to claim that the brain indeed realizes *some* function, even if we cannot pick it out among the indefinite number of compossible functions. In other words, we might have no warrant to claim that any particular member of the function-set is realized by the brain, but we can perhaps still claim that some member of the set is realized.

But the warrant for this latter claim that the brain realizes *some* function, and thus that the activity is identical to some particular function, is only entailed by the possibility of actual identification. If we cannot tell what function the machine is performing with regard to its physical properties, there's no telling whether it's actually performing any particular function whatsoever. The position that it surely *must* be performing some idealized function X, that we yet cannot possibly identify, presumes that there necessarily is such a normal function associated with the computer (or brain), a presumption for which there can be no warrant only with regard to its physical constitution.

---

<sup>201</sup> Berger 2011, p. 365

Also, this last counter-claim against Krikpe's argument amounts to a form of type-identity position (the mental token is said to necessarily be an instance of the function-type, broadly considered, and the mental type is thus regarded as identical with the physical function). Therefore it is also subject to the problems discussed under 6.4.2.1. and 6.4.2.2.

#### 6.4.7. Incommensurability arguments

This type of argument works towards establishing the incommensurability of the phenomenal and the physical in principle, due to qualitative differences immediately available to the subject from the first-person perspective. Upon this premise, it is then argued that the mental and the physical are ontologically distinct, and that any sort of complete reduction of the phenomenal to the physical is impossible. This entailment can then be used to ground all forms of dualism and non-reductive materialism, as well idealism, indirectly, if something akin to the indispensability arguments are provided.

Nagel's and Sprigge's bat-argument is probably the most famous of incommensurability argument. It rests on the introspectively immediately available fact of one's own first-person perspective, and provides a definition of consciousness based in this, which amounts to saying that there is something which it is like to be a certain individual at a certain time – there is something which it is like to be a bat, even if we cannot conceive of it, due to the bat being an animal whose perception of the world is probably very unlike our own. This fact is essentially phenomenal and subjective, and can thus not in principle be reduced to any purported physical constituents according to any reductive method which involves ignoring the appearances of an entity, and redefining it in terms of something other than the appearances as such.<sup>202</sup> Arguments related to the qualia problem can generally be regarded as incommensurability arguments, such as Jackson's famous knowledge argument.<sup>203</sup>

We will examine a development of this basic line of reasoning that was recently presented by Lynne Baker, which is arguably related to the somewhat similar knowledge argument.

---

<sup>202</sup> Timothy Sprigge's version of the argument is found in his "Final Causes", in *Proceedings of the Aristotelian Society*, Suppl. Vol. 45, 1971, while Nagel's independent formulation can be found in his "What is it Like to be a Bat", 1974.

<sup>203</sup> Jackson's argument is first presented in his "Epiphenomenal Qualia" from 1982

#### 6.4.8. Baker's argument from phenomenal self-reference

This argument, just like Nagel's and Sprigge's, purports to establish that there is a phenomenological datum, i.e. a fact stemming from direct observation, that is exclusively available from a subject's first-person perspective, which precludes any and all forms of physicalist reductionism. Similarly to the knowledge argument, it is further contended that apprehension of this datum must be considered knowledge, since 1) it cannot even in principle be translated to statements regarding objective physical facts in the third person (where the physical is, again, conceived of as devoid of intentionality, along post-Cartesian lines), and since 2) it uniquely produces objectively available material effects.

If this argument works, it's surprisingly strong. It adds two factors over and above what Jackson's knowledge argument gives us, or that are at best only implicitly presented within that particular framework. Like Jackson's argument, it establishes the ontological independence and irreducibility of the mental, the subjective or the phenomenal, which can be expressed by way of its satisfying the Leibnizian non-identity criterion discussed under 6.4.5.

However, Baker's argument anchors this ontological independence not just in subjectively available, simple qualia, but in two further factors: 1) a change immediately available to the subject itself, which takes place exclusively within the phenomenal first-person perspective and is thus in principle intranslatable to the third person, and 2) the intersubjectively available material effects that this change produces.

Simple qualia, such as objective redness, while implicitly irreducible to the third-person level due to the entailed intentionality, are at least in principle objectively describable, and this, in turn, arguably makes reductive physicalism less implausible. Baker's phenomenological datum, like Sprigge's and Nagel's, is not even in principle objectively describable. Moreover, it is argued that this datum uniquely produces concrete effects which are intersubjectively available. For this reason, the argument not only serves to establish non-identity, but by virtue of the intersubjectively available effects, further undermines any purportedly empirically grounded denialism or eliminativism.

So how is the argument structured? The argument in question analyzes the subjective quality referred to by Nagel and Sprigge in terms of a first-person perspective. This is considered an attribute of human persons which consists in the ability to directly *conceive of oneself as oneself in the first person, without recourse to names, descriptions or other tools of refer-*

*ence from the third-person level of reality.*<sup>204</sup> This, incidentally, also enables us to ascribe this ability to other subjects via ostension, discussed under 6.4.5.

The argument can be illustrated with the following basic kind of thought experiment pertaining to a misidentification of the own subject with an external entity: if the reader imagines her- or himself entering a darkened room through a door leading back to a lit hallway, faintly illuminating the new surroundings while you're searching for the light switch. Suddenly you make out another person standing in a similar doorway on the other side of the room, seemingly also fumbling for the switch on the opposite wall. For a moment you ascribe this other person this position and action, as well as some external, visually present attribute such as his or her wearing a blue jacket, which can be expressed in the sentence A: "He is standing in the doorway, wearing a blue jacket, and is searching for the light switch". Suddenly, however, you realize that you're actually looking at a reflection of yourself in a mirror on the opposite wall of the murky room, generating something like sentence B: "I believe that I\* am standing in the doorway, wearing a blue jacket, searching for the light switch", where the asterisk marks your self-reflexive ability to conceive of yourself as yourself in the first person. The point is precisely that the transition from sentence A to sentence B cannot be expressed with regard to the objective, third-person level of reality. Every attempt to describe sentence B in the third person will produce examples like "X is visible in the mirror, wearing a blue jacket" with X referring to the name of the reader, rather than an equivalent of sentence B, implying that the belief that it is you yourself\* who is visible in the mirror, cannot be conveyed by simply listing attributes in the third person regarding an individual seen in the mirror. The particular change that takes place when you realize that it is really yourself you're seeing in the mirror, expresses the datum previously discussed, i.e. the fact stemming from direct observation that is exclusively available from a subject's first-person perspective, which precludes any and all forms of physicalist reductionism. This datum is expressed by Baker as follows:

(D) There is a difference between conceiving of oneself qua oneself in the first person, and to conceive of someone in the third person who actually is oneself (perhaps unbeknownst to the thinker).<sup>205</sup>

This insufficiency of objective third-person language is further emphasized by the particular thought experiment Baker utilizes, which is similar to the

---

<sup>204</sup> Baker 2013 pp. 32 ff

<sup>205</sup> Ibid. p. 64

one provided above. It regards an individual in a market, pushing a shopping cart carrying a damaged package of sugar whose contents, without the shopper's knowledge, spills on the floor. This is then noted by the clerks. They announce the fact on the speakers, describing the transgressor, who after a moment or two realizes that he or she herself\* is the one referred to, and acts to remedy the situation.<sup>206</sup>

Another way of exemplifying the datum can be made as follows:

Assume that the infamous, yet unidentified serial killer known as Jack the Ripper, was actually the asylum inmate James Kelly, as some recent investigators have suggested. Unless something equivalent to the datum is assumed, we would, when confronted with all objectively available facts of the Ripper's identity, know the *very same things* that James Kelly learns when he recovers from amnesia in the custody of Scotland Yard and realizes that *he himself* indeed is the serial killer in question. Since *we ourselves* would indeed not be (apologies to any undead Victorian vampire-serial killers reading this), any reductive analysis such as functionalism falls short. Moreover, this difference clearly produces markedly divergent results in terms of the measurable behaviour of subjects, even if they're fed with the same objective information. Accordingly, Kelly, having realized that he himself is the killer, will exhibit other behaviour and other attitudes regarding the future or possible or desirable actions relative to his current situation than another person would. For instance, Kelly would perhaps feel that he himself would need to conceal his identity from the authorities and act accordingly with regard to his own safety. This would, in turn, definitely prompt a whole different neurological and limbic response in Kelly, when compared to another person attaining the relevant objective knowledge about the human body referred to with this name - information such that James Kelly realizes that James Kelly is the Ripper; James Kelly is now fearful and desires to conceal himself from the London Metropolitan Police, et c.

At this point, it might be argued that a machine may in principle be programmed to behave exactly as a human person in a similar situation, so that the behavioural response is identical. While this is true, the difference is that in this case, the machine will be programmed to act upon a certain type of input equivalent to objective information. The human person, on the other hand, can have complete, objective information regarding the situation (e.g. objective omniscience with regard to the store in the sugar spill-example), and still not be prompted to act unless her first-person perspective actually becomes conceptually related to the information regarding a particular body

---

<sup>206</sup> Ibid. pp. 49-51

in a particular place; i.e. unless she realizes herself to be situated in relation to the particular body associated with the sugar spill.

A robot might for instance be prompted to act if e.g. the weight of shopping cart #1429 is suddenly reduced, which executes a command to analyze and perhaps replace the cargo of the particular cart, but this will exclusively be a programmed response to new objective information fed to the machine. For the person, the behavioural response is instead exclusively elicited by something which cannot even in principle be described with regard to objective data on the physical level.

If we in spite of Kripke's critique allow for self-referential functionalistic analyses in response to this issue, it nonetheless only pushes the problem a step back. Accordingly, if we go back to the Ripper-example and suggest that the behavioural difference observed in relation to the Kelly-body simply lies in that the Kelly-body now has transitioned into a state wherein it *refers to itself* in terms of the Ripper, this implicitly concedes that first-person self-reference cannot be translated into objective third-person information. Namely, *Our* knowledge that the Kelly-body so does makes no such difference, and neither does actually the Kelly-body's objective knowledge of the very same thing (such-and-such a person refers to himself as Jack the Ripper), unless, precisely, his identity and the first-person self-reference is taken for granted.

In these examples, actual, intersubjectively available changes in the subject's behaviour stemming from nothing but the realization that it is him-\*or herself\* which is the actual object of a particular reference is highlighted, implying that the subject's realization that it was him- or herself being considered all along, while falsely believing that another person was in question, actually produces material effects – that the *specific beliefs* caused by his/her identification of his/her own self-concept in the first person with a particular description, produces intersubjectively available material effects.

The state of affairs that you yourself\* is the very same subject as the one which was previously externally identified by the equivalent of sentence A above, is, again, impossible to adequately translate into third-person language. This is exemplified in that the objective and publicly available change in behaviour produced by the realization expressed by sentence B would never have been brought about by any further factual information regarding the external subject X, not yet realized to be yourself.



The gist of the argument against attempts at physicalist reductionism is then that these cannot affirm (D) without assuming the robust FPP as a premise, which supposedly undercuts the reductionist's position in that one cannot disregard (D) without failing to acknowledge important features of reality, nor affirm (D) without attesting to ontologically subjective facts which are essentially irreducible to the objective third-person level of reality. This again emphasizes the non-eliminative reductive physicalist's dilemma:

(D), or the equivalent thereof, can namely not be denied without implying eliminativism, yet it can obviously neither be affirmed without acknowledging the non-identity of the mental and the physical and thus the former's effective irreducibility.

The only way to deny (D) is to show that the intentional can be effectively reduced to the objective physical level without remainder, otherwise there will necessarily be a difference between conceiving of oneself *qua* oneself in the first person, and objectively conceiving of a person in the third person who actually happens to be oneself (perhaps unbeknownst to the thinker). The ontological bracketing-strategy of non-eliminative reductive physicalism therefore does not work as a response to this argument, since unless reductive metaphysical identity is maintained, the distinctiveness of the mental item implies non-identity via (D). If the mental item actually exists *qua* mental, and the phenomenological datum is necessarily entailed by the first-person nature of anything mental, the mental item cannot be conserved without also implying the datum (D).

The item-conservative route is thereby denied with regard to the introspectively available phenomenological datum, and one important benefit of Baker's argument is precisely that it clarifies this, which has otherwise only been alluded to or implied in other incommensurability arguments, such as John Foster's problem of constitution. Foster's argument entails that the non-eliminative reductionist in rejecting eliminative materialism, must necessarily accept the *sui generis* character of the mental facts, and that, on this assumption, it becomes untenable to claim that the mental facts are non-mentalistically constituted, since first-person phenomenal events must necessarily be constituted in terms of *how they are constituted for, and appear to a subject*.<sup>207</sup> Baker's argument elaborates on this latter claim by way of the phenomenological datum, which cannot in principle be reconciled with item-conservative reductionism.

---

<sup>207</sup> Foster 1991, p. 155

The argument covers all non-eliminative strategies discussed under 6.4.2., and any functionalist response to the argument is subject to all the same limitations. This includes recent, advanced developments of the position such as, for instance, David Rosenthal's HOT-theory or Uriah Kriegel's or Thomas Metzinger's self-representational theories of consciousness.<sup>208</sup> Insofar as they are developed on a basis of functionalist premises, they provide no qualitatively different option in response to this form of the incommensurability argument than those we have already discussed, notwithstanding all their other merits and broader relevance.

## 6.5. Conclusions

Given the two main lines of argument explored in this chapter, together with the untenability of eliminative materialism as such, as well as the problems associated with conceptual reductionism, the dilemma of reductive physicalism appears fatal.

To begin with, if the position does indeed imply eliminativism, it must overcome the many decisive problems of incoherence and conceptual self-refutation we've referred to, including the important issue of whether the subject actually can be rational in affirming a theory that negates the subject's own existence. On the other hand, if the position is item-conservative yet reductionist, the theory must, if it's a form of conceptual reductionism, effectively respond to the identification-problem and address the multiple-realizability issue. If the position is a variant of non-eliminative metaphysical reductionism, it must be shown that it does not collapse into some form of non-reductive materialism or dualism (such as, for instance, Davidson's anomalous monism) and at the same time succeeds in conserving the mental items as such without recourse to any strategy of conceptual reductionism, which is arguably a quite obvious contradiction. Also, non-eliminative metaphysical reductionism must also show how metaphysical identity does not actually presume the possibility of eliminative materialism (or it will also have to answer the critique directed against this position).

Moreover, *any* form of reductive physicalism must effectively defeat the indispensability strategy. It must accordingly, contra Taliaferro, show that

---

<sup>208</sup> David Rosenthal's position is outlined in his *Consciousness and Mind*, Oxford: Clarendon Press 2005. For Kriegel's position, cf. Uriah Kriegel, *Subjective Consciousness: A Self-Representational Theory*, Oxford: Oxford University Press, 2009. Metzinger's position is presented in Thomas Metzinger, *Being No One. The Self-Model Theory of Subjectivity*, Cambridge: MIT Press 2011. Of these examples, only Thomas Metzinger explicitly rejects non-reductive materialism, instead opting for eliminativism.

the physical can indeed be conceived of, defined and identified without any recourse to ostension whatsoever. It must respond to the Kripkean critique of functionalism, since any form of reductionism maintaining that particular activities of the mind is identical to purely physical processes, must distinguish between the states of some equivalent to an abstract causal network and the purely physical states of the brain realizing that causal network, and thus employ idealization in identifying and distinguishing the activities of the mind from one another. As an aside, outright eliminativism is in principle exempt from Kripke's argument, but insofar as this position finds its warrant in functionalist analyses, it's to some extent undercut by the argument.

Finally, any and all forms of reductive physicalism must respond to the incommensurability arguments, which in and of themselves are sufficient to render the position untenable, since they rule out all forms of reduction in principle.

It may also be added that the historical causes and support of reductive physicalism as an ontology, i.e. its association with the particular industrial economics of early modernity, and its role as supporting ideology, arguably suffices to explain its emergence without regard to any actual arguments in its favour (which indeed seem to have been quite few and inconclusive). Thus, if reductive physicalism and associated positions never were a truly, philosophically credible alternative to their predecessors, and if their emergence can be explained in terms of ideology rather than their actual defeat of all competing metaphysics, their basic plausibility is diminished.

No models that are actually capable of overcoming the aforementioned obstacles, while retaining the distinctiveness from both eliminativism and non-reductive physicalism have so far been suggested, and with regard to several of the arguments that have been reviewed throughout the chapter, there are good reasons to believe that this is actually impossible in principle. Until there have been presented such comprehensive theories which can adequately address all the problems associated with an elimination or a reduction of the subjective, the positions criticised should at the very least therefore be considered implausible – yet a wholesale rejection of reductive physicalism as such seems warranted in light of the incoherence of eliminativism and the non-eliminative reductionist's ostensible dependence on the phenomenal, as well as due to the support for incommensurability.

## 7. The problem of interaction and non-monist metaphysics

In this chapter, we will examine the classical problem of interaction – the notion that a pluralist or dualist metaphysics with qualitatively different substances (or properties) in play seems to have certain difficulties in explaining how interactions between such differing entities are at all possible. We will also ascertain what implications the possible responses to this problem has for the idea of post-mortem survival as we've previously defined it. The investigation of the problem of interaction is due to its great significance as an obstacle to immaterialist and non-reductive physicalist ontologies alike, and the indirect relevance of these matters for the transcendent problem of death, the question of post-mortem survival.

What is the significance of the problem of interaction relative to the idea of the survival of persons beyond the corporeal death? Simply put, interaction is relevant to the issue of post-mortem survival for much the same reasons that it is taken to actualize the question of human free will. To the extent that survival is supported by e.g. a dualist ontology, be it substance dualism or the property variety, the question of interaction is relevant to the extent that a rejection of the idea may undermine the veracity of dualist or non-monist ontologies. Conversely, the question of interaction is also significant insofar as an affirmation of interaction may then substantiate survival-supporting ontologies.

Basically, the affirmation of intersubstantial interaction is more congenial to the idea of post-mortem survival than a rejection of it. In detail, this is for two main reasons. The problem of interaction is a significant detriment to the ontological positions which are capable of grounding the phenomenological criterion of personal identity which we have utilized in response to the problems of post-mortem survival in chapter 4. Also, the affirmation of the intersubjectively accessible causal efficacy of the immaterial subject or the basic self can be argued to be a necessary foundation of the rational affirmation of its actual ontological independence. That is to say, if it can't actually cause anything, its ontological independence can more obviously be questioned.

Causation is a key issue with regard to the problem of interaction. The denial of transsubstantial interaction is often expressed in terms of the endorsement of the principle of the causal closure of the physical (if the physical level is a causally closed system, where do the effects of the non-physical come into play?). Accordingly, all of the suggested responses we will examine ultimately rest on the compatibility of interaction and the particular causal metaphysics that are endorsed. We will first present the issue of the interaction problem as such, then examine the most important responses suggested, and finally discuss the viability of the causal metaphysics entailed by the responses.

In this discussion, the notions of *substance* and *property* will be central, mainly for reasons of clarity. The use of these contested concepts should not be taken to imply an endorsement of any particular ontology, but rather as an heuristic measure to illustrate and explicate the interaction problem.

A minimal, heuristic understanding of the notion of *substance* is the traditional idea that a substance is something basic which has properties. Properties, then, is that which substances possess. Following upon this, the idea that *ontologically different* substances can interact and influence one another is at the heart of the problem of interaction – yet we don't need to speak of substances per se for this problem to become actualized. Anything will do, insofar as we're within the framework of a non-monist metaphysics where the two or more classes of existent entities cannot in principle be reduced to one another due to qualitative differences.

## 7.1. The Cartesian origins of the interaction problem

We generally associate the mind-body problem and the problem of transsubstantial interaction with Descartes. Cartesian dualism can be said to present a comprehensive reductionist ontology as a response to the early-modern break with the more holistic Aristotelian metaphysics. The resulting position is a strict separation of mind from matter, with no intentionality or any immaterial attributes present in bodies as such, relegating each of the two substances to a metaphysical realm distinct from that of the other.

This severance is in principle, strictly speaking, no modern novelty – a clear division of the mind and body was acknowledged by Plato some two thousand years prior to Descartes in relation to the Plato's Theory of Forms, according to which the immaterial and timeless and universal forms or ideas dwell in a realm separate from the time and space of ordinary, material reality, a level of reality wherein the human soul also ultimately resides. A de-

tachment of mind from the material dimension also seems to be attested in the Tao Te Ching by Laozi, roughly contemporaneous to Plato. According to later interpretations of this work, the Tao can be identified with the qi, the life force or active principle of the living being, of which the mind (or rather something which might be best translated as *perceptivity* or *intentionality*) was considered an aspect, distinct from the physical body as such.<sup>209</sup>

However, Descartes' approach was original in that he explicitly considered the mind and the body to be two clearly separate objects of entirely different types, which at least in principle were functionally independent of one another, while the older accounts, e.g. the Aristotelian, never set them apart so radically, considering the immaterial form a necessary aspect of physical bodies, actualizing the formless matter. With Descartes, we accordingly face something of a Humpty-Dumpty situation, and ever since, all the king's horses and all the king's men have been quite unable to put the two back together again. Hopefully, we shall find ourselves more adequately prepared for this task at the end of this chapter, at least when compared to equines and professional mercenaries.

## 7.2. Cartesian substance dualism

The metaphysical dualism which Descartes suggests is a quite straightforward affair. Its basic premise is, again, that the mental and the physical are to be understood as distinct and separate substances, unlike, for instance, as according to contemporary property dualism, wherein the mental and the physical are not as sharply distinguished, as the mental is considered somehow arising from the physical level of reality. Substance dualism therefore doesn't merely distinguish between mind and matter on the grounds that neither is reducible to the other and that they therefore are ontologically distinct, but also claims that they are essentially qualitatively and ontologically different substances.

The concept of substance can helpfully be approached according to the basic Aristotelian definition, which was assumed by Descartes and still is widely accepted in contemporary philosophy. This definition is in line with the basic Latin concept of *substantia*, meaning *that which stands under*,<sup>210</sup> and indicates something which exists by itself beneath all its accidental features, capable of being modified and having attributes, and which may function as

---

<sup>209</sup> Jan Yün-Hua, "Problems of Tao and 'Tao Te Ching'", *Numen*, Vol. 22, 3, pp. 208-234 1975

<sup>210</sup> substance. (n.d.). *Collins English Dictionary – Complete & Unabridged 10th Edition*. Retrieved March 11 2014 from <http://dictionary.reference.com/browse/substance>

a cause (but still isn't merely an event of some kind).<sup>211</sup> The mental and the physical are on Cartesianism, in line with this, taken to be two separate substances, but of radically different kind or quality with regard to each other: the physical is extended in space, and has attributes which may be perceived by the human sense organs, while the mental is neither found in space, nor ever at all perceived by the sense organs, although perhaps directly or immediately encountered by phenomenal consciousness as such.

Reality is thus on substance dualism split in twain in a quite radical fashion. The mental and the physical, or mind and matter, are of entirely different natures and would seem to constitute two separate realms with little in the way of connection between them. To be sure, Descartes himself emphasized how intimately intertwined his body and his mind (the thinking thing which constituted his essence) must be, and that the situation isn't at all comparable to the loose relation of, say, a car to its driver (or perhaps of a carriage to its coachman to avoid the anachronism). This he argues since he for instance seems to experience physical pain as a phenomenally present, inescapable sensation and does not merely perceive injury by way of reason and the senses, by way of the calculating intellect, like perhaps how a shipmaster would detachedly note the damage to his ship's hull after having collided with an iceberg.

After having separated the two, it was in the early modern period still generally assumed that this immaterial part of human beings was causally efficacious on the physical realm – yet this clear separation of the body and the mind now adopted, implied some rather unforgiving difficulties in explaining how the two actually related to one another. How could something entirely immaterial conceivably influence the realm of mechanical relations, populated by nothing but objects and phenomena extended in space, which by means of physical action effected change upon one another? And from this position, it was but a short step to supposing that this immaterial mind simply did not exist – for since we have no obvious point of contact between matter and mind, and since the latter cannot be weighed, perceived or measured by any conceivable means, why should we assume that it's actually real?

In essence, this is the problem of interaction and one of its commonly proclaimed consequences, and, as has been previously mentioned, also one of the main rationales for modern reductive physicalism.

---

<sup>211</sup> Howard Robinson, "Substance", *The Stanford Encyclopedia of Philosophy*, Spring 2014 Edition, Edward N. Zalta (ed.). Retrieved 2014-03-31 from <http://plato.stanford.edu/archives/spr2014/entries/substance/>

### 7.3. The problem of interaction

If we accept substance dualism, then, the problem basically consists in the difficulties in explaining *how* an interaction would be possible (it can helpfully be stated as an issue of incommensurability: how could, for example, the subjectively experienced beauty of a musical piece ever by causal influence affect the size of a physical object?), and the lack of clear cut empirical evidence that such interactions actually take place.

To utilize the well-worn billiard metaphor, we may observe how the pool cue strikes the cue ball, which in turn, unfortunately (since it was the opening strike) sets the 8-ball in motion, sending it down the pocket, which allows us to reasonably infer that we've just witnessed how physical objects causally interacted with other physical objects – incidentally with the bothersome result that we lost the bet and consequently were forced to buy everyone in the bar a round. But the influence of the immaterial mind upon, or role within this series of events, is arguably not obvious. With regard to the mind's influence upon this type of causal chains, we may possibly argue like Erwin Schrödinger once did, and contend that we have a direct, incontrovertible experience of the causal efficacy of our minds on our bodies.<sup>212</sup> Yet, this line of reasoning can always be countered by the claim that this perception is somehow misleading, that what we normally perceive as voluntary action really is predetermined, or in any case caused by something other than what we experience as will and action. Interaction between mind and body thought of as different substances thus seems difficult to explain, as well as verify, at least in any straightforward manner.

Now, given this quite troublesome problem of explaining interaction, and the absence of any point of contact between the physical and the non-physical, what really becomes of dualism as a metaphysical position? Is the problem of interaction formidable enough for us to deem dualism so improbable that it cannot be rationally maintained, or can there still be reasons for us to hold on to it?

To begin with, it seems that the problem of interaction taken by itself cannot undermine dualist metaphysics, even though it's generally assumed to be severely problematic for the Cartesian variant of the position.<sup>213</sup> If we for instance were to accept that there are no possible solutions to the problem, and that interaction accordingly never can take place, substance dualism as a metaphysics can nonetheless still function. We may for instance assume that

---

<sup>212</sup> Klaus Mainzer, *Thinking in Complexity: The Complex Dynamics of Matter, Mind and Mankind*, Heidelberg: Springer-Verlag 1994, p. 163

<sup>213</sup> Cf. Lycan 2013



both the mental and the physical exist as separate substances although they never interact, and with regard to apparent interaction maintain that any causal relations between them we which may perceive are purely coincidental (or perhaps are orchestrated by their common origin as on occasionalism or similar positions). To utilize yet another vehicular metaphor, it's perfectly coherent to maintain the occasionalist position that the mind is the equivalent of a young child in the passenger seat of a car, unwittingly handling a toy steering wheel attached to the dashboard while falsely believing that he actually controls the vehicle, when it in fact is driven by the amused parent next to him.

Nevertheless, if we assume that the problem of interaction is impossible to solve from the dualist's perspective, a number of indirect difficulties of some magnitude seem to appear.

The basic intuition mentioned above, that the absence of any point of convergence between mind and matter gives us reason to also doubt the ontological independence and even existence of the immaterial mind, isn't to be taken lightly. If we accept that no connections at all are evident to us, and that the interaction as such is conceptually problematic, it at the very least seems necessary to put forth independent arguments for the independent existence of this immaterial aspect of humans, as was exemplified in chapter 6.

Another difficulty which the assumed non-interaction entails for the Cartesian is that it would seem to undermine the freedom of the will. Obviously, this isn't necessarily a fatal flaw (substance dualism is indeed compatible with determinism), but since libertarianism in its various forms is a position of great relevance for not least the theistic context wherein substance dualism has been widely acknowledged, it would seem a limitation on the part of dualism if it could not readily accommodate the idea of human free will.

This conflict is somewhat self-explanatory; if the immaterial mind which is equated with the individual self cannot influence anything whatsoever in the external world, what then becomes of its purportedly free will and intentionality? The substance dualist who accepts non-interaction may perhaps then argue some form of compatibilism, claiming that the will as such is nonetheless to be considered free, even though it's not causally efficacious in relation to physical matter. Or she might choose a compromise and argue along the lines that the mind's intentions may have consequences in the immaterial, mental realm, the spiritual theatre or the like, and that its limitations in the physical sphere therefore is of no real relevance for the freedom of the will.

Anyhow, it would seem that the problem of interaction causes enough *prima facie* difficulties for the position of classical substance dualism and associated perspectives that the dualist ought not accept it as a given, but instead try and overcome it, if at all possible. There are three main problems in line with what has previously been sketched:

1. The assumption of non-interaction tends to reduce the plausibility of claims towards the existence of the mind as a non-corporeal entity, unless idealism is assumed. On non-interaction, the immaterial mind cannot affect sense-organs or measuring instruments, neither directly nor indirectly. Yet things that exist are usually somehow measurable, either directly or via their concrete effects, and therefore, the plausibility of the mind's existence as an immaterial substance is undermined.

Moreover, on non-interaction, it's possible to argue that the mind really has no empirically accessible actual effects in any sense. This is problematic for the dualist since the notion of an actual entity completely lacking in any causal efficacy of any form whatsoever, seems positively incoherent.

2. Non-interaction reduces the scope of human free will to the extent that it cannot influence physical reality, in case the latter is considered independent of the mind, which renders the intentions of the immaterial mind irrelevant for embodied human existence. Thus, if human freedom or effective intentions must entail that our minds are able to directly influence such things as the motion of our own bodies, non-interaction renders such freedom and efficacy conceptually problematic.

3. Finally, non-interaction together with the assumption of the existence of an immaterial mind to some extent undermines dualism in favour of idealism, since it renders the independent existence of a material reality able to influence the mind implausible (which of course isn't a problem in case one is ready to embrace idealism).

A further issue that needs to be mentioned arises in relation to theism. Traditional religious metaphysics according to which the divine is portrayed as a non-material being or entity, influencing and communicating with and creating the physical universe, seems to presuppose something analogous to transsubstantial interaction. Insofar as theism is held to be true or plausible, non-interaction in general therefore seems to be a less preferable position.

For these reasons alone, there seems to be ample reason for the dualist who also accepts the idea of post-mortem survival to also maintain the idea of intersubstantial interaction, and defend it if possible.

### 7.3.1. Basic responses to the problem of interaction

Having portrayed the problem as such, we will now turn to the most important responses to the problem of interaction presented in relation to a defense of dualism and other forms of non-materialist metaphysics. We will present the responses, and assess what metaphysical commitments they entail, as well as which positions on the nature of causation are endorsed.

#### 7.3.1.1. Negating the principle: Swinburne's response

Swinburne addresses the problem of interaction by criticizing the assumption of the principle of causal closure of the physical, whereas this principle simply prohibits that anything non-physical may causally influence that which is physical. Swinburne states the principle as the position that all physical events, insofar as they have causes, only have strictly physical causes.<sup>214</sup> The principle of the causal closure of the physical can, formulated thusly, be considered tantamount to the denial of transsubstantial interaction, or at least as implying such a denial.

Swinburne's response to the interaction problem is twofold. He maintains that the principle of causal closure is unfounded, and that we have direct, *prima facie* empirical evidence for the interaction of the physical and the mental. With regard to the principle of causal closure, Swinburne argues that it cannot be justified, and therefore cannot be held rationally, whereupon the interaction problem collapses.

The direct *prima facie* evidence for interaction, Swinburne finds in such situations as an injury causing us pain. Here an introspectively inevitable, phenomenally present pain-event is evidently caused by some physical event, such as our spilling boiling water on ourselves or stepping on a vengeful honeybee. These are according to Swinburne paradigm examples of causal connections.<sup>215</sup>

In terms of the lack of justification for the principle of causal closure, the argument is twofold.

---

<sup>214</sup> Swinburne 2013, p. 104

<sup>215</sup> *Ibid.* p. 105

First, Swinburne attempts to show that contemporary science provides no evidential support for the non-efficacy of intentions. He begins with an analysis of the evidential relevance of neuroscientific research, and concludes that at best, neuroscience can only provide support for the position that we're not always aware of previous intentions after the fact, and that we do not always precisely gauge the timing of our intentions. Libet-type experiments, which seem to provide the most relevant evidence in this case, cannot, according to Swinburne, give us any reason to disregard the causal efficacy of intentions.<sup>216</sup> Moving on to an analysis of the evidential relevance of physics, Swinburne's conclusion is that contemporary non-deterministic interpretations of quantum theory do not support the principle of causal closure, whereas an appeal to physics cannot definitely rule out the efficacy of intentions.

Secondly, Swinburne argues that the principle of causal closure necessarily cannot be justified empirically. Swinburne maintains that true memories of previous experiences in humans necessitate that the experience has actually caused a brain event, which may later function as the cause of a phenomenally present memory of the experience. But since a justified belief in a theory in turn requires justified belief in that it can provide true predictions, the principle of causal closure then fails, since we cannot have justified belief that a theory can provide true predictions without deferring to the memory of experiences. Thus, Swinburne argues, even if Libet-type experiments *could* provide evidential justification in support of the principle of causal closure, we could never actually utilize this evidential support without deferring to memory, which then entails the negation of the principle of causal closure.<sup>217</sup>

Swinburne however never directly addresses the metaphysical gap that conceptually underlies the problem of interaction on dualism and related ontologies, instead settling with criticizing the principle of the causal closure of the physical according to the above lines of reasoning.

The conceptual difficulty of actually providing an exhaustive description of *how* the mental possibly could influence the physical on the assumption of the reductive dualist account of early modernity, has made this type of response into something of a go-to strategy for defenders of interactionist dualism, who many times content themselves with ostensibly demonstrating that the principle of causal closure is unfounded. William Hasker emphasizes this

---

<sup>216</sup> Ibid. p. 106-108

<sup>217</sup> Ibid. p. 117-118

conceptual difficulty in his *Metaphysics*, maintaining that mind-body interaction will always, necessarily, remain “deeply mysterious”.<sup>218</sup>

### 7.3.1.2. Broadening causality

The obvious way to directly address the metaphysical issues adhering to the notion of transsubstantial interaction is to target the understanding of causality at the basis of the interaction problem. This is generally done by simply selecting and defending a notion of causality that does not imply the truth of the principle of causal closure of the physical.

In the recent anthology *After Physicalism*, Uwe Meixner, in the article “The Naturalness of Dualism”, argues that an adequate theory of causation cannot rule out the conception of causation in such a way that entities which are *not* mind-independently physical can interact with those that are.<sup>219</sup> Meixner’s position is basically that there are no conceptual problems in asserting that transsubstantial causation occurs. His argument is that the behavior of an animal (*A*), can possibly be the effect of a phenomenal experience (*B*), even though (*A*) can in principle occur without there existing any “realm of consciousness” (*C*), which, however, is not the case regarding (*B*). The point is that the *prima facie* ontological independence of *A* with regard to *C* does not imply that *A* necessarily cannot be the effect of *B*, which however is ontologically dependent upon *C*.<sup>220</sup>

Such a response, however, is still unable to satisfactorily address an assertion of metaphysical impossibility based in a strong metaphysical gap, which would imply that *A* could not possibly be the effect of *B* since they’re ontologically distinct and as the deeper metaphysics of causation does not permit of such interactive events. In specifically arguing against such an assertion, one needs to provide an account of the metaphysics of causation that indeed does permit of such interactive events. We will examine the possibilities of providing such an account later in this chapter.

### 7.3.1.3. The compromise of emergent dualism and property dualism

Various forms of emergentism such as property dualism attempt to address the problem of interaction by narrowing or otherwise mitigating the ontological gap.

---

<sup>218</sup> Hasker 1983, p. 58

<sup>219</sup> Uwe Meixner, “The Naturalness of Dualism”, in Benedikt Paul Göcke, *After Physicalism*, Notre Dame: The University of Notre Dame Press 2012

<sup>220</sup> Meixner 2012. p. 34-35

William Hasker's emergent dualism portrays the mind as a simple substance that somehow emerges from the material level of reality, which is taken to at least moderate the severity of the problem of interaction since on this account, the rearrangement of matter must necessarily be able to influence the character or the emergence of the mental substance.<sup>221</sup> This would imply that the mental substance is something akin to a higher level of physical properties, which then arises from particular configurations of purely physical matter, yet qualitatively differ from the basic physical substance to such a degree that they must be considered ontologically different (e.g. by including intentionality), and then feed back into the physical causal processes.

Property dualism can provide a similar response with regard to the problem of interaction. On this position, immaterial *properties* emerge from a physical basis, and the rearrangement of matter can thus influence the non-material property in a similar sense as that described in the last paragraph.<sup>222</sup>

However, it's difficult to see exactly how either emergent dualism or property dualism can really make a dent in the problem of interaction as such. Surely, if it is assumed that non-material substances or properties may readily emerge from a physical substrate, the possibility of at least one-way causal interaction must necessarily follow – but if this very event of emergence is metaphysically impossible due to the ontological gap, the emergentist is obviously faced with fundamentally the same type of problem as is the classical interactionist dualist. In other words, if physical particulars are exhaustively analyzable and understandable in terms of objective physical properties as on the post-Cartesian conception of matter, and if it is the case that transsubstantial interaction is metaphysically impossible, how can it be the case that strictly physical particulars either exhibit or instantiate irreducibly mental properties, or effect the emergence of an irreducibly mental substance?

With regard to property dualism, this difficulty seems decisive. If a particular attribute A is properly applied to entity X, then there must be something in or about X that grounds the applicability of attribute A to entity X. But if this attribute is irreducibly mental, and if there's nothing irreducibly mental in or about the physical particular X that supposedly instantiates attribute A, there's nothing that could reasonably ground its possessing this mental attribute. In other words, the metaphysical gap at the root of the problem of interaction still remains, no less daunting.

---

<sup>221</sup> William Hasker, *The Emergent Self*, Ithaca: Cornell University Press 1999

<sup>222</sup> Nino B. Cocchiarella, "Property Theory", in Edward Craig (ed.) *Routledge Encyclopedia of Philosophy*, New York: Routledge 1998

Emergent dualism faces similar objections. Here, the contention is not only that a certain physical particular (of a sufficient level of complexity, we may add) may engender the emergence of non-physical properties, but that it actually causes the emergence of a bona fide non-physical substance. Conceptually, though, it does not matter much whether the emergence regards a substance or a property, since both are necessarily dependent on a particular configuration of matter. If the material substrate ceases to exist, so does the emergent attributes as well as the emergent substances. An emergent substance is thus similarly dependent upon the material substrate as is the emergent property. But what carries or grounds the persistence of this mental substance? The complex physical system seems to be the only viable candidate, but it's unclear as to how this is to both facilitate the emergence of, and then support the persistence of a substance wholly ontologically other. The metaphysical gap seems as much of a problem as always.

#### **7.3.1.4. The hylemorphic response**

The problem of interaction is, as we have seen, a product of the reductive view of matter and mind in terms of clearly separated substances which was established in relation to Cartesian dualism as an indirect consequence of the repudiation of Aristotelian hylemorphism.<sup>223</sup> Therefore, a possible response to the issue is to be found in returning to an Aristotelian variant of dualist metaphysics which arguably does not generate the problem of interaction in the first place.

How does then hylemorphism ostensibly sidestep the interaction problem? Basically, it endorses a more complex view of causal processes in comparison to the mechanistically inclined early-modern variant, and conceives of the nature of the soul and the body quite differently from Cartesianism, utilizing something which might be described as a holistic, integrative model of the soul and body. On hylemorphic dualism, the physical body and the immaterial soul are intimately connected, really integrated. The soul and the body are not independent substances as on the Cartesian account, but actually entirely inconceivable without one another. The immaterial soul is the form which informs the body, and the body is informed by the immaterial soul, whereas there can be no physical objects apart from their forms. Matter-in-itself can in principle not exist without the form, and a physical object is thus to be considered something like a composite entity of one single substance that has two metaphysical components.<sup>224</sup>

---

<sup>223</sup> Cf. Nancey Murphy, *Bodies and Souls*, Cambridge: Cambridge University Press 2006, p. 45

<sup>224</sup> David S. Oderberg, "Hylemorphic Dualism", in *Social Philosophy and Policy* 22, 2005

Moreover, the hylemorphic account employs a set of four causes. While the subsequent Cartesian-Newtonian model is generally taken to maintain only two basic aspects of causal processes, the material and the efficient cause, the Aristotelian metaphysics involves four. The material cause is simply what a thing is made of: the material cause of a gold ingot is the element of gold. The efficient cause is whatever produces the effect that is the event or complete object: that which brings the gold ingot as such into existence, or that which produces the specific effect when a certain event is considered to be caused. The material and efficient cause are readily translated into mechanistic action, and when the other aspects of causation are removed from the picture, the interaction problem is more or less inevitable. The notion that an immaterial substance somehow enters into mechanistic action as such and thereby produces a material effect is as we have seen clearly problematic.

Instead, however, the hylemorphic account describes the situation as one in which the immaterial soul *constitutes* the matter of the body in terms of the body's formal cause. An intentional event ostensibly caused by a human person in terms of an agent therefore does not entail mechanistic transsubstantial interaction in the same sense as it would on the Cartesian account. It must rather be considered a complex event involving one composite substance and all four causes. An event such as one's responding to an e-mail then involves the mechanistic level of causation, i.e. the material and efficient causes, but also the non-material formal causes of the substances involved, and the final causes in terms of the event's intentional aspects wherein will and conscious agency are found.

This is then argued to avoid the problem of interaction since only one substance is involved, and as this substance already involves metaphysically non-physical as well as physical aspects or components which are argued to be inconceivable as existing separately. Also, since the types of causes ascribed to the non-physical components are not mechanistic in character, but of an entirely different order, the problem of explaining how the immaterial soul could possibly "collide" with the material particulars of e.g. the human brain and produce a specific physical effect, does not as obviously arise. On the hylemorphic account, it's in principle impossible to describe the complex event of a human agent causing an effect without reference to all the four causes; there's no exclusively material side of the event that can conceivably exist on its own.

The issue of exactly how the metaphysically immaterial components or aspects of a particular body can actually produce effects with regard to the purely physical components does therefore arguably not arise, since no such components *can actually exist* in isolation. Yet it seems that some equivalent



of the problem of interaction can still be formulated in terms of the immaterial form interacting with the material substance, or the conscious intention actually producing an effect in terms of the behavior of physical objects.

Thus, to fully come to terms with the issue of interaction, it seems that the hylemorphist too must finally defer to some appropriate model of causal metaphysics as such. This, however, is already implied by the Aristotelian-Thomist metaphysics upon which hylemorphism traditionally was founded. Basically, this ontology presents the outline of a solution to a general problem of causal metaphysics which, arguably, any causal theory cannot really do without. The final cause, the inbuilt intentionality of the Aristotelian-Thomist system namely provides an explanation for why, precisely, a certain cause consistently leads to a particular effect rather than another, in a situation where a multitude of effects are metaphysically possible.<sup>225</sup> Teleology thus gives us what is arguably a necessary explanation for the consistency of particular causal effects. And as teleology is not material, the hylemorphist may argue that a kind of interactionism must be presumed by *any* model of causal metaphysics that involves regular relationships between causes and effects – implying that any reasonable model of causation must allow for interaction.

So how could a more fleshed out “appropriate model” of causal metaphysics allowing for even transsubstantial interaction look in detail? In exploring this, we will now examine models of causal metaphysics as a possible means of directly addressing the problem of interaction. Such models may of course be effectively combined with the basic responses we’ve just discussed, rather than just operating separately.

## 7.4. The problem of interaction and the metaphysics of causation

Causation, as commonly understood, entails the interaction of at least two distinct entities in such a way that a change is produced in one of them as a consequence of the other acting upon it. An entity which persists or acts in such a way that some phenomenon is produced as the result is considered the cause, whereby that which obtains from this event is the effect.<sup>226</sup> The link-

---

<sup>225</sup> Cf. Edward Feser, *The Last Superstition*, South Bend: St. Augustine’s Press 2012, ch. 2c

<sup>226</sup> “causation, n.”. *OED Online*, June 2017, Oxford University Press, retrieved from: <http://www.oed.com/view/Entry/29140?redirectedFrom=causation& 2017-08-12>

age between these two parties, however defined and understood in detail, is the relationship of causation.<sup>227</sup>

But even these seemingly innocuous statements presuppose a great deal which upon closer inspection isn't uncontroversial. The *distinct entities* in the above paragraph echo the Cartesian substances, utilizing which we will tend to approach reality as populated by item-like objects bearing fixed attributes. Contemporary physics has been taken to indicate that such a description of material reality is inadequate in many circumstances. Furthermore, the assumed definition of the causative relationship as actual *production* of change by the interactive action of one such entity-substance upon the other, seemingly disregards the issue of the empirical inaccessibility of causation-in-itself.<sup>228</sup> To be sure, something evidently takes place in such circumstances we identify as instances of causal interaction, yet it's been notoriously difficult to pinpoint exactly what this is, and whether it can be said to be common across all different types of phenomena where causation is invoked. Such issues have tended to direct contemporary theories of causation towards less pregnant conceptualizations of the causal relationship as such.

#### 7.4.1. Four basic models of causation

There are four major contemporary approaches towards an analysis of causation, which may be considered basic foundations for more particular analyses. As a popular contemporary example, albeit with old roots, we have *the counterfactual theory of causation*, which in its early forms appears in the writings of Hume himself as a possible model compatible with his critique of legacy principles.<sup>229</sup> The counterfactual theory has an intuitive appeal in a basic sense: we tend to think of causation in terms of something which makes a difference from what would otherwise have been, whereas had the cause been non-existent, its effects would neither have emerged, or at the very least not in the same fashion.

The approach can thus be summarized as the notion that for X to have caused Y, Y could not have taken place (or existed) if X were absent. X causes Y iff the absence of X would have rendered Y impossible. The counterfactual theory is generally expressed using modal logic, due to its depend-

---

<sup>227</sup> Nicholas Bunnin, Jiyuan Yu (eds.), "Causation", *The Blackwell Dictionary of Western Philosophy*, Oxford: Wiley-Blackwell 2004, p. 107

<sup>228</sup> Cf. David Hume, *A Treatise of Human Nature*, Oxford: Oxford University Press (Clarendon Press) 1960/1740

<sup>229</sup> David Hume, *An Enquiry Concerning Human Understanding*, Oxford: Oxford University Press 2007/1748, sect. VII

ency on unactualized possibilities, which are effectively subsumed under the modal category of possible worlds.<sup>230</sup>

There are a few problems with this approach. A major drawback with the counterfactual theory lies in its *prima facie* inability to take the context-sensitivity of the causal relationship into account. It tends to assume that causation is a basic type of relationship, common across all different types of phenomena, with the result that it cannot readily distinguish between what might be termed background conditions and what more intuitively would be considered the actual, direct causes of a particular phenomenon (cf. the distinction between causation *per se* and causation *per accidens*). The counterfactual theory could for instance possibly identify the invention of the aeoli-pile in the 1<sup>st</sup> century as a cause of the 1969 moon landing, which in some respects can be seen as problematic. Other issues, which we will not explore here, reside in the theory's approach to the temporal direction of causation, and its conceptual adherence to causal chains and associated problems of transitive causal effects.

Counterfactual theories were advanced to address some of the problems associated with the *regularity theory* (or *covering law theory*) of causation. The starting point of this latter approach is the Humean notion that cause C is a cause of effect E iff C belongs to a set of conditions which taken together are sufficient for the emergence of effect E, given (inductively inferred) laws of nature.<sup>231</sup> The contemporary regularity account owes a great deal to J. L. Mackies INUS-model, which can be considered a general refinement of the regularity theory. The INUS model mitigates some of the issues which has plagued the "naïve" regularity theory, such that C in the above example may very well belong to a certain set of minimal conditions sufficient for the emergence of E, even when C is itself an effect of E; or when C and E together share the same cause. The INUS approach instead considers what we normally would think of as the direct proximal cause of a certain event (the efficient cause in Aristotelian terminology)<sup>232</sup> as an *insufficient* but *non-redundant* part of a condition, which is itself *unnecessary*, but *sufficient* for the effect to occur. As an example, say that the lighting of a campfire is caused by the flame of the lit match touching the kindling, the presence of cured firewood and the absence of adverse weather. According to the INUS approach, the lit match is then analyzed as an in itself insufficient, yet non-redundant part of a total condition which is sufficient for the effect (the light-

---

<sup>230</sup> Peter Menzies, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy* (Spring 2014 ed.), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual> 2014-09-20

<sup>231</sup> A. J. Ayer, *Language, Truth and Logic*, London: Penguin Books 2001, pp. 40-42

<sup>232</sup> Aristotle, *Metaphysics*, V, 1013a

ing of the campfire) to occur, yet not actually necessary (as there are other potential ways in which the effect might have emerged).<sup>233</sup> Since cause C according to INUS is unnecessary for the emergence of E, there's no formal contradiction if C itself actually happens to be a consequence of E, or if C and E share the same cause.

A further approach which has recently garnered attention, in part due to the intensified analysis of the philosophical implications of quantum mechanics of the late 20<sup>th</sup> century, is the *probability theory of causation*. The basic idea of this perspective is that causation is in effect whenever an event or state of affairs makes another particular event more likely to occur. So if X is to be considered the cause of Y according to the probabilistic approach, the presence of X as a fact or an event must have rendered the appearance of Y more likely. An early, basic version of this model was presented by Hans Reichenbach in the 1950s, and equates causation with the relationship that holds between the cause C and the effect E iff the probability that E occurs is greater given C than otherwise.<sup>234</sup> This can be formalized as:

$$P(E|C) > P(E)$$

Since causation in the common understanding stated above refers to the actual bringing about of an event given certain factors, this model has a somewhat counterintuitive character when applied in analysis of particular past events we know from experience, and the putative connections between them. This becomes obvious if we consider that the notion of increasing the likelihood of some event occurring, does not by itself entail it actually doing so, unless of course the increase is from  $P < 1$  to  $P = 1$  (from less than 100% probability of occurrence to 100%). The implication is that successful causation on the probability model doesn't necessarily mean that the event in question actually takes place. Developments of the theory has grappled with this issue among others, suggesting relatively simple solutions such as allowing that the causative effects of an event doesn't actually have to bring about the emergence of the caused effect, but merely needs to push it further towards the edge of actuality within the realm of possibilities. Another possible solution is to postulate an actual multiverse along the lines of Everett's many-worlds response to the measurement problem of quantum mechanics, implying that every successful cause, for some particular event, rendering  $P > 0$  actually does bring that event about in modal reality.<sup>235</sup>

---

<sup>233</sup> J. L. Mackie, *The Cement of the Universe*, Oxford: Oxford University Press 1980, pp. 62 ff

<sup>234</sup> Cf. Hans Reichenbach, *The Direction of Time*, University of California Press 1956

<sup>235</sup> David Wallace, *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*, Oxford: Oxford University Press 2012, pp. 35-39

The fourth major contemporary theory isn't so much an analysis of the causative relationship itself, as a resignation to the seeming impossibility of attaining one. This type of approach is known as the *primitivist theory of causation*, due to its holding that causation is an entirely irreducible or ontologically primitive phenomenon or relationship. A fundamental argument for this view lies in philosophy's failure to provide an adequate analysis of causal connections, while another rests upon the alleged circularity of the other accounts, which have been said to implicitly presuppose causality in some of their central notions.<sup>236</sup> The *event* is an example of this, which has been argued to be inexplicable without reference to causation, since the event by definition modulates properties, which in turn are defined with regard to their causal roles.<sup>237</sup> A parallel can be drawn between the primitivist account and the Kantian response to Hume's critique of the concept of causation, according to which causation as an a priori conceptual category must be regarded an indispensable basis for structured thought as such, which implies that causation as such can never be accessible via reductive analysis.<sup>238</sup>

There are of course several other interesting perspectives on causation which have been omitted, yet these four approaches make up the general conceptual foundation upon which more specific theories of the causal relationship have been constructed, which may be considered refinements or combinations of the basic approaches. Such theoretical artifacts as the structural process approach, suggested as a solution to some of the problems associated with the probability theory, would then for instance be considered a refinement of the latter,<sup>239</sup> while the agential manipulability theory explored by von Wright, where causation is approached as possible agential action, can be seen as a variant of the counterfactual theory's theme.<sup>240</sup> The mechanistically-inspired analyses of causation as energy transfer or contiguous change would also seem to fall into this category.

In summary, we have a diverse set of causal theories which permit no extensive consensus upon the nature of the causal relationship as such. It seems that all would minimally agree, primitivism included, that causation as a concept describes a situation wherein it seems that an event, entity or state

---

<sup>236</sup> D.M. Armstrong, *A World of States of Affairs*, Cambridge: Cambridge University Press 1997

<sup>237</sup> Sydney Shoemaker, "Causality and Properties", in Peter van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor*, Dordrecht: D. Reidel Publishing 1980, pp. 109-135

<sup>238</sup> Stuart Hackett, *The Resurrection of Theism*, Austin: Wipf and Stock 1957, pp. 37-113

<sup>239</sup> Peter Menzies, "Probabilistic Causation and the Pre-emption Problem", *Mind*, 105, pp. 85-117, 1996; Ned Hall, "Two Concepts of Causation", in J. Collins, N. Hall (eds.), *Causation and Counterfactuals*, Cambridge: MIT Press 2004, pp. 181-204

<sup>240</sup> G. H. von Wright, *Explanation and Understanding*, Ithaca: Cornell University Press 1971

considered the cause exerts some form of influence, yet differ regarding how the cause ought to be defined, and what particular type of influence is to be regarded as a causal effect. Important to note is that none of these basic perspectives gives us much information about the exact nature of the causal relationship as such, something which is emphasized by primitivism. This is significant, since it's this particular aspect of a causal theory which is of greatest relevance for the interaction problem.

## 7.4.2. Causal-model approaches to the problem of interaction

With this overview of the fundamental contemporary approaches towards causation, we can now move on to a review of related, possible responses to the problem of interaction.

### 7.4.2.1. The proxy approach

Since the problem of transsubstantial interaction concerns whether and how ontologically disparate entities may influence one another, one commonsensical approach is to explore the possibilities of an interface between them. This could be termed the *proxy approach*, as it attempts to discern by what means we might be able to bridge this metaphysical gap between mind and matter (or whatever ontologically distinct entities, substances or parts one considers) which seems to prevent us from conceiving an efficient causal interaction. Its main question would then be whether we can provide an interface environment for them in which interaction is conceivable, or if the interaction could take place via some sort of proxy. This proxy-model of interaction can be conceived of in four different basic forms. Assuming substance dualism as an example, we may suggest a physical proxy, such as the entire brain, or as perhaps infamously considered by Descartes, the pineal gland, which by virtue of some particular property may function as a point of contact for the disparate substances.<sup>241</sup> We could also postulate a mental equivalent, which for the similar reason of possessing special properties would function as the transsubstantial bridge. A third option would be to consider some sort of hybrid of the two substances playing the role of conduit for causal interaction, while a fourth version of the proxy approach is to invoke a transcendent deity or spiritual theatre as an interface.

However, the proxy model is beset by a number of significant difficulties, which with regard to the problem of interaction render it a questionable solution at best. First of all, it's not at all obvious how a proxy-entity of either substance could manage to bridge the ontological cap. To claim that a physical proxy can act as a channel of causal influence without suggesting how

---

<sup>241</sup> Descartes most complete treatment of the pineal gland and its role in channeling the influence of the soul is found in his *Les passions de l'âme*, Amsterdam: Lodewijk Elsevier 1649

this is possible, say with regard to some unique attribute, merely masks and disregards to respond to the basic issue. If the mental and the material cannot obviously interact, there's no *prima facie* reason to believe that interaction can be facilitated with the aid of a particular privileged object that is either material or mental. Moreover, the conceivable unique attributes which supposedly render the privileged object a functional proxy, must also be physical on the assumption that the proxy object is reducibly physical, whereas they seem to offer little in the way of a solution to the interaction issue as such. These same general difficulties would also hold if the proxy entity is considered an environment or a context rather than a specific object, and seem to be equally troublesome whether the proxy is considered physical or non-physical.

The notion of a hybrid proxy might seem more promising at first glance, but faces similar problems. In the absence of an explanation of transsubstantial interaction, the assumption of a hybrid would be an obvious case of *petitio principii* as the functional merging of the two substances in the hybrid needs to bridge the ontological gap as direly as the monosubstantial proxies do. Insofar as the previously discussed hylemorphic response is to be considered a kind of hybrid proxy solution, however, it avoids begging the question by arguing that actual material objects are actually metaphysically impossible without their immaterial aspects (i.e. form and telos), and upon this maintains that all causal relationships necessarily must involve an affirmation of a form of interactionism. On this position, the mind-body-connection is no stranger than the form-matter relationship we see everywhere around us.

Divine intervention or transcendental interactionism is less problematic in terms of a proxy approach than those variants previously mentioned, as the postulating of a transcendent deity either as proxy object or context of interaction doesn't necessarily face the ontological gap. Indeed, this approach could address the issue by denying that such a gap has any relevance for the causal relationships in question. This, however, presumes if not classical theism with God as an omnipresent, omnipotent agent, then at least the ontological primacy of the transcendent realm, and that the material and immaterial substances under discussion are in some sense derivatives of this. Depending on how one frames the response, it might collapse into something along the lines of the occasionalism of Malebranche and Leibniz, denying transsubstantial interaction in other terms than God's foreknowledge of agential intentions and complete command of physical phenomena.

In summary then, the proxy response to the problem of interaction on a dualist premise seems beset by certain important difficulties. The most promising version seems to be the hylemorphic variant insofar as this one is considered

a proxy response, which however depends on the viability of its particular metaphysical assumptions. The transcendental proxy variant is a possible position, but entails the dismissal of, or radical reinterpretation of interaction to such an extent that the subject cannot without difficulty be considered having an independent causal character (if God is the actual causal agent electing to realize a subject's intentions, it's not entirely obvious that anything equivalent to an immaterial self of the human person causes said effects). This latter problem is also found with Malebranche's and Leibniz' occasionalism or parallelism.

The proxy model as a causal metaphysics seems compatible with, and could plausibly work in concert with all three basic response strategies discussed under 7.3.1.

#### **7.4.2.2. The primitivist response**

Another conceivable and quite common approach to the problem of interaction opens by protesting a tacit premise of the issue, namely that *intrasubstantial* interaction as such is a straightforward, non-problematic affair. This approach assumes a position akin to that of the primitivist theory of causality: that causality is an inexplicable, ontologically primitive fact which cannot be subjected to any further reductive analysis. The argument is then that since we really have no idea what's going on in the case of intrasubstantial interaction, and yet for various reasons accept it as relatively unproblematic, we should neither disregard the notion of transsubstantial interaction merely on the grounds that we cannot provide a detailed description of the relevant causal relationship at hand.

Nancy Cartwright, in her article "Causation: One Word, Many Things", provides what she terms a "thick" account of causation, basically maintaining that causation ought not be approached in terms of a single, monolithic concept, since "there are qualitatively different kinds of causal relations imbedded in different kinds of systems."<sup>242</sup> The position described in this article is somewhat akin to a primitivist account, in that it is proposed that we cannot possess a unified, all-encompassing understanding of causation as a delimited phenomenon in itself.

Further, Charles Taliaferro seemingly endorses an equivalent position in his *Consciousness and the Mind of God*, arguing that the experiential fact of mental-physical causal interaction is "no more or less mysterious" than exclusively intrasubstantial causal interaction. Taliaferro further argues that the

---

<sup>242</sup> Nancy Cartwright, "Causation: One Word, Many Things", *Philosophy of Science*, Vol. 71, No. 5 2004



ostensible indeterminism at the deeper layers of material reality which the physical sciences have unveiled implies that mental causation cannot be ruled out with regard to classical deterministic physics.<sup>243</sup>

This kind of argument receives further implicit support in the fact that our knowledge of the fundamentals of physical reality is comparatively limited. This is to say that we really don't have a satisfying, exhaustive understanding of what physical properties and/or substances *are*, (what we do have is a decent idea regarding what they *do*). Physical entities as described by contemporary science is rather a very diverse crowd, which no theoretical approach has yet successfully reduced to any common fundamental, yet we nonetheless assume that they interact without difficulties. The primitivist may further argue that since the major contemporary theories of causal interaction do not provide a detailed analysis of the very causal relationship itself, they thereby do not exclude transsubstantial causal interaction. As these theoretical approaches give us insufficient information regarding the intrinsic nature of the phenomenon of causal interaction for us to conclude that transsubstantial causation is impossible, there's simply no *prima facie* reason to think that the interactionist position is excluded by today's most prominent causal theories.

Indeed, few of these theories seem to necessarily and in principle entail the problem of interaction insofar as they do not very specifically describe the necessary nature of the causally related parties. The counterfactual theory, for one, poses no problem for the interactionist, as the latter's position entails that if some mental events were not to occur, certain physical events would neither, allowing for counterfactual causation in that if mental event M were absent, physical event P would not take place. The regularity theories would also in principle permit interaction, if we allow that M belongs to the set of conditions sufficient for the emergence of P, e.g. by virtue of some inherent lawlike psychophysical regularities. Similarly, the probability-raising theories seem perfectly compatible with interactionism, exemplified by the notion (or some equivalent thereof) that one's first-person experience of intense pain increases the likelihood of one's vocal cords producing a certain sound. The primitivist approach may of course also be embraced by the interactionist dualist, while perhaps needing to explain why transsubstantial causation

---

<sup>243</sup> Taliaferro 1994, p. 220-224. Several other defenders of interactionist dualism tend to assume a similar position regarding the interaction question, if they at all analyze it more closely (see Klawonn 2011; Swinburne 2013) arguing along the lines that materialism or dual aspect theory/property dualism simply give rise to more severe difficulties than assuming some mind-matter nexus with the adjunct problem of interaction, and therefore aren't preferable to interactionist substance dualism. The important difference is that primitivism as defined here can be employed to argue against the severity of the interaction problem as such, while this other similar position acknowledges it as formidable yet does not consider it decisive.

seems to take place in such an ordered fashion, again, possibly by some set of psychophysical laws, or by invoking the final causality or inherent intentionality of the Aristotelian-Thomist account.

Certain particular causal theories other than the general approaches previously mentioned, such as the mechanistically inspired approaches like the energy-transfer model, would on the other hand pose a problem for the interactionist.<sup>244</sup> These theories, however, face certain decisive obstacles, not least in the argument from our limited knowledge of physical reality implied above, and in the dubious character of their narrow description of the nature of the causal relationship (there are for instance noteworthy counterexamples to the energy-transfer description of causality in that many effects on the contrary seem to be produced by the actual interruption of transferred energy, such as the shape of a block of ice or the extinguishing of a fire).

To sum up, what has here been called the primitivist response in effect amounts to a questioning of the significance of the purported metaphysical gap for three distinct reasons: 1) current causal theories do not on principle rule out transsubstantial causal interaction due to their inability to define the causal relationship which arguably is ontologically primitive. 2) Our limited knowledge of fundamental physical reality seems to imply that a complete reductive analysis of causation is currently unattainable, and 3) that even intrasubstantial interaction is beset by ontological difficulties similar to those assumed in the problem of interaction. The primitivist response is an interesting approach to the problem of interaction which seems unencumbered by any severe difficulties. However, as it doesn't propose a solution by actually explaining the phenomenon of transsubstantial interaction and clearly demonstrating its possibility, but merely contends that our current inability to reductively analyze the causal relationship implies that it cannot be excluded, the primitivist response as such perhaps might be thought to have a somewhat unsatisfactory flavor. Moreover, it lacks the finality that a hypothetical thoroughgoing analysis of causation coupled with a demonstration of the possibility of the transsubstantial variant would have, since the primitivist response rests on the contingent fact of an inadequate state of human knowledge which may possibly be augmented in the future.

The primitivist response is thus compatible with Swinburne's response in 7.3.1.1., according to which the principle of the causal closure of the physical is questioned. One way to undermine this principle and support Swinburne's argument would be to assume something akin to the primitivist account of causal metaphysics. It would also be a way to provide a broadening

---

<sup>244</sup> David Fair, "Causation and the Flow of Energy", *Erkenntnis*, 14:215-219, 1979

of the notion of cause exemplified by Meixner's line of reasoning in 7.3.1.2., addressing the main issue with Meixner's argument, i.e. the lack of an appropriate causal metaphysics. The Aristotelian account described in 7.3.1.4. also seems compatible with the primitivist response to the problem of interaction, insofar as this response does not conflict with the four interrelated causes. Arguably, the Aristotelian-Thomist position may even imply a primitivist account, or something akin to the "thick" account suggested by Cartwright, in that the assumption of four distinct types of causation precludes any monolithic, unified notion of causation as a single, delimited phenomenon. In relation to the Aristotelian-Thomist position, the primitivist response can in particular be fruitfully implemented with regard to the potential problem of explaining how the immaterial form can effectively interact with the material substance, or the conscious intention actually producing an effect in terms of the behavior of physical objects: we don't know how they interact, but neither do we know precisely what happens in terms of the intrasubstantial causation endorsed by classical modernity, and since no physical bodies are conceivable without the immaterial form, we simply must presume that interaction takes place.

#### **7.4.2.3. Causality-modifying direct interaction**

A third conceivable response to the problem of interaction from causal metaphysics is to assume the conclusion that transsubstantial causal relations are possible from a model of direct interaction rather than a proxy approach. However, what has been said regarding the latter covers most aspects of any purported solution assuming direct interaction. Given the ontological gap, any model of direct interaction must show that it can be bridged without resorting to monism, turning into a proxy approach, and without modifying the notion of interaction to such an extent that the relationship between the entities in question is no longer causal (i.e. by assuming something akin to parallelism or Jungian synchronicity as a basis for connection). If we as an example of a direct interaction model consider a suggestion of transsubstantial "energy"-transfer as an explanation of causal effects, its proponents would be hard pressed to show how this model does either not entail an abandonment of interactionist dualism, or is an equivalent to a proxy solution with all the problems that approach entails. This difficulty obtains since said form of "energy" either must be regarded as an ontologically basic factor, common to both the mental and the material, which implies some form of monism, or assumed to function as a third party proxy belonging to one of the substance categories. If E is considered material, the assumption that mental phenomena and effects consists of E undermines their ontological independence, and vice versa if E were to be considered mental.

Nevertheless, there seems to be at least one conceivable way to maintain interactionist dualism on a direct interaction approach, which is to assume that the mental influences physical phenomena by modifying the relevant processes of *intrasubstantial* causation. So on the premise that intrasubstantial physical causation is explained with regard to energy transfer, this solution would for instance entail that the causal efficacy of the mental is possible by virtue of it somehow modifying or interfering with this energy transfer, without exerting direct influence on the very physical bodies as such, supposedly avoiding the incoherent position that the mental entities itself are emitting or producing physically efficacious energy. This model appears functional (and possibly more congruent) in relation to the alternative conceptions of causation as well: if we explain intrasubstantial physical causation with the probability theory, the interactionist may argue that mental-physical causation takes place in that the mental somehow modifies the probability interface between two or more physical events. The counterfactual and regularity theories likewise seem neutral regarding this approach, and the primitivist position does not obstruct it.

Certain interpretations of the observer-effect discussed in relation to quantum theory is very similar to this position, in that they describe this effect as a modification of the probability that certain actual effects materializes, rather than an effect upon concrete physical events or bodies as such. The next chapter will explore these issues more closely. This position is also conceptually similar to the Aristotelian understanding of causes as the actualization of potentialities.

## 7.5. Conclusions and further considerations

The problem of interaction, although possibly overrated in terms of just how damaging it actually is for the immaterialist ontologies, nonetheless poses them a significant obstacle. However, the possible responses reviewed, especially in regard to causal metaphysics, seem to reduce its severity and provide grounds for accepting the possibility of transsubstantial interaction and/or interaction between ontologically disparate attributes or aspects of composite entities.

A response to the problem of interaction, if not strictly necessary for the non-materialist ontologies, does indeed seem to be of vital importance. The dualist of various sorts can in principle assume something like occasionalism, but such positions are hampered by other difficulties, and call for an explanation of the co-occurrence of the mental and physical events. Strictly speaking, all metaphysical positions excepting the eliminative materialist variant of reduc-

tive physicalism and idealism are affected by the interaction problem. Thus, excepting an adequate response to the problem of interaction, the viability of the non-materialist ontologies is impaired.

Since the immaterialist or non-reductionist ontologies according to chapter 4 are the only ones compatible with the fulfillment of our criteria for post-mortem survival, it follows that the assumption of non-interaction indirectly would weaken the case for the metaphysical possibility of post-mortem survival, unless, of course, a strongly convincing case for metaphysical idealism could be put forth.

What has been insufficiently emphasized so far, however, is the potential *positive* consequences of an adequate response to the problem of interaction on the part of the non-materialist or phenomenalist ontologies. Since the problem of interaction is one of the major stumbling blocks for these positions – arguably the only really significant one – a response to the problem would greatly improve their outlook.

In other words, since the central difficulty with the various forms of dualism is the problem of interaction, an effective response to this problem, together with the many benefits of the non-materialist ontologies, such as their ability to do justice to our immediate phenomenal reality, renders this set of positions obviously preferable.

Whether the possible responses to the problem discussed in this chapter are ultimately successful, and which of them are to be preferred, is beyond our scope. However, the suggested responses provide several clearly viable options for the defender of interactionism, which gives us ample reason to regard the problem of interaction as surmountable. In particular, the hylemorphic and primitivist responses both stand out, whereas the hylemorphic one seems to be the most promising insofar as its metaphysical preconditions can be successfully supported.

But we haven't yet said our final word on the interaction issue. In our next chapter, we will not only assess the compatibility of the non-materialist ontologies with regard to contemporary physics. We will also touch upon the possible avenues of support for the non-materialist ontologies which can possibly be garnered from contemporary physics, including the additional ways the interpretations of, and theories from quantum mechanics, may provide us to address the problem of interaction.

## 8. Quantum mechanics, phenomenal consciousness and transsubstantial interaction

### 8.1. Introduction

This chapter will explore possible avenues towards supporting phenomenalist ontologies (all ontologies except reductive physicalism), as well as ways of mitigating the interaction problem, with the use of data and theoretical constructs imported from quantum physics.

The relevance of this field for ontology chiefly consists in its implications for our understanding of the nature of physical matter as such, and the concordance between acts of consciousness and certain material processes persistently attested to by certain interpretations of quantum physics. We will mainly examine John von Neumann's suggested interpretation of the wave-function collapse and Henry P. Stapp's contemporary theoretical refinement of this position, as well as the role and possible implications of John Bell's interconnectedness theorem.

### 8.2. The metaphysical implications of quantum mechanics

With the last chapter's overview of basic contemporary causal theories and the available general approaches towards an amelioration of the problem of interaction, we are now to assess the contributions to the ontological issues that interpretations of data and theoretical considerations from the field of quantum mechanics can provide. As previously indicated, this is far from a shot in the dark, due to the particular character of this scientific discipline and its marked break with the Cartesian-Newtonian paradigm in terms of the ontology of matter. Ever since its inception, quantum physics has had a close relationship to questions of the nature of consciousness.<sup>245</sup> This relationship

---

<sup>245</sup> Niels Bohr's work is an influential and early example of this, emphasizing the importance of a phenomenological approach to attain an understanding of the nature of quantum phenomena. See Niels Bohr, "Discussion with Einstein on Epistemological Problems in Atomic Physics", in Schilpp, P.A (ed.), *Albert Einstein: Philosopher-Scientist*, Fine Communications 2001, pp. 200-241

chiefly stems from the experimental discovery that the very act of measurement itself (whatever this exactly entails is a matter of debate, as will presently be explained) verifiably alters the future state of a physical system at the quantum level, compared to a situation where such a measurement is absent.<sup>246</sup>

To fully appreciate the central role of measurement for the theoretical peculiarities of quantum mechanics, and its associated relevance for issues of consciousness, we will have to briefly revisit the early history of quantum theory.

### 8.2.1. Planck, Einstein and the quantum stone age

This scientific discipline which eventually came to entail such profound consequences for our understanding of the nature of reality has a rather modest origin. The earliest days of quantum theory began with Max Planck's quite prosaic work on a solution to certain inconsistencies between empirical results and theoretical predictions regarding the energy distribution between different frequencies of electromagnetic radiation, presented in 1900.<sup>247</sup>

More exactly, the basic problem investigated by Planck prior to the formulation of his eponymous constant, was the question of why the spectrum of *black body radiation* did not conform to theoretical predictions at low frequencies of emitted thermal radiation, yet nonetheless could be successfully approximated at high frequencies. As is well known, a heated object (every object with a temperature above absolute zero) emits thermal radiation at frequencies corresponding to its temperature. At the temperatures of our everyday surroundings, this radiation is in the infrared spectrum, but above about 500 degrees Celsius, bodies begin to emit visible light. The frequency of radiation emitted by the idealized non-reflective black body (as is the case with a complete vacuum, no perfect black body actually exists) ought then correspond to nothing but the temperature of the object, as it doesn't reflect any incoming radiation. This was familiar in the mid-19<sup>th</sup> century, and it accordingly was assumed that the frequencies of black body radiation could be approximated with regard to nothing but the object's temperature. However, this proposal was only successful in relation to high frequency radia-

---

<sup>246</sup> Paolo Grigolini, *Quantum Mechanical Irreversibility and Measurement*, World Scientific Publishing Co. 1993, pp.2-7

<sup>247</sup> Max Planck, "Über das Gesetz der Energieverteilung im Normalspektrum", *Annalen der Physik*, vol. 309, issue 3, 1901, pp. 553-563

tion – the theoretical approaches of the late 19<sup>th</sup> century could for some reason not correctly predict the spectrum at longer wavelengths.<sup>248</sup>

To remedy this situation, Planck proposed that the equations of motion for waves of light are a set of *harmonic oscillators*, assigning each possible frequency its own such oscillator equation.<sup>249</sup> A harmonic oscillator is simply any system, which if forced out of its position of equilibrium, responds with a restoring force proportional to the previous displacement.<sup>250</sup> Still, this in itself did not suffice to produce the empirically correct formula, until Planck also added that the vibrational energy of these oscillators, these undulating waves of light, was not to be considered a continuous, divisible quantity extending towards infinity, in the manner energy distribution was thought of at the time. Instead, Planck, merely as a heuristic measure of his mathematics, assumed that the energy of the oscillator for each possible frequency, was a *discrete, finite quantity* proportional to the frequency in question. That is, something similar to a particle. The specific ratio between this discrete energy quantity and the associated frequency, rendered what today is known as the Planck constant, denoted  $h$ , which finally allowed him to correctly predict black body emissions.<sup>251</sup> Planck's constant was later called the "quantum of action" as it has the dimensions of energy times time, which is a quantity known as "action" in classical physics.<sup>252</sup> The implications of this were initially disregarded by the physics community, as the restriction upon energy implied by Planck's solution was fundamentally alien to classical physics.

In 1905, however, utilizing Planck's discovery in his work regarding the photoelectric effect (the emission of electrons by certain metals when the metal is struck by light), Albert Einstein showed that empirical data associated with this phenomenon could only be explained if the incoming light actually was regarded as the discrete packets of energy Planck had proposed (again, rather than waves extending to infinity, as according to classical electromagnetic theory), which definitely established the particle-nature of light.<sup>253</sup> In the view of 19<sup>th</sup> century physics, electrons in metal were thought of as suspended in the electromagnetic field of the atoms, whereas light con-

---

<sup>248</sup> Thomas Kuhn, *Black-Body Theory and the Quantum Discontinuity, 1894-1912*, Oxford University Press, Oxford 1978, p. 8-12; R. Bowley, M. Sánchez, *Introductory Statistical Mechanics*, Clarendon Press: Oxford, 1999, p. 164-165

<sup>249</sup> Planck 1901 (Herein Planck uses the concept of *resonator* for what would today be called an oscillator)

<sup>250</sup> Raymond Serway, John Jewett, *Physics for Scientists and Engineers, Vol I*, Belmont: Brooks/Cole, CENGAGE Learning: 2010, pp. 434-437

<sup>251</sup> Planck 1901

<sup>252</sup> Herbert 1985, p. 35

<sup>253</sup> Rusty L. Meyers, *The Basics of Physics*, Westport: Greenwood Press 2006, pp. 171



sidered a wave then supposedly could dislodge these electrons much in the same way as waves in water may eject debris upon the shore, with the implication that a more powerful wave would send the debris further up the coastland in comparison to a weaker one. According to actual photoelectric experiments, however, this was not the case – the energy of the ejected electron is always the same for the faintest light as well as for the most powerful beam, light of a higher intensity merely dislodges a greater number of electrons. The analogy is that a tsunami very well would toss a lot of debris ashore, but only *as far up the coast* as the faintest ripple on the surface of the water could project a tiny floating leaf.

This strange result was explained by Einstein's assumption that light essentially behaves as a swarm of particles, each carrying the very same quantum of energy as determined by the frequency of the light when they interact with the electrons in the target metal.<sup>254</sup> These particles were later named *photons*.

### 8.2.2. Wave-particle duality and the measurement problem

This, of course, directly implied the wave-particle duality issue, since light by more than a century of previous experiments had been shown to be a wave, which further on led to the quantum measurement problem wherein our main interest lies. Light can actually very easily be shown to display wave-like properties – if one looks at a small, point-like source of light, like a distant street lamp, through a tiny hole (around 0.1 mm) in some opaque object, e.g. a sheet of aluminium foil, a pattern characteristic of any wave being forced through a minute opening will be clearly visible – yet due to the work of Planck and Einstein we have to consider it particulate as well.

The measurement problem essentially hinges upon the fact that elementary particles, when measured upon, evidently present themselves as the definite, discrete entities assumed in Planck and Einstein's approaches, yet the pattern formed by these particles prove that between measurements, they behave exactly like waves.<sup>255</sup> In quantum mechanics, these patterns are interpreted as the effect of a *probability wave*, which in itself carries no energy in the way that everyday waves in matter do, but rather describes the probability

---

<sup>254</sup> Herbert 1985, p.37-38

<sup>255</sup> Ibid. p. 89; Perhaps obviously, the incongruity between these results lie chiefly in three wave attributes utterly incompatible with those of matter as classically conceived: waves can spread out over a vast area, while matter is confined to a comparatively very small one; a wave can be split in an infinity of ways, with different parts travelling in opposite directions, whereas material entities only travel in one single direction; and waves can pass through one another without difficulty, where material entities would collide.

that a particle or cluster of particles will be measured in a given position or with a particular momentum.<sup>256</sup>

The wave-particle duality is nowhere as evident as in the famous double-slit experiment, which establishes that light, projected towards a receptive detector screen (normally coated with phosphorous) through two slits in an obstructing plate, exhibits an interference pattern, while it's still observed as separate particles when the light engages with the detector screen.<sup>257</sup> An interference pattern is simply the result of the interaction of two or more waves, as when an ocean wave is split in two by openings in a barrier similar to those of the double-slit experiment, and these waves later physically interact in the water. Curiously enough, this interference pattern persists even when the detector screen only receives one single photon at a time. If we turn down the intensity of our projector to such a degree that only one of Planck and Einstein's discrete quantities pass through the perforated plate at a time, we nonetheless detect an interference pattern, implying that the single photon actually passes through both slits simultaneously, in a form similar to that of a wave. This state of the system (in this case a single-particle system) is described by the *probability wave function*, which normally is denoted  $\psi$ . The wave function is nothing else than the mathematical description of the system considered as a wave, which enables us to predict the system's future state by applying the Schrödinger equation.

However, if we modify the setup further, and place a detector at one or both of the slits in the plate, effectively determining whether a photon passes through either, the *interference pattern disappears*. This indicates that the act of determining the definite attributes of the quantity of light in question, somehow inhibits the photon's wave state which would have enabled it to effectively pass through both openings, and instead, in some sense, reduces it to a definite particle with matter-like behaviour.

The question of why we never observe a particle in what is known as the *superposition*, passing through both slits simultaneously, but always as a definite, discrete entity, is the very essence of the measurement problem. We must also note that these peculiarities are not, according to quantum theory, confined to the realm of the infinitesimal (the double-slit experiment has for instance been performed on enormous particles consisting of more than 800

---

<sup>256</sup> Henry P. Stapp, *Mindful Universe: Quantum Mechanics and the Participating Observer*, Berlin: Springer-Verlag 2011, pp. 26-27

<sup>257</sup> Brigitte Falkenburg, *Particle Metaphysics: A Critical Account of Subatomic Reality*, Leipzig: Springer-Verlag 2003, p 283

atoms, with similar results)<sup>258</sup>, but do indeed regard the larger objects of our everyday realities. This is often illustrated with Erwin Schrödinger's well known thought experiment involving some radioactive material and the unethical treatment of felines, the point of which being that unmeasured entities, no matter what their size is, according to quantum theory are described by the same wave function as the unmeasured photon discussed above. Nothing in quantum theory itself explains why measurements consistently yield definite results, nor why we actually experience a system described by the wave function in terms of the tangible, "classical" objects of our everyday reality. The standard explanation, what was originally considered the orthodox view,<sup>259</sup> modeled on Niels Bohr's original Copenhagen interpretation, in an attempt to explain this peculiar fact maintains that acts of observation somehow collapses the wave function, which in turn forces the system to assume one definite state among a multitude of possibilities. Accordingly, when measured at either opening, the photon discussed above must "choose" one of the two slits in the plate.<sup>260</sup>

### 8.2.3. Von Neumann, Stapp and the Copenhagen orthodoxy

So where is the photon when we're not measuring it? What does the probability wave function actually represent as passing through both slits simultaneously? Such questions were hotly debated in the early period of quantum theory around the 1920s, epitomized by the Bohr-Einstein debates.<sup>261</sup> About this time, Niels Bohr argued that we ought not speak of any hidden, objectively existing reality beneath what the wave function describes, that quanta actually are indeterminate between measurements, while Einstein's contrary, yet intuitively appealing position can be roughly summarized in his statement that

I don't believe that when I am not in my bedroom my bed spreads out all over the room, and whenever I open the door and come in, it jumps into the corner.<sup>262</sup>

Schrödinger's thought experiment previously mentioned was originally a similar attempt to emphasize the absurdity of Niels Bohr's Copenhagen in-

---

<sup>258</sup> Sandra Eibenberger et al, "Matter-wave interference with particles selected from a molecular library with masses exceeding 10000 amu", *Physical Chemistry Chemical Physics* 15, pp. 14696-14700

<sup>259</sup> The contemporary Copenhagen interpretation is commonly known as the "orthodox view" simply due to its historical position as the majority view

<sup>260</sup> Michael A. Parker, *Solid State and Quantum Theory for Optoelectronics*, Boca Raton: CRC Press 2010, p. 258-259

<sup>261</sup> Herbert 1985, p. 44

<sup>262</sup> Quoted in Hilary Putnam, "A Philosopher Looks at Quantum Mechanics (Again)", *British Journal for the Philosophy of Science*, Vol. 56, Issue 4, 2005, pp. 615-634

terpretation. It can be summarized as a hypothetical situation in which the outcome of a measurement equivalent to the double-slit experiment determines the ultimate fate of an unlucky cat trapped inside a box rigged with a lethal trap. The setup is such that there's a 50-50 chance that the wave function collapses in such a way that the trap is triggered, yet our only means of measuring the position of the relevant quantum particle is to actually open the box and find out what has happened to the cat inside. However, since not only the photon in question, but the entire experimental apparatus is a physical system described by the wave function, the indeterminacy of the unmeasured photon must according to the Bohr/von Neumann interpretation of quantum theory apply to the cat in the box as well. The unmeasured cat would thereby be both alive and dead until the observation determines its definite state, just as the photon in the double-slit experiment evidently passes through both slits simultaneously.<sup>263</sup> Of course, the obvious possibility that the cat would itself be capable of performing the relevant observation is excluded for the sake of argument.

John von Neumann, the renowned Hungarian mathematician, complicit in the development of nuclear weapons, game theory and modern computers, was the first to establish a comprehensive mathematical framework for quantum mechanics, presented in what's colloquially known as the "quantum bible", his *Mathematische Grundlagen der Quantenmechanik* from 1932.<sup>264</sup> In this work he additionally tackled the quantum reality question, providing a proof in favor of Bohr's position, while he with a hypothesis of his own also expounded upon the orthodoxy's implied claim that the observation inherent in the measurement act somehow initiates the collapse of the wave function.<sup>265</sup> This proof, although not entirely conclusive as we soon shall see, essentially showed that unmeasured entities cannot possess actual, innate attributes of their own, since the assumption of such attributes arguably violated the mathematical formalism of quantum theory.<sup>266</sup> Assuming the validity of von Neumann's proof, it must thus be emphasized that the indeterminacy of the unmeasured entities is metaphysical, and not merely epistemological. It would thereby be through the very event of observation that the entity in question acquires particular attributes.<sup>267</sup> Von Neumann's reworking of Bohr's position was based on the idea that observations are extraordinary events due to the fact that they involve a non-physical consciousness exter-

---

<sup>263</sup> Christopher Gerry, Kimberly Bruno, *The Quantum Divide*, Oxford: Oxford University Press 2013, pp. 151-152

<sup>264</sup> Herbert 1985, p. 25

<sup>265</sup> John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton: Princeton University Press 1955, pp. 417-437

<sup>267</sup> Jeffrey A. Barrett, "A Quantum-Mechanical Argument for Mind-Body Dualism" *Erkenntnis* 2006, 65:97-115

nal to the physical systems involved, and therefore are able to produce the collapse of an otherwise continuous wave function.

Given that physical systems according to quantum theory necessarily are subsumed under the wave function, and since definite states are only encountered in relation to observations, the physical systems themselves do not possess the ability to violate the wave function's linear progression according to which they evidently develop when unobserved. In other words, the probability wave should not be able to disturb itself. In accordance with this, von Neumann distinguished between two basic processes active in the world, which he called *process 1* and *process 2*. Process 2 is simply the orderly, mechanically controlled evolution of any physical system taking place between measurements, which is controlled by the Schrödinger equation,<sup>268</sup> while process 1 is considered to be

...the basic probing action that partitions a potential continuum of physically described possibilities into a (countable) set of empirically recognizable alternative possibilities.<sup>269</sup>

This means that process 1 selects a certain physical system out of many possible such systems in the world, and can be thought of as a kind of macro-measurement which produces a particular wave function (such as the one describing the wave in our double-slit experiment, passing through both openings), which then may be determined as having particular attributes. Neumann's position was that consciousness, considered external to the wave function, is necessary for the initiation of process 1, which in turn is a prerequisite for the emergence of seemingly definite attributes (the actual passage through one slit rather than the other). Henry Stapp, an American physicist and contemporary supporter of von Neumann's position, retains this terminology, yet introduces two more processes. He calls this "selection" of the outcome, of the definite passage, *process 3*, and specifies it as a statistically specified choice of the outcome of the probing action. Thus, if we elect to probe our single particle system in Schrödinger's experiment, the probabilistic process 3 simply determines whether or not the cat gets whacked. Stapp also introduces what he calls *process 0*, intending to complete the picture. Process 0 denotes the seemingly free choice of the particular consciousness which initiates process 1.<sup>270</sup>

Accordingly, on the Stapp/von Neumann description, we have:

---

<sup>268</sup> von Neumann 1955

<sup>269</sup> Stapp 2011, p. 24

<sup>270</sup> Ibid.

Process 0 – the “free choice” of the conscious experimenter on where and how to perform the probing action of process 1

Process 1 – the probing action which selects a certain physical system and initiates the collapse of the wave function

Process 2 – the evolution of the wave function of an undisturbed physical system

Process 3 – the “choice on the part of nature”<sup>271</sup> of the particular result of the collapse of the wave function, initiated by process 1

In effect, Stapp suggests that the automatic process 2, which as we have seen generates a set of classically incompatible events and does not self-determine, is reduced to a definite state via the statistically specified choice of process 3, brought about by the probing action of process 1, which in turn must be initiated by the external input of the conscious observer (process 0).<sup>272</sup>

This ordering of the processes could perhaps seem somewhat skewed, since the basic undisturbed system evolving according to the Schrödinger equation is denoted by the second process, yet we will for the sake of consistency utilize the terminology suggested by Stapp and von Neumann. As is likely obvious to the attentive reader, this set of processes is arguably beset by the problem of interaction. The connection between process 0 and process 1, given that the collapse-initiating consciousness cannot be an aspect of the physical system whose hypothetical states due to the collapse are rendered *actual* rather than *potential*, is at least on the face of it subjected to the very same difficulties as e.g. Cartesian interactionist dualism: if we assume that consciousness is a real entity other than a physical system, not subsumable under the wave function, we have essentially attained an ontological dualism. Moreover, if it's given that consciousness as an immaterial entity may exercise causative influence over the progression of a physical system (by disturbing the wave function) the criteria for interactionism is also satisfied.<sup>273</sup> However, we may note that the reinterpretation (or possible negation) of the physical ontological category as such, which is entailed by the Copen-

---

<sup>271</sup> This formulation was popularized by Paul Dirac, first stated in 1927. Cf. Shimon Mahlin, “The Collapse of Quantum States: A New Interpretation”, *Foundations of Physics*, vol. 23, issue 26, June 1993

<sup>272</sup> Stapp 2011, pp. 29-34

<sup>273</sup> The consciousness-causes-collapse theory can in principle be joined to other ontological positions than interactionist dualism. Since our interest lies in the particular approaches to the interaction problem which may be modeled on the theory, these possibilities will not be extensively addressed here.

hagen interpretation and the von Neumann/Stapp expansion thereof, may perhaps serve to reduce the severity of the ontological gap, thus lessening the impact of the interaction problem. This possibility will be discussed in detail at the end of the chapter.

The problem of interaction was not much discussed by von Neumann himself, and neither by his contemporary physics community. An influential article by Eugene Wigner from 1961 explores some of the ontological implications of von Neumann's position, but offers little in terms of the interaction problem, not to speak of any solutions thereof.<sup>274</sup>

The problem is almost – but not entirely – lost on Stapp as well. He proposes a somewhat ingenious approach to the issue of foreseeable effects of intentional acts, which can be indirectly related to the matter of interaction. He does, briefly, address the matter head on with a notion of intentional frequency-modulation of process-1 phenomena, but this suggestion isn't very solid, philosophically. As was previously stated, Stapp suggests that the automatic process 2, which generates a set of classically incompatible events and does not self-determine, is reduced to a definite state via the statistically specified choice of process 3, brought about by the probing action of process 1, which in turn must be initiated by the external input of the conscious observer (process 0).<sup>275</sup> The fact that process 3 is a probabilistic affair (we never know beforehand where the electron will hit the detector screen or which path the photon will take) would seem to pose a problem for the position that an external consciousness brought about any particular results specifically related to its intentions. For the purpose of this chapter, this probabilistic indeterminacy of the hypothetical effects of the observer's intentions is not an issue, since we're only interested in the very possibility of interaction as such. However, Stapp offers a reasonable quantum-theoretical solution to this particular problem in the application of something known as the *Zeno-effect*. Without going into any detail, the relevance of this effect can be summarized in that it enables a process 1-query to be intentionally repeated in rapid sequence, or insistently, by the hypothetical conscious observer, if e.g. the subsequent process 3-choice happens to be preferred by the observer. This increases the *likelihood* that the outcome of the process 3-event is established in the relevant system, making intentional influence logically possible with regard to quantum theory.<sup>276</sup>

---

<sup>274</sup> Eugene Wigner, "Remarks on the Mind-Body Question", I. J. Good (ed.) *The Scientist Speculates*, London: Heinemann 1961, pp. 284-302

<sup>275</sup> Stapp 2011, pp. 29-34

<sup>276</sup> Ibid, pp. 35-39

However, this has no impact on the problem of interaction as such, as the precise metaphysical relation between processes 0 and 1 remains unaddressed. Stapp's brief, direct treatment of the interaction issue emphasizes that his and von Neumann's approach does not imply that the mind acts directly upon physical matter via some force, but, discussing the relationship between mind and brain, instead claims that the effect

...is associated with a *modulation of the frequency* of certain process 1 actions that act directly upon large-scale patterns of neurological activity. This modulation of frequencies is achieved, strictly within the pragmatic framework (the orthodox Copenhagen model), by exploiting certain human 'free choices' that are allowed within that pragmatic framework.

This assumes that the observing consciousness can intervene from outside of the physical domain, and influence the collapse of the wave function, which in turn determines the attributes of what we experience as physical reality. We may note that this doesn't necessarily entail the strict mind-matter-interaction of classical substance dualism, but rather gives us a situation in which the subjective first-person perspective somehow influences the ontologically speaking extremely slippery superposition of a hypothetical entity. Stapp, favoring a Whiteheadian process metaphysics, argues that the interaction problem is not really an issue since all we have are processes rather than substances influencing one another, while he nonetheless advocates some form of dualism.<sup>277</sup> This feat of semantics seems to accomplish little in the way of a philosophically viable response to the problem, and Stapp's metaphysics aren't very clear in this regard. Notwithstanding which interpretation of quantum theory one favours, however, we seem forced to approach the reality behind the wave function as something quite different from the Newtonian billiard balls of classical mechanics.<sup>278</sup> But the notion that this fact alone can provide a basis for addressing the interaction problem seems doubtful. Additional, auxiliary approaches towards this end will therefore be examined later in the chapter.

#### 8.2.4. Non-orthodox approaches

As we saw before, not everyone agreed that this arguably counter-intuitive understanding of reality suggested by the Copenhagen account was the most preferable interpretative option. Indeed, Einstein, one of quantum theory's most important founders, persisted in his doubts for the remainder of his life,

---

<sup>277</sup> Ibid. p. 106

<sup>278</sup> Ibid. pp. 55ff



maintaining that a model wherein objective but hidden variables determined the measurement results could be coherently formulated. Before we move on, we must therefore address the alternatives to Bohr and von Neumann's old orthodoxy. The set of alternative interpretations which will be discussed during the remainder of the chapter is not intended to be entirely comprehensive. It purports to cover the majority of basic perspectives espoused within the contemporary physics community as summarized by Herbert in 1987, Tegmark in 1997 and Schlosshauer, Kofler and Zeilinger in 2013, with particular focus on the models that have positive or negative ontological impact with regard to such views of the human subject that conceptually support the metaphysics entailed by our solution to the problems of post-mortem survival.<sup>279</sup>

There are several basic interpretative models, the exact number of which depending on where one draws the line between closely related positions. Some distinguish between merely two fundamental interpretative positions; those that affirm the privileged role of the observer and those that do not.<sup>280</sup> It is also possible to distinguish between four basic positions – the accounts related to the influential Copenhagen interpretation; the opposing no-collapse theories; the objective-collapse theories; and the subjectivist positions.<sup>281</sup> Tegmark, for different reasons, focuses on five basic accounts, Herbert on eight and Schlosshauer et al on eleven.

All in all, there are many positions to choose from. In keeping with our historical outline and our purpose as such, an important set of positions to mention, aside from those related to the orthodox interpretation, are the early category of perspectives which are sometimes labeled *neorealism*, and arguably diminish the role of the observer, a variant of which was defended by Einstein in opposition to Bohr. Another early important position which enjoys some contemporary metaphysical influence is Everett's many worlds-interpretation or variants thereof, which deny both the observer's role and (ostensibly) the existence of any objective hidden variables.

---

<sup>279</sup> Herbert 1987, pp. 157-251; Max Tegmark, "The Interpretation of Quantum Mechanics: Many Worlds or Many Words?" *Fortschritte der Physik*, 46, 1998; Maximilian Schlosshauer, Johannes Kofler, Anton Zeilinger, "A Snapshot of Foundational Attitudes Toward Quantum Mechanics", *Studies in the History and Philosophy of Modern Physics* 44, 2013

<sup>280</sup> Adán Cabello, "Interpretations of quantum theory: A map of madness" <http://arxiv.org/abs/1509.04711v1>, 2015, retrieved 2016-07-20

<sup>281</sup> Dimitris Lazarou, "Interpretation of Quantum Theory – An overview", <http://arXiv:0712.3466v2>, 2009, retrieved 2016-07-20

### 8.2.5. Neorealism and hidden variables

The neorealist perspectives (e.g. Bohm's pilot wave model; time-symmetric theories and stochastic models) basically maintain that the quantum world, in spite of every indication otherwise, nonetheless is populated by some type of "ordinary objects" (ordinary in the sense that they have definite attributes even when not measured upon) rather than the ephemeral non-objects of the orthodox interpretation. Einstein, Podolsky and Rosen's suggested hidden variables would be one example of such ordinary objects, determining the outcomes of empirical experiments.

The American physicist David Bohm was in 1952 the first to successfully formulate such a neorealist interpretation of quantum theory, while, however, Frenchman Louis de Broglie had made a similar although ultimately incomplete attempt some 25 years earlier.<sup>282</sup> In essence, Bohm's solution to the measurement problem was to partition the ephemeral quantum entities into actual and definite particles and waves, contrary to the orthodox position which instead considers the particle and the wave distinct attributes of a single substance. Upon this, he also added a novel mechanical element which he termed the *pilot wave*. This pilot wave functions as a kind of tentacle in relation to the surrounding environment, which senses any changes occurring anywhere in the world, and then responds to these by modifying its attributes. Moreover, the pilot wave communicates the information of these changes to the electron, which in turn also reacts to the event. Thus, the pilot wave, and thereby the quantum particle in question, will actively respond to your measurements by a corresponding change in their attributes. These attributes are on the Bohmian model innate to the particle and the pilot wave, but seem to be contextual and randomly changing, since the omnipresent pilot wave renders them immediately responsive to any change in the total environment of the entire universe, including your intent to make a particular measurement (if such an intent is taken to be causally efficacious).<sup>283</sup>

The model undermines the fundamental role ascribed to the observer on the von Neumann/Stapp position, implying that observations do not necessarily have causal effects. It's globally determinist, and assumes the existence of hidden variables (the pilot wave) and denies the collapse of the wave function which is integral to many other ontological interpretations of quantum

---

<sup>282</sup> B.J. Hiley, "Some Remarks on the Evolution of Bohm's Proposals for an Alternative to Standard Quantum Mechanics", preprint, 2010, retrieved from: [www.bbk.ac.uk/tpru/BasilHiley/History\\_of\\_Bohm\\_s\\_QT.pdf](http://www.bbk.ac.uk/tpru/BasilHiley/History_of_Bohm_s_QT.pdf) 2014-11-24; Herbert 1985, p. 49.

<sup>283</sup> Gerard Gouesbet, *Hidden Worlds in Quantum Physics*, Dover Publications, New York, 2001, p. 168-170

theory. Moreover, it provides a tentative solution to the Schrödinger paradox, and evades wave-particle duality.

In a somewhat ironic twist, however, Bohm's original neorealist position actually necessitates a move towards another form of non-realism. Bohm's model, in the omnipresent pilot wave's immediate reaction to any and all changes in the universe, does not function without some equivalent to faster-than-light signaling or the negation of all actual spatial distance. This was early on taken to be a significant drawback of the model: FTL-signaling violates special relativity, and the denial of actual distance seemed problematic for other reasons. As we shall soon see, this problem is far from fatal, although there are other issues with the neorealist interpretation which tends to impair its plausibility as an alternative to the orthodox position.

Again, Bohm's model cannot work without faster-than-light signaling or the denial of physical distances. This was thought to be a devastating flaw of the hidden variable pilot wave-solution, until Irish physicist John Stewart Bell proved that either superluminal connections, or the ontological negation of all actual spatial distance, given certain uncontroversial assumptions necessarily must be a feature of *any* model of physical reality anchored in empirical fact, a position which since has been conclusively experimentally verified.<sup>284, 285, 286, 287</sup>

### 8.2.6. Bell's interconnectedness theorem and non-locality

Initially, Bell was examining how it was possible for Bohm to produce a model of the electron endowing it with objective attributes in the unmeasured state, when von Neumann had proved that precisely this type of interpretation was logically impossible on quantum theory. Simply put, a loophole in the proof, which had been used by Bohm, was discovered. In the mid-1960s, Bell took it upon himself to examine exactly what this loophole looked like, and how it was exploited by Bohm's model. Basically, what von Neumann's proof prohibited, was the reproduction of the empirical results of quantum theory on any model which combined ordinary objects "in reasonable ways", which incidentally excluded electrons adjusting their attributes

---

<sup>284</sup> Marissa Giustina, Alexandra Mech et al, "Bell violation using entangled photons without the fair-sampling assumption", *Nature*, 497, pp. 227-230, 2013

<sup>285</sup> Markus Ansmann, H. Wang et al, "Violation of Bell's inequality in Josephson phase qubits", *Nature* 461, pp. 504-6, 2009

<sup>286</sup> M. A. Rowe, D. Kielpinski et al, "Experimental violation of a Bell's inequality with efficient detection", *Nature* 409, pp. 791-794, 2001

<sup>287</sup> S. J. Freedman, John F. Clauser, "Experimental test of local hidden-variable theories", *Physical Review Letters* 28, 938, 1972

via an omnipresent, invisible field instantaneously reacting to any changes made in the configuration of measuring devices.

Von Neumann's proof was thus never completely decisive. More exactly, it assumes that a *linear operator* is associated with every measurement situation, and that any result which can ever be measured is what is known as an *eigenvalue*, possessed by this linear operator. The proof shows that on these assumptions, no further hidden variables, determining the attributes of the measured particles can exist, implying the completeness of quantum theory.<sup>288</sup> However, Bohm could circumvent this proof by the rather bold assumption that the *operators need not always correspond to the observed properties of measured objects*, which in turn opens for the possibility of the arguably unreasonably context-sensitive behaviour of particles that his pilot wave-model assumes.<sup>289</sup>

Having clarified the issue, Bell set out to ascertain whether an argument could be made towards specific constraints on any model of physical reality whatsoever, with the aid of nothing but hard empirical data, rather than the somewhat less robust edifices of quantum theory which were assumed in von Neumann's case. This ultimately led him to what today is known as the Bell interconnectedness theorem, which most accurately can be described as a family of results ostensibly showing the impossibility of a local realistic interpretation of quantum mechanics, and indeed, any such model of physical reality.<sup>290</sup> Bell's original paper from 1964 was basically a response to an argument put forth in the 1930s by Albert Einstein and two other prominent physicists, Nathan Rosen and Boris Podolsky. This argument, known as the EPR-paradox, purported to show that a hidden-variable model was preferable to the orthodox account due to the latter entailing certain counter-intuitive results, which would imply that quantum mechanics is an incomplete physical theory. The EPR-paradox is thereby effectively an argument against quantum theory.<sup>291</sup>

---

<sup>288</sup> John F. Clauser, "Early History of Bell's Theorem", in N. P. Bigelow, J. H. Eberly et al (eds.), *Coherence and Quantum Optics VIII: Proceedings of the Eight Rochester Conference on Coherence and Quantum Optics*, New York: Kluwer Academic/Plenum Publishers 2003, pp. 23-24

<sup>289</sup> Dürr, D., Goldstein, S., Zanghì, N., "Quantum Equilibrium and the Role of Operators as Observables in Quantum Theory", *Journal of Statistical Physics* 134, 2009, pp. 959-1055

<sup>290</sup> Shimony, Abner, "Bell's Theorem", *The Stanford Encyclopedia of Philosophy*, (Winter 2013 Edition), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2013/entries/bell-theorem> 2014-12-08

<sup>291</sup> John Stuart Bell, "On the Einstein Podolsky Rosen Paradox", *Physics* vol. 1, no 3, 1964, pp. 195-200

The argument regards a certain class of quantum attributes ascribed to entities in what is known as an *entangled state*, which due to this fact are predictably correlated in a certain fashion. So if particle P and particle Q together make up an entangled system, described by the same Schrödinger equation, a measurement on a certain attribute of particle P rendering  $A(x)$  will thereby inform us that  $A(y)$  simultaneously holds for Q. The important thing is that we don't have to measure any attributes of the second particle to find this out, since the correlation is necessary. This implies that if particle P is separated from particle Q by a distance equivalent to that between the sun and Alpha Centauri, we nonetheless instantly determine the conjugate attribute of one of these by measuring that of the other. To Einstein and his associates, assuming what is known as *local realism*, this implied that the attributes in question must have been there all along, or more accurately, that they were determined by objectively present hidden variables unknown to quantum theory.<sup>292</sup>

In other words, the fact of this type of correlation was according to the EPR-team irreconcilable with the orthodoxy's claim that the attributes of quantum entities do not actually exist prior to measurement, albeit on the assumption of local realism. The locality-part of the assumption is in this case simply the position that no physically separated systems may causally interact with one another immediately, i.e. with a speed above that of light, or alternatively, that spatial separation actually exists with regard to physical systems; that the distance between distinct objects is in reality there. The reality aspect is within this context the assumption of *counterfactual definiteness*, the position that there would be some fact of the matter as to the results of any quantum-level measurement that actually was not performed, e.g. as with any hidden-variable theory.

In 1964, Bell devised a hypothetical experiment intended to once and for all determine whether local realism actually held. The outcome of the first such experiment, performed in 1972, indicated that it did not.<sup>293</sup> This result has been verified in every subsequent empirical investigation since, with the last hypothetical loophole for photons finally closed in 2013.<sup>294</sup> Local realism is thereby incompatible with all currently available empirical data, with the consequence that it must be regarded as a supremely untenable position.

The proof itself is rather simple, and is based on an experimental setup similar to the one discussed in the EPR-argument. We accordingly begin with a

---

<sup>292</sup> Dipankar Home, Andrew Whitaker, *Einstein's Struggles with Quantum Theory: A Reappraisal*, Springer Science: New York 2010, pp. 121-123

<sup>293</sup> Freedman & Clauser 1972

<sup>294</sup> Giustina & Mech et al, 2013

light source producing pairs of photons in the entangled state, yet for the Bell-experiment, they travel in opposite directions towards detectors, measuring an entanglement attribute collectively belonging to the two photons, called polarization correlation. This type of correlation is determined by setting the detection devices at different angles, producing an either-or correlation measurement for the property of polarization at a certain angle between the two receptors, denoted  $\theta$ . So if we are to measure the polarization correlation, or PC, of a certain photon pair, we will set detector A at angle X and detector B at angle Y, respectively, and then compare whether the binary polarization readings match. So if the polarization value for A and B both are either 0 or 1, they match, otherwise, they do not. It must be noted that according to quantum theory, the polarization correlation reading only depends on the relative angle  $\theta$  between the detectors, rather than the specific values of X or Y.<sup>295</sup>

Thus, a PC-measurement for a particular angle  $\theta$  is concerned with the fraction of matches registered with a series of emitted photons. If the detectors are set at the exact same angle, all measurements yield matches, with  $PC = 1$ . If the separation is great enough, no matches are detected, with  $PC = 0$ . At intermediate values of  $\theta$ , the polarization correlation varies between 0 and 1. Bell's theorem regards what exactly happens with this fraction of matches when we alter the relative angle  $\theta$  between the detectors.<sup>296</sup>

Given that the PC varies between 1 and 0 in relation to the relative angle, there is a particular value of  $\theta$  where the  $PC = \frac{3}{4}$ , a value where we on average thus register a match for every three out of four measurements. Let us say that an altering of the angle of either detector for  $\phi$  degrees yields this particular value of  $\theta$ . So if we either move detector A or detector B for  $\phi$  degrees, we get a PC equal to  $\frac{3}{4}$ .

Here's where Bell's proof comes in. On the assumption of local realism, nothing that happens at detector A can alter the results registered at detector B, or vice versa. This assumption, however, leads to a prediction which is contrary to empirical data. Given the above, altering both detectors  $\phi$  degrees, with a total misalignment of  $2\phi$  degrees, the total error rate cannot on the assumption of local realism be greater than two misses out of every four measurements. This prediction is an example of what is known as a *Bell inequality*.<sup>297</sup> Astonishingly, this particular Bell inequality has been conclu-

---

<sup>295</sup> Herbert 1985, p. 215-216

<sup>296</sup> S. M. Blinder, *Introduction to Quantum Mechanics: In Chemistry, Materials Science and Biology*, Elsevier Academic Press, Burlington 2004, pp. 273-274

<sup>297</sup> Douglas Hemmick, Asif Shakur, *Bell's Theorem and Quantum Realism*, Springer, Heidelberg, 2002 pp. 51-52

sively violated in actual experiments, yielding a consistent error rate of three misses on every four measurements if the misalignment of the detectors is of  $2\varphi$  degrees, a result which is ostensibly impossible to obtain in the absence of non-local effects.<sup>298</sup>

The verdict is thus that our reality, according to all available data, cannot be accurately modeled on the assumption of local realism. But what does this really mean? In fact, it's not certain which aspect of local realism we are to surrender. If we accept that superluminal connections are possible, or that actual physical distances do not exist, disregarding the locality-part of the assumption as defined previously, the contradiction is clearly eliminated. But we can also attain the same goal by dropping realism according to the stated definition, and deny the counterfactual definiteness that hidden variables provide, implying that Einstein's bed actually does jump into the corner whenever he opens the door to his hotel room. The question of which solution is the most preferable has not been satisfyingly answered, nor very thoroughly explored in research, yet an admission of superluminal connections poses such problems for the special theory of relativity that a jettisoning of realism in the above sense rather than locality is often preferred.

However we choose, there are certain implications for our metaphysical options, as well as, indirectly, for the problem of interaction.

If we embrace superluminal information-transfer as a universal phenomenon, a significant aspect of the assumed basis for the interaction problem is indirectly undermined, since local, mechanical interaction can no longer be considered a necessary prerequisite for causal effects. On non-locality, direct interactions may still be retained as a prerequisite for causal relations, since there really are no separate systems even though they appear to be. However, since every system in the universe must be considered locally and immediately interconnected in spite of being seemingly separated, a disregard of transsubstantial causal relations merely due to the apparent lack of some equivalent to local mechanistic connections becomes problematic. Admittedly, interaction between separate physical systems are ontologically different from transsubstantial interaction, yet if our reason for disregarding the latter simply is the absence of any obvious mechanical interaction, it is undermined by the assumption of non-locality.

---

<sup>298</sup> Cf K. Edamatsu, "Generation of entangled photons in semiconductors", in Benoît Deveaud, Antonio Quattropani, Paolo Schwendimann, *Quantum Coherence in Solid State Systems*, Società Italiana di Fisica: Bologna 2009, p. 21ff; A. J. Duncan, "Two-Photon Decay of Atomic Hydrogen", in H. Narumi, I Shimamura (eds.), *Atomic Physics 10*, Elsevier Science Publishing Company: New York 1987, p. 134-139

If we, on the other hand, were to drop the possibility of hidden variables and remove the realism-aspect of the pair, we're faced with the want of an explanation of the actual fact of definite measurement results (of why we encounter no Schrödinger cat-zombies), which is most obviously addressed with the aid of some form of metaphysical pluralism (which in turn presumes interactionism to be able to actually explain the fact of definite measurement results). This latter implication will be discussed in detail in the end of the chapter.

So given the consistent violation of the Bell inequalities which are independent of any particular interpretation of quantum theory, the two possible metaphysical implications both indirectly tend towards enabling an affirmation of phenomenalist metaphysics, either by undermining the problem of interaction or by necessitating the existence of extra-physical levels of reality which in turn presupposes a form of interactionism.

### 8.2.7. Criticisms of Bell's theorem: superdeterminism, the time-symmetric theories and the transactional interpretation

Nonetheless, the disavowal of local realism does not reign entirely supreme in the physics community. A critique of Bell's interconnectedness theorem based on the possibility of global determinism was early on put forth, and the backwards causation introduced by the so-called time-symmetric theories, as well as the transactional interpretation of quantum mechanics arguably also allow us to possibly keep local realism.

The essence of the global determinist critique, based in a position which is often called *superdeterminism*, is that local realism can be saved if we assume that every event in the universe by necessity is strongly predetermined, including the experimenter's choice of parameters with regard to the measurement of mutually exclusive quantum attributes.<sup>299</sup> To clarify, a fundamental aspect of quantum theory is Heisenberg's uncertainty principle. This principle describes the empirical fact that certain correlated attributes of quanta, known as complementary variables, relate in such a way that the extent to which we know one variable, limits knowledge of the other. For instance, if we know a photon's position with perfect accuracy, its momentum becomes impossible to ascertain. In line with the uncertainty principle, Bell's theorem assumes that with or without any hidden variables, when having measured the momentum of the photon in question, I *could* in actuality have measured the position instead, i.e. that there really would have been a position value resulting from such a counterfactual measurement. Superde-

---

<sup>299</sup> Gerard 't Hooft, "The Free-Will Postulate in Quantum Mechanics", *arXiv*, 2007, retrieved from: [arxiv.org/abs/quant-ph/0701097](http://arxiv.org/abs/quant-ph/0701097), 2014-12-07



terminism simply entails that there's no choice, and therefore never was any possibility of attaining position value, since my act and the outcome of the measurement was entirely predetermined from the beginning.

The implication of superdeterminism is accordingly that the universe already "knows" beforehand what particular attribute the experimenter was going to choose to measure, with the consequence that particle P and particle Q in the example above, from the very beginning of time was predetermined to be apprehended from a certain point of view, and thus never could have yielded different measurement results, nor even actually examined in another manner.<sup>300</sup> Bell's theorem on the other hand assumes that alternative results actually are conceivable, i.e. assumes the possibility of basic counterfactual definiteness of local hidden variables which is necessary for the Bell inequalities to obtain (if I had been measuring the position rather than the momentum, I would have received some particular position value due to the influence of local hidden variables). However, strong global determinism actually eliminates counterfactual definiteness by virtue of the hidden variables, specifically due to their superdeterministic character. Thus, no such alternative, counterfactual measurement results would ever have been possible.

The assumption of superdeterminism would eliminate the need for superluminal connections and/or the particular kind of non-realism discussed above, since, again, both the attributes of the particles and the intentions of the experimenter was predetermined locally from the very beginning of the universe's existence. Bell responded to this kind of criticism in 1985 by pointing out that the hypothetical type of hidden variables in question would need to have their initial conditions set exactly right, working together in an incredibly elaborate conspiracy inducing the false appearance that the Bell inequalities were violated in every test throughout the history of the universe, so that it would falsely appear to the observer that superdeterminism did not hold.

In other words, there's no obvious reason that the inequalities consistently would actually *seem* to be violated even on superdeterminism, implying the strangest sort of conspiratorial coincidence, even though it's logically possible in a strict sense. This imbues the superdeterminist critique with a decidedly ad hoc-character.<sup>301</sup> In effect, since even particles created at the opposite ends of the known universe on a deterministic local-realist model would have to contain the complete genetic information regarding each other (as

---

<sup>300</sup> Paul Davies, Julian Brown, (eds.) *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics*, Cambridge University Press: Cambridge 1986 pp.45-46

<sup>301</sup> John Stuart Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, New York 2004, ch. 12

well as every other particle in existence), superdeterminism necessitates the existence of an extremely complex hidden internal structure capable of carrying and conveying this information, which is something we currently have no reason to suspect actually is present, and which might be metaphysically impossible.

An additional way of evading Bell's theorem's rejection of any model of reality based in local realism (the assumption of locality together with the presence of hidden variables) is, again, to introduce the notion of retrocausality. Assuming backward causation, the violations of Bell inequalities can then arguably be interpreted as local effects, albeit working backwards all the way down to the emission of the entangled pair of particles.

The main positions implying retrocausality are the so-called "time-symmetric theories", another type of interpretation of quantum mechanics first proposed by Satoru Watanabe in 1955.<sup>302</sup> Similar positions have also been suggested by John Wheeler and Richard Feynman, and more recently, John G. Cramer with his transactional interpretation of quantum mechanics.<sup>303</sup> The obvious problem of this kind of interpretation is its introduction of the intuitively suspect, arguably paradoxical notion of backward causation, which also ostensibly contradicts experimental results supporting irreversibility in physics. The time-symmetric theories (of the Watanabe-Wheeler-Feynman type) are also difficult to reconcile with other important empirical data, specifically the necessity of assuming a particle's interaction with its own emitted field in accounting for certain relativistic effects.<sup>304</sup> However, it has been argued that a time-symmetric account could possibly recover local realism without resorting to retrocausality.<sup>305</sup> All in all, these positions currently seem to be minor, with no respondent in three recent surveys of the interpretative preferences of the physics community claiming to espouse them.<sup>306</sup>

---

<sup>302</sup> Satoru Watanabe, "Symmetry of physical laws. Part III. Prediction and retrodiction." *Reviews of Modern Physics* 27.2, 1955

<sup>303</sup> John Cramer, "The Transactional Interpretation of Quantum Mechanics", *Reviews of Modern Physics* 58, 1986

<sup>304</sup> Ruth Kastner, *The Transactional Interpretation of Quantum Mechanics*, Cambridge University Press: New York 2013, p. 199-200

<sup>305</sup> Huw Price, "Does Time-Symmetry Imply Retrocausality? How the Quantum World Says 'Maybe'", *Studies in History and Philosophy of Modern Physics* 43, 2012

<sup>306</sup> Schlosshauer et al, 2013; Christoph Sommer, "Another Survey of Foundational Attitudes Towards Quantum Mechanics", <http://arxiv.org/pdf/1303.2719v1> 2013. retrieved 2016-07-20; Travis Norsen, Sarah Nelson, "Yet Another Snapshot of Foundational Attitudes Toward Quantum Mechanics", <http://arxiv.org/abs/1306.4646v2> 2013, retrieved 2016-07-20

### 8.2.8. Remaining alternatives to the old orthodoxy

Aside from Bohm and de Broglie's neorealistic hidden-variable account and the time-symmetric theories, there are a few other models that also deny the wave-function's collapse and diverge from Bohr's orthodox account.

The basic premise of the no-collapse position is that the wave function never actually collapses into a definite state, and thus that an equivalent of Stapp's process 2 described above is everything that ever occurs in the physical universe.<sup>307</sup> One of the main difficulties with this class of perspectives beyond what has already been discussed regarding Bohm's pilot wave model and the positions related to time-symmetric theories, is that the no-collapse position has severe conceptual problems explaining the definite measurement results which are phenomenally available to the experimenter (in lieu of an actual collapse of the wave function). Stapp argues that we on a no-collapse model of reality ought to experience a "smearing" of the possible measurement results equivalent to the probability distribution of the wave function, and that the fact that we do not, begs an explanation.<sup>308</sup> This is also known as the "preferred basis-problem", i.e. the problem of the apparent lack of a basis for the definite measurement results relative to the superposition described by the wave function.

Everett's many-worlds account is, again, a prominent position among the other no-collapse interpretations. This model evades the measurement problem illustrated by Schrödinger on the assumption that every possible outcome of the experimental measurement actually takes place, and that reality as such branches into many worlds all simultaneously actually existing, while we ourselves however only experience one of the outcomes. Yet, the preferred basis-problem remains. As Stapp contends, here also the divergence between our limited phenomenal reality and the actual realization of all possible outcomes does indeed beg an explanation.

Another minor no-collapse theory is the many minds-account advanced by Heinz-Dieter Zeh in 1970, as a suggested variant of Everett's interpretation.<sup>309</sup> This model entails that the mind, rather than the world as such, branches into many causally efficacious observers.

Beyond the no-collapse theories, there are finally interpretations that alongside the Bohr/von-Neumann account acknowledge the wave function's col-

---

<sup>307</sup> D. Z. Albert, B. Loewer, "Interpreting the Many Worlds Interpretation", *Synthese* 77, 1988, pp. 195-213

<sup>308</sup> Stapp 2011, p. 62-63

<sup>309</sup> Heinz-Dieter Zeh, "On the interpretation of measurement in quantum theory", *Foundations of Physics* 1, 1970, pp. 69-76

lapse, yet nonetheless diverge from the orthodox account. The three most prominent are the transactional account, which has already been mentioned in relation to the critique of Bell's theorem, the objective-collapse theories, and the relational account.

The relational account is to a great extent a development of the Bohr/von Neumann orthodoxy, and is therefore ontologically quite similar. It clearly underwrites the non-materialist or non-reductionist metaphysics of the person amicable to post-mortem survival, as the causal efficacy of the observer's intentions is fundamental to the relational interpretation: all states are observer-dependent.<sup>310</sup> According to the relational account, observer-independent states are actually entirely inconceivable, echoing bishop Berkeley.

The objective collapse-theories are also related to the Copenhagen interpretation, in that a collapse of the wave function is presupposed. Some objective collapse-theories, such as Roger Penrose's model are functionally similar to the orthodox interpretation with regard to the observer's possible role, and mainly differ from the orthodox account with regard to the ontological status of the wave function and its collapse.<sup>311</sup> Other objective collapse-accounts, such as the one advanced by Giancarlo Ghirardi, Alberto Rimini and Tullio Weber regard the collapse as independent of any intervening observer, and thus as *objective* in a further sense. According to the Ghirardi-Rimini-Weber account, the wave function spontaneously collapses, although extremely rarely – the rate at which an individual particle emerges from a wave-function collapse would arguably be around the rate of once every hundred million years, according to their theory.<sup>312</sup> Spontaneous-collapse accounts have been criticized for being less amenable to general and special theory of relativity when compared with the orthodox account and related positions.<sup>313</sup>

### 8.3. Consensus and consequences

With all these quite exhilarating avenues of interpretation and experimental results, and arguably little in the way of any final arbitration between the interpretative models, what can we actually make of the situation? Is there

---

<sup>310</sup> Carlo Rovelli, "Relational quantum mechanics", *International Journal of Theoretical Physics* 35, 1996, pp. 1637-1678

<sup>311</sup> Cf. Roger Penrose, *The Emperor's New Mind*, Oxford University Press: Oxford 1999, pp. 475-481; Roger Penrose, Stuart Hameroff, "Consciousness in the universe: A review of the Orch-OR-theory", *Physics of Life Reviews* 11, 2014

<sup>312</sup> Bell 2004, pp. 201-212

<sup>313</sup> Stapp 2011, p. 63; Philip Pearle, "Quasirelativistic quasilocal finite wave-function collapse model", *Physical Review A*, 71, 2005

any tangible consensus among researchers in the relevant fields, and what philosophical conclusions relevant to our purpose may be drawn from this plethora of positions, with or without such a consensus?

First and foremost, it must be stated that none of the interpretations of quantum mechanics *exclude* the category of ontologies amenable to post-mortem survival as discussed in the previous chapters. That is, even as roughly half of the interpretative models that are commonly suggested by researchers and philosophers active in the field do not ascribe the observer any privileged role, they do not by necessity exclude immaterialist or non-reductive ontologies – nor even the possible causal efficacy of the observing subject. These models simply imply that one does not need these features to square one's picture of reality with quantum physics as a theory of material reality. On the other hand, the remaining interpretative options actually maintain that the privileged status of the observing subject is either absolutely necessary for our description of reality to function in concert with quantum theory, or that such a privileged position at least can be naturally factored into the model.

Whether or not there is any meaningful consensus to speak of regarding favoured interpretations is currently difficult to say, as there have been no really comprehensive inquiries made on this subject. The surveys that are available and also relate to the relevant populations all have way too few respondents, together with other methodological issues, which make it impossible to base anything too conclusive on them.

What can be said, however, is that the Copenhagen orthodoxy, which most clearly enabled the privileged role of the observer, is probably still a major contemporary position in the physics community. This is likely in part due to its earlier, almost unchallenged reign which lasted into the 1980s,<sup>314</sup> and can be seen in the fact that it placed first or second in two of the three surveys available (albeit fourth in the third).<sup>315</sup> Moreover, all three surveys, somewhat surprisingly, entail that a significant portion of the respondents acknowledge the ontological distinctiveness of the observing subject. In the survey wherein the orthodox account placed most poorly, Travis Norsen and Sarah Nelson's repetition of Schlosshauer, Kofler and Zelinger's survey at a different quantum physics conference, 25% of the respondents nonetheless maintained that such an ontological distinctiveness is real. In the other two surveys, an overwhelming majority stated the same position.

---

<sup>314</sup> John Gribbin, Mary Gribbin, Jonathan Gribbin, *Q is for Quantum: An Encyclopedia of Particle Physics*, Simon & Schuster: New York 2000, pp. 87-90

<sup>315</sup> Schlosshauer et al 2013; Sommer 2013; Norsen and Nelson 2013

Obviously, these results cannot be taken as representative for the vague category of “the physics community”, nor even some more clearly delineated subset thereof. Nonetheless, they show that the orthodoxy is still a live option within this community. Moreover, they make it plausible that the ontological distinctiveness of the observing subject, as von Neumann maintained must be a necessary feature of reality if we are to take quantum theory seriously, is accepted by many researchers working in the field.

Thus far, it seems obvious that, at the very least, a weak criterion of scientific compatibility is satisfied insofar that fully coherent and oft-endorsed interpretations of quantum theory are compatible with, or even imply the metaphysical positions that are amicable to a post-mortem survival. However, the argument can be made a bit stronger with regard to the implicit undermining of the problem of interaction that is entailed by certain aspects of basic quantum theory as well as most explanations of the Bell inequalities, and also, the possible avenue towards immaterialist ontologies from several of the non-orthodox interpretative models as well.

### 8.3.1. Ontological implications

We have already seen that according to the orthodox interpretation, the quantum object in the wave state has no definite attributes of its own and cannot as such be located in ordinary space-time. It has no extension in the normal sense of the word, as the wave state carries no physical energy, but rather represents something that best can be described as the probability distribution of possible experiential data in relation to space-time, i.e. the conscious experiences of definite attributes. As far as we know, according to the orthodox interpretation of quantum theory and several other associated perspectives, this is indeed the very essence of ordinary matter, yet the description as such is vastly different from the understanding of matter inherent to classical physics.

Thus, assuming the orthodox, or the orthodox-like interpretations of quantum mechanics, the proponent of interactionist dualism and/or non-reductive materialist ontologies seems to have certain tools at her disposal unavailable in relation to mere classical Newtonian physics. Most obviously, the issue of the metaphysical gap is, at least on the face of it, rendered less severe by postulating interaction between a non-physical consciousness and the ethereal wave function rather than between Descartes’ *res cogito* and *res extensa*. The often cited problem with direct interaction in the violation of the conservation of energy also seems to lose much of its momentum, as no energy necessarily needs to be transmitted for the modulation of Stapp’s process 1 actions and the collapse of the probabilistic wave function.

The interactionist is also provided with the opportunity of posing an indirect argument for interaction by the arguable necessity of an external trigger for the hypothetical collapse of the wave function as such. As previously stated, quantum theory subsumes every physical system in existence (indeed, the entire physical universe) under the wave function. These systems are in principle unable to realize themselves as definite entities by initializing any collapse, which is why something external to process 2, the totality of physical systems described by the wave function, is necessary for the emergence of the definite attributes of physical reality. If it can be made plausible that we need this external trigger, two important implications follow. First, we have a clear argument for some form of immaterialist metaphysical pluralism, in the non-identity of the external trigger and the wave function as representing the totality of all interconnected physical systems. Moreover, we then have indirect support for interaction, since it may be further contended that quantum theory demands the actual interaction of the physical system in the superposition, and something external which is not reducible to such a system, whether or not we can explain the actual causal relationship at hand.

In relation to the preceding chapter; for such a model of interaction to be a viable option in addition to those previously referenced, it must be shown that it does not collapse into monism or a proxy approach, and that it does without modifying the notion of interaction to such an extent that we can no longer distinguish a causal relationship. However, it seems obvious that the notion of interaction inherent in the consciousness-causes-collapse model does not need to violate our commonly held notions of causation as such. It is in principle compatible with the counterfactual and the probabilistic theories discussed in the preceding chapter, since e.g. Stapp's process 0 without contradiction can be conceived of as a necessary prerequisite for the collapse, as well as considered a probability raising factor in relation to the collapse as an outcome. The covering law theory also seems viable in this context (yet needs to assume some form of lawlike regularity with regard to process 0 and the wave function collapse) and the primitivist option obviously poses no problem either.

It's however not entirely clear that quantum-mechanical interaction of the consciousness-causes-collapse variant can evade either slipping into the territory of an immaterialist monism such as classical idealism, or becoming a version of the somewhat problematic proxy approach.

If we consider the wave function to be ontologically distinct from determinate physical reality as we experience it, the ontological gap between the wave function and a non-physical mind is possibly reduced. Still, this discrepancy is thereby transferred to the relationship between the wave function and determinate physical reality, and this notion of interaction then faces

similar problems as the proxy approach: even if interaction between the non-physical consciousness and the ethereal wave function is conceivable, how can the equally non-physical wave function “interact” with substantial physical reality by causing its emergence? A possible line of argument could be to insist that while interaction between ontologically distinct substances is indeed problematic, the notion of one distinct substance *causing the emergence* of another isn’t necessarily beset by the same difficulties. Regarding the possible collapse into monism, quantum-mechanical interaction of the consciousness-causes-collapse type will risk turning into some form of idealism if the non-physical aspect of reality is taken as ontologically basic. Thus, if in the interest of saving interaction, we emphasize the non-physicality of the wave function as the fundamental nature of matter, with the implication that the substantiality of the matter and its ostensibly definite attributes are regarded as non-independent and ephemeral, some form of idealism seems impossible to avoid. The same result seems likely if we instead were to utilize the metaphysics hinted at by Stapp and try to avoid the ontological gap by considering the interacting parties something like processes rather than anything else. Such a strategy is clearly also limited to either dismissing the aspects of reality we normally consider substantial and determinate as non-independent and ephemeral, or being forced to face the transferred ontological discrepancy between definite matter and the ethereal wave function mentioned above.

### 8.3.2. The unorthodox path towards immaterialism

It seems clear that the hypothetical collapse of the wave function can be utilized to support immaterialist ontologies, and that the implications of Bell’s theorem provides further lines of argument towards this end.

However, the American philosopher Jeffrey Barrett has shown that even independently of this approach, quantum theory seems to lead us away from reductive physicalism. In a 2006 article, further developed and defended in 2014, Barrett argues that the mere persistence of definite measurement results can be taken to imply some form of substance dualism or pluralism, even on the no-collapse interpretations of quantum theory.<sup>316</sup> The gist of his argument is that given certain central, yet uncontroversial explanatory constraints, any interpretation of quantum theory seems to imply a pluralist metaphysics. Again, as we have seen, the Copenhagen orthodoxy clearly takes us in this direction, and the von Neumann-account obviously entails an immaterialist metaphysics. What’s perhaps not as obvious is that several of the

---

<sup>316</sup> Jeffrey A. Barrett, “Quantum Mechanics and Dualism”, in A. Corradini and U. Meixner (eds.), *Quantum Physics Meets the Philosophy of Mind*, New York: de Gruyter 2014, pp. 65-82



other interpretations can also be taken to underwrite such metaphysical implications. Barrett's main point is that even if we assume that no collapse of any wave function takes place, the fact of definite and consistent measurement results is peculiar enough to warrant a strong variety of mind-body dualism, given four basic and widely accepted explanatory constraints – no collapse; no branching; state completeness and empirical consistency:

Consider a system that is initially in a superposition of states where the observable being measured has incompatible determinate values. Given that there is no collapse of the quantum-mechanical state, the linear dynamics predicts that, after a measurement interaction, a good observer will end up in an entangled superposition of recording mutually incompatible measurement results. Since there is no branching to yield different post-measurement observers with different determinate measurement records, there will be only one post-measurement observer with one measurement record. By state completeness, the post-measurement quantum-mechanical state is a complete representation of the physical state. And, by the linearity of the dynamics and empirical consistency, each of the records represented in the post-measurement state are in fact possible measurement results.

But, under these conditions, there is nothing in the complete post-measurement physical state that selects one of the possible records as actual. So insofar as the observer has a determinate measurement record, its value must be determined by something that is not determined by the complete physical state. The commitment to such a parameter is a commitment to a strong variety of mind-body dualism insofar as it is the value of this non-physical parameter, not the complete physical state, that determines the observer's experience and beliefs concerning the outcome of the measurement.<sup>317</sup>

This is a line of reasoning which seems to have been overlooked in the literature on the metaphysical implications of quantum mechanics.

Basically, Barrett's argument is that the definite measurement results can only be explained with regard to something outside of the physical system described by the wave function, *even if* no actual collapse of the wave function is presumed. This presumes a few explanatory constraints. First, the no-branching clause is assumed, since otherwise, it could just be maintained that the observer simply branches into a multitude of mutually exclusive subjects, each with a different measurement result. State-completeness is simply the rejection of any hidden, determining variables, and the constraint of empirical consistency is the assumption that the counterfactual measurement results were actually possible, i.e. that superdeterminism does not hold.

---

<sup>317</sup> Barrett 2006

On these assumptions, there's nothing that can possibly explain the definite measurement results, other than the interaction of something non-physical that is not described by the wave function, and the physical system, which is.

Moreover, Barrett argues that quantum theory indirectly will lead us towards metaphysical pluralism even in the absence of certain of these explanatory constraints:

If we accordingly remove the state completeness-constraint and accept the possibility of hidden physical variables as e.g. Bohmian mechanics do, we still need to make a distinction between the always determinate hidden variable of the pilot wave or equivalent, and of the non-determinate physical system described by the Schrödinger equation. And even if this is possible without some form of metaphysical pluralism, the consistent and definite measurement results (the fact that we for some reason never perceive a system in the superposition) as previously stated, do indeed still beg an explanation. It may well be argued that determinate particle positions produce definite measurement results, but why should we expect determinate mental records to consistently supervene on determinate particle positions? According to Barrett, the assumption of metaphysical pluralism in the manner of Wigner or von Neumann effectively solves the problem, while reductive physicalism arguably forces us to resort to ad-hoc explanations that at the worst amount to the assumption of a conspiratorial cosmic illusion, i.e. that the definite mental records or experiences and the definite particle positions just happen to coincide by chance.<sup>318</sup> Also, see Bell's perspective on superdeterminism above. Barrett further claims that the removal of the no-branching constraint, and the assumption of the possibility of the many-worlds and/or the many-minds interpretation implies a metaphysical pluralism that is at least as problematic as classical dualism. It's not clear exactly how he purports that this would be argued, but it seems obvious that a definite branching of the observer into different subjects, where each is limited in terms of what they definitely and exclusively experience, assumes some form of metaphysical distinction between them.

In summary, Barrett's argument seems convincing. All of the explanatory constraints he suggests are reasonable, and widely accepted among the interpretative models of quantum theory. What's more important – the definitive repudiation of those of Barrett's explanatory constraints, a repudiation which could have provided an explanation of determinate measurement records, has no positive empirical support.

---

<sup>318</sup> Ibid.

Moreover, the rejection of the explanatory constraints aside from the state completeness-constraint can in principle *not* be supported empirically. Rejecting non-collapse is obviously not an issue, since if collapse *is* assumed, Barrett's (as well as von Neumann's) argument can effectively be brought to bear on this particular assumption. Superdeterminism is, aside from being extremely implausible, potentially metaphysically impossible and theoretically non-parsimonious, impossible to prove with empirical data. Branching-theories like Everett-models where the world and/or the subject splits into several, mutually inaccessible ones, are likewise impossible to support empirically (unless the split is somehow negated, and direct access to several such subjects is enabled, which however arguably would undermine the basic proposition of the branching-theories). And as we have seen, Barrett's argument works even on the assumption of hidden variables. This is, again, since the consistent and definite measurement results still require an explanation which implies some form of metaphysical pluralism, wherein the observer must receive a privileged position if we are to explain why determinate mental records consistently supervene on determinate particle positions.

In contrast to the insufficient empirical support for the repudiation of Barrett's explanatory constraints, his proposed solution along the lines of Wigner or von Neumann is a different issue. Mind or phenomenal consciousness is immediately present to the subject, and must thus be said to function as an empirically viable basis for any explanation invoking it as such. Moreover, as we have seen in chapter 6, the proposed elimination of the subjective and the intentional is incoherent.

For this reason, an indirect argument for transsubstantial interaction can be constructed with regard to quantum theory, assuming Barrett's argument and the criticism of reductive physicalism of chapter 6, on the basis that the only empirically viable interpretations of quantum theory imply some form of such interaction.

## 8.4. Conclusions

Since several of the interpretations of quantum theory (if not most of them if we accept Barrett's line of argument) are clearly indicative of, or can be taken to imply a non-reductive immaterialist ontology, the criterion of scientific compatibility with regard to the metaphysical basis of post-mortem survival is taken to be fulfilled. Whether or not quantum mechanics *strongly* indicates a non-reductive immaterialist ontology is beyond the scope of this dissertation, but the mere possibility that it does, together with the popularity of several of the interpretative models that indicate this, will suffice to satis-

fy our moderate criterion of scientific compatibility as described in chapter 2.

Regarding the problem of interaction on classical substance dualism, in light of the accounts of causation discussed in chapter 7, and the possible meta-physical implications of certain important interpretations of quantum theory as well as of the Bell theorem, the problem doesn't seem to be as decisive as it's often portrayed.

The basic approaches of direct interaction or interaction by proxy are in and of themselves incapable of providing an entirely convincing solution without recourse to an amicable analysis of causation, although the proxy approach might perhaps show some promise as a response in combination with the quantum-mechanical model, if one could plausibly construct an argument to the effect that the notion of the collapse of the wave function causing the emergence of definite physical reality is somehow less problematic than direct interaction between a non-physical mind and physical reality. The category of causality-modifying direct interaction is also a conceivable option for the proponent of dualist interactionism, which could be supplemented by the quantum-mechanical model in relation to the probabilistic theory of causation, e.g. on the assumption that the observer's immaterial consciousness is capable of modifying some aspect of the probability wave in accordance with which the definite attributes of physical reality then emerges. Still, it seems that such a position cannot avoid the ontological discrepancy between the wave function and definite matter.

On the whole, the most favorable options from the perspective of interactionist dualism and other positions burdened by the problem of interaction, seems to be the assumption of the primitivist or hylemorphic responses of chapter 7.

The primitivist response has a particular affinity with much of what has been written in this chapter. It entails emphasizing that the major contemporary theories of causation do not, and possibly cannot, provide any exhaustive analysis of the causal relationship as such, which would imply that transsubstantial causation cannot be excluded on principle. The primitivist position may be further supported by the orthodox Copenhagen interpretation of quantum mechanics, in that it explicitly demands something outside the physical box to initiate the collapse of the wave function, which would provide foundations for an argument stating that the empirical data demands an equivalent of transsubstantial interaction, even if we have no idea what it entails, or how we exactly would define such a transsubstantial causal link. Finally, assuming collapse of the wave function, the von Neumann/Stapp model can be utilized as an ad hoc description of quantum-mechanical causal

interaction which also refutes the energy conservation objection. This is due to the fact that it clearly shows that no energy needs to be transferred in modulating the frequency of the probabilistic wave function, which as long as we accept that the wave function somehow determines the attributes of definite matter (a basic premise of quantum theory) implies that the interaction of the conscious observer and the wave function doesn't violate the conservation laws. The von Neumann/Stapp model is also clearly compatible with the hylemorphic response, as the metaphysics underlying the latter emphasize the actualization of potentialities in the description of causal relationships, which mirror the role of the observer in the von Neumann/Stapp model.

Also, the basic empirical support amassed for Bell's theorem either provides an additional line of argument against certain important presuppositions of the interaction problem, or indirect support for a non-reductive immaterialist ontology, with relative independence in relation to the particular interpretations of quantum theory. On the abandonment of the locality-aspect of local realism, which follows from the experimental violation of the Bell inequalities, the local-mechanistic understanding of causation which is presupposed in the problem of interaction is severely impaired, and thereby cannot *prima facie* be assumed as an ultimate criterion for either intrasubstantial or transsubstantial interaction. Furthermore, assuming non-realism in the relevant sense, and a rejection of the possibility of hidden variables, the preferred basis problem makes metaphysical pluralism an attractive option, i.e. as an obvious response to the question of why we actually experience definite measurement results.

Finally, "the unorthodox path" provided by Barrett, especially in combination with our critique of reductive physicalism in chapter 6, does indeed seem to enable a powerful line of argument from the point of view of the proponent of interactionist dualism. Not only does the no-collapse theories on Barrett's argument necessitate something akin to transsubstantial interaction in terms of explaining a central empirical fact, namely the definite measurement results and the correlation between determinate mental records and the determinate particle positions (assuming neorealism and hidden variables) – it also implies that the only empirically viable explanatory constraints upon interpretations of quantum mechanics either leads us to interactionist dualism, metaphysical idealism, or at least some form of non-reductive physicalism of an interactionist character that allows for non-material properties if not substances (cf. ch. 4 with regard to the phenomenal criterion and non-reductive physicalism) if we are to explain the persistence of definite measurement results.

## 9. Conclusions

You know my history. You know why my withers are quite unwrung by the fear that I was bribed – that I was lured into Christianity by the hope of everlasting life. I believed in God before I believed in Heaven. And even now, even if – let's make an impossible supposition – His voice, unmistakably His, said to me, 'They have misled you. I can do nothing of that sort for you. My long struggle with the blind forces is nearly over. I die, children. The story is ending,' would that be a moment for changing sides? Would not you and I take the Viking way: 'The Giants and Trolls win. Let us die on the right side, with Father Odin.'<sup>319</sup>

C. S. Lewis

### 9.1. On natural immortality

As was Kierkegaard's concern when arguing against the "natural immortality" proposed by his Hegelian contemporaries, the notion that one by supporting the metaphysical possibility of post-mortem survival at the same time risks diminishing the miraculous character of such an occurrence, and thereby may undermine the necessity of trust in the undeserved grace of God – which otherwise would be the only conceivable recourse to one hoping for eternal life – has been a recurrent fear of mine throughout this project. With frygt og bæven ought we work out our salvation, as Paul maintained, yet if post-mortem survival is in our nature by default, Kierkegaard's position was that our need for God's salvific grace at least would seem less acute.

In terms of a remedy for such concerns, it may be remarked that post-mortem survival clearly does not necessarily imply salvation in terms of the Judaeo-Christian theological scheme, and even less so does the mere metaphysical possibility of survival. The notion that there are forms of post-mortem survival which are indeed not very preferable at all has been a part of most theistic eschatologies that have contained ideas of an afterlife. Moreover, even on the assumption of soteriological universalism, the Classical theist who hold that by necessity, God's activity continuously preserves

---

<sup>319</sup> C. S. Lewis, *Letters to Malcolm: Chiefly on Prayer*, New York: Harcourt, Brace & World Inc. 1964, p. 120

every contingent entity in existence, must ascribe this fact to the undeserved grace of God in that we are thus created and sustained.

But we are perhaps getting ahead of ourselves here – before addressing indirect consequences of a successful case such as the one this dissertation presents, the issue of whether or not it can actually be deemed successful, and in what sense, must obviously be attended to.

## 9.2. Evaluating the case

All in all, the case that has been presented is rather complex. We first encounter the outline of a piece of circumstantial evidence in chapter 3, the historical presence of the intuition of continuity, evidence which can be taken in two directions. The more radical interpretation would be that the mere fact that human beings have an inherent tendency to form beliefs in post-mortem survival, also provides a slight epistemic support for the position that such survival actually is possible, provided the assumption that the beliefs human beings strongly tend to form are also true more often than not.

The less radical interpretation of this piece of evidence merely entails that the issue of the veracity of these beliefs as such, is one worthy of our fullest attention and critical exploration, since their emergence seems all but inevitable in human cultures.

In the following chapter, number 4, it is argued that the significant conceptual difficulties with the notion of post-mortem survival could be circumvented by utilizing what we called “the phenomenological criterion” for post-mortem survival, such that if the phenomenal, subjective self persists or reemerges beyond physical death, the person in question is taken to have survived.

The viability of this criterion, however, is dependent upon the truth of certain metaphysical preconditions, as is elaborated in chapters 5, where a model of the human self in accordance with the criterion is presented. More specifically, the viability of the criterion presupposes either a theistic background metaphysics; i.e. a non-naturalist global ontology incompatible with reductive physicalism – or at the very least a non-reductive philosophical anthropology, which has been variously denoted with “phenomenalism” or “non-materialism”, that is equally incompatible with universal reductive physicalism. Thus, if the criterion is to be even possible to fulfill, the obstacle of reductive physicalism must be somehow averted.

However, already at this stage, a first support for the rational permissibility of afterlife beliefs discussed in chapter 2 is arguably established. If the phenomenological criterion is coherent, and can be utilized to effectively address the conceptual problems associated with post-mortem survival, then afterlife beliefs must be regarded as rationally permissible to hold insofar as the supporting background metaphysics which the criterion must presuppose (i.e. the non-reductive/phenomenalist metaphysics), can within reason be said to possibly hold. Thus, since the criterion is coherent, and would work if the background metaphysics were in place, beliefs in the possibility of post-mortem survival are rationally permissible to the extent that this set of metaphysical positions can be rationally held. Yet, a much stronger support for the metaphysical possibility of post-mortem survival will be provided if reductive physicalism as such can be effectively criticized, and if the arguably debilitating problem of interaction which affects most of the non-reductive ontologies can be addressed. Moreover, we are thus far yet to see the criterion of scientific compatibility satisfied.

The criticism presented in chapter 6 provides a response to reductive physicalism, and gives us several decisive reasons to reject the position as incoherent and empirically inadequate.

Chapter 7, which regards the problem of interaction, provides grounds for considering the problem of interaction as less than decisive from the point of view of the non-reductionist or phenomenalist ontologies. This is for two reasons. 1. The contemporary views on the metaphysics of causation do not necessarily imply an equivalent of the problem of transsubstantial interaction, and several in fact seem perfectly compatible with such a kind of phenomena. 2. There are several workable responses to the problem of interaction, which are especially viable in concert with an appropriate position on causal metaphysics.

Chapter 8 anchors the scientific compatibility of the case. This compatibility must be considered granted, only in that several of the most influential interpretations of quantum theory clearly allow for a causal role of the non-physical observer (i.e. an observer outside of the wave function described by the Schrödinger equation), as well as the fact that the Bell theorem and the impressive empirical support thereof allows us to weaken the impetus of the interaction problem insofar as the Bell theorem implies the possibility of non-local causal interactions on the rejection of superdeterminism.

What is more, however, the arguments provided towards the conclusion that all interpretations of quantum mechanics which involve certain explanatory constraints (the negation of which cannot be supported empirically even in principle) actually imply an equivalent to transsubstantial interaction if they



are to explain certain empirical facts central to quantum theory (i.e. the presence of definite measurement results), entails a strong scientific compatibility of the non-reductionist or phenomenalist metaphysical positions.

In other words, if the best available set of theories regarding physical reality, which quantum mechanics is generally taken to provide, and the only empirically viable interpretations thereof actually tend to imply a phenomenalist metaphysical pluralism if they are to explain a fundamental empirical fact, the scientific compatibility of the background ontologies that provide an alternative to monist reductive physicalism must be considered strong.

All in all, the case presented with these chapters recounted, clearly establishes the rational permissibility of belief in the possibility of an afterlife. Whether or not it entails that it's rationally obligatory to affirm such a possibility pivots on the strength of the criticism of reductive physicalism in chapter 6, since reductive physicalism blocks the possibility of post-mortem survival as we've considered it, as well as the additional support which can be provided for the non-reductive/phenomenalist ontologies in addressing the problem of interaction that plagues most of them, and in anchoring their scientific compatibility.

However, the weight of the arguments provided, and the successful anchoring of said scientific compatibility, together with the criticism of the interaction problem that is presented, makes it very difficult to deny the actual metaphysical possibility of post-mortem survival, and the rationality of affirming this possibility within today's scientifically inclined or even scientific intellectual milieu. The conclusion is therefore that with regard to the evidence reviewed, and thus within the context of contemporary science and philosophy, the metaphysical possibility of post-mortem survival cannot rationally be denied.

### 9.3. Concluding remarks on theism

Finally, something must be said of the importance of theism for this dissertation's topic, an issue which has so far not been extensively touched upon, yet which has in some sense hovered in the background. Theism just in and of itself renders post-mortem survival metaphysically possible, which can be seen throughout chapter 4. What's more, theism in the classical sense is likely the most important conceptual factor in terms of actualizing post-mortem survival.

Mere metaphysical possibility namely does not entail that something will actually occur. It may be metaphysically possible for an acorn to grow into an oak, but not without the requisite context and nutrients. Similarly, it may be metaphysically possible for a human person to survive physical death, yet bringing about actual survival in terms of the preservation or re-emergence of the phenomenal self, may well require divine intervention. In other words, the situation may well be that a human person is such a being that can possibly survive corporeal death, but that the actualization of such survival is yet dependent upon certain external factors (factors which are necessarily present with theism).

For these reasons, it must be remarked that classical theism as such provides a stronger epistemic relevance in relation to afterlife beliefs than any other single factor.

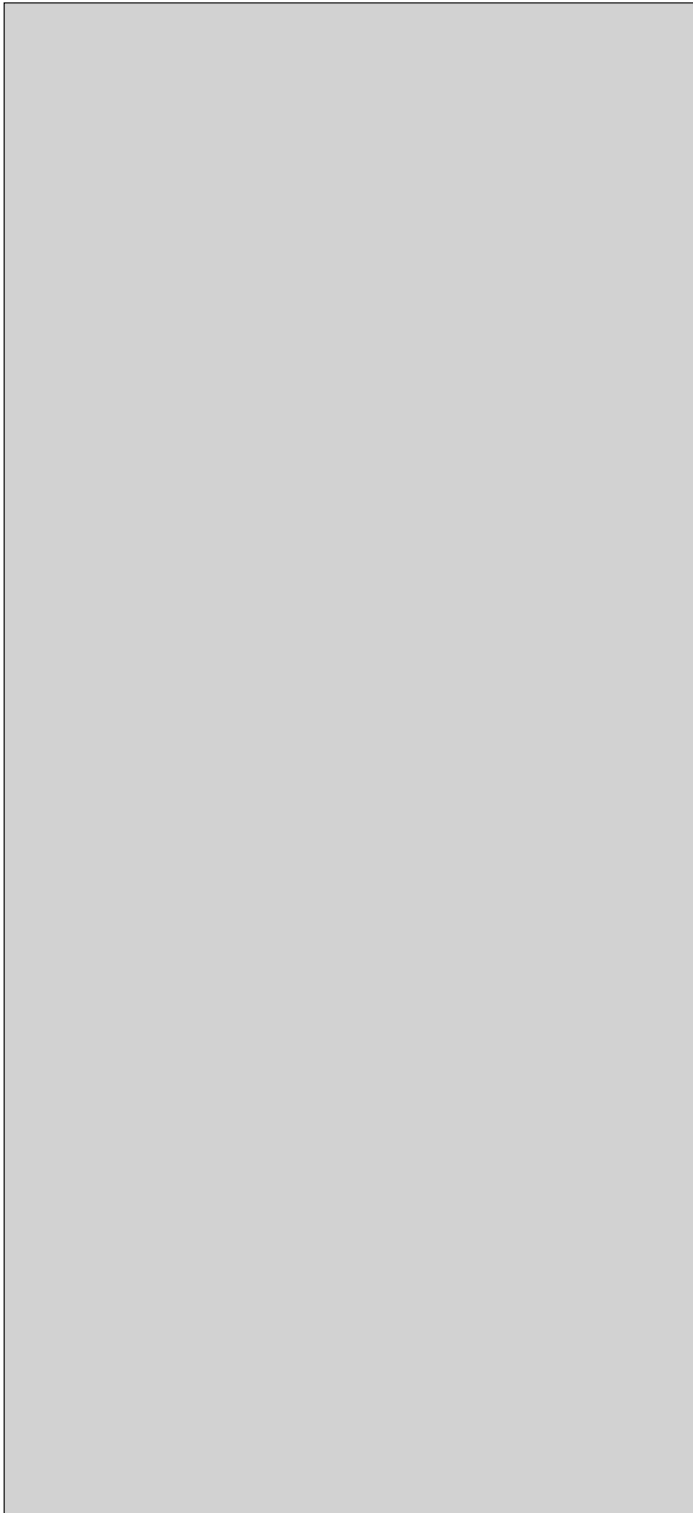
Moreover, theism is the only obvious aspect of a worldview that can render a human person's actual post-mortem survival *probable*, since even if we were to assume that survival can occur without some form of divine intervention, this does not imply that it's at all likely to occur – it may just as well be astronomically improbable. The presence of a divine agent, on the other hand, with the loving intention of preserving human persons beyond corporeal death, can clearly render actual survival probable.

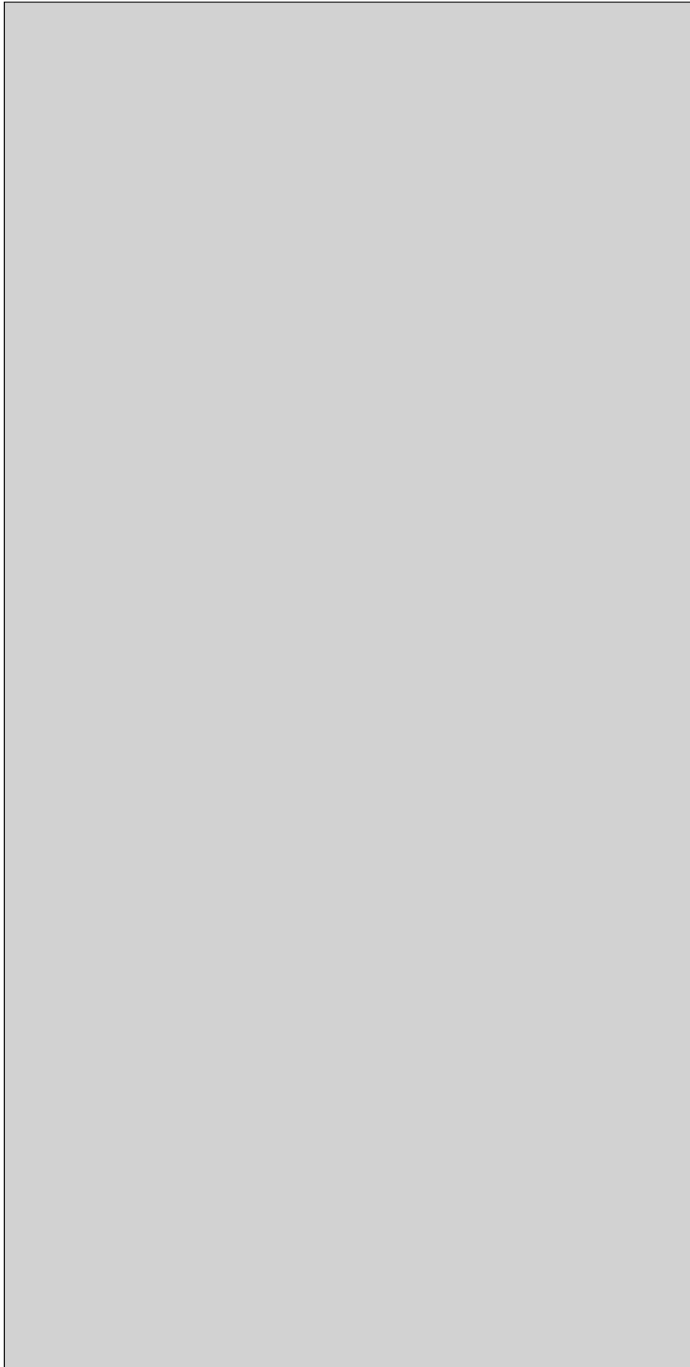
In terms of axiological issues, theism also has a clear importance. The value of the afterlife, if any, must be anchored in a proper eschatological setting, since mere survival is not necessarily desirable. Classical theism provides such an eschatological setting in accordance with which post-mortem survival is both valuable and meaningful. Lacking a constructive eschatological setting, however, afterlife beliefs in and of themselves can potentially become an existentially problematic factor in a person's worldview, engendering negative consequences for the individual and her community. Correspondingly, in the words of St. Ambrose,

God did not establish death in the beginning, but gave it as a remedy. ... after the transgression of Adam and man's condemnation to long labors and unbearable sorrow, his life became wretched. Consequently, an end had to be established for evils, so that death might restore what life had lost. For, unless grace should breathe upon it, immortality would be rather a burden than an advantage.<sup>320</sup>

---

<sup>320</sup> St. Aurelius Ambrosius, "On the Death of His Brother Satyrus", in Leo P. McCauley (trans.), *Funeral Orations by Saint Gregory Nazianzen and Saint Ambrose*, Washington: Catholic University of America Press 1953, p. 216





---

<sup>321</sup> E. E. Cummings, G. J. Firmage (ed.), "if i believe", in *Complete Poems, 1904 – 1962*, Liveright Publishing Corp. 1992, pp. 47-49

## 10. Bibliography

- Abner, Shimony, "Bell's Theorem", *The Stanford Encyclopedia of Philosophy*, (Winter 2013 Edition), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2013/entries/bell-theorem> 2014-12-08
- Achilli, Justin et al, *Vampire – The Masquerade*, 20th anniversary ed., White Wolf Publishing 2011
- Albert, D. Z., Loewer, B., "Interpreting the Many Worlds Interpretation", *Synthese* 77, 1988
- Alexander, Eben, *Proof of Heaven*, New York: Simon & Schuster 2012
- Allen, James P., *The Ancient Egyptian Pyramid Texts*, Atlanta: Society of Biblical Literature 2005
- Althusser, Louis, *Lenin and Philosophy and Other Essays*, London: Monthly Review 1971
- Ambrosius, St. Aurelius, "On the Death of His Brother Satyrus", in Leo P. McCauley (trans.), *Funeral Orations by Saint Gregory Nazianzen and Saint Ambrose*, Washington: Catholic University of America Press 1953
- Andrea, Alfred J. (ed.) *World History Encyclopedia*, Santa Barbara: ABC-CLIO 2011
- Ansmann, Markus, Wang, H. et al, "Violation of Bell's inequality in Josephson phase qubits", *Nature* 461, 2009
- Aquinas, St. Thomas, Fathers of the English Dominican Province (trans.), *Summa Theologiae* 1920. Retrieved from <http://www.newadvent.org/summa/index.html> 2017-10-14
- Aquinas, St. Thomas, *On the Truth of the Catholic Faith: Summa Contra Gentiles*, New York: Image Books 1956
- Archer, John, *The Nature of Grief – The Evolution and Psychology of Reactions to Loss*, London: Routledge 1999
- Armstrong, D.M., *A World of States of Affairs*, Cambridge: Cambridge University Press 1997
- Aristotle, W. D. Ross (trans.), *Metaphysics*, Whitefish: Kessinger Publishing 2010
- Ayer, A. J., *Language, Truth and Logic*, London: Penguin Books 2001
- Baker, Lynne R., "Death and the Afterlife", in Wainwright, William J. (ed.), *The Oxford Handbook of Philosophy of Religion*, New York: Oxford University Press 1997
- Baker, Lynne R., *Naturalism and the First-Person Perspective*, Oxford: Oxford University Press 2013
- Bar-Yosef Mayer, Daniella E., Vandermeersch, Bernard, Bar-Yosef, Ofer, "Shells and ochre in Middle Paleolithic Qafzeh Cave, Israel: indications for modern behaviour" *Journal of Human Evolution*, Volume 56, Issue 3, March 2009
- Bar-Yosef, Ofer, "The chronology of the Middle Paleolithic of the Levant", in Aka-zawa, T., Aoki, K., Bar-Yosef, Ofer (eds.), *Neandertals and Modern Humans in Western Asia*, New York: Springer-Verlag 1998

- Bardale, Rajesh, *Principles of Forensic Medicine and Toxicology*, Jaypee Brothers Medical Publishers Ltd., New Delhi 2011
- Barrett, Jeffrey A., "A Quantum-Mechanical Argument for Mind-Body Dualism" *Erkenntnis*, 65 2006
- Barrett, Jeffrey A., "Quantum Mechanics and Dualism", in A. Corradini and U. Meixner (eds.), *Quantum Physics Meets the Philosophy of Mind*, New York: de Gruyter 2014
- Barrett, Justin, *Why Would Anyone Believe in God?*, Lanham: AltaMira Press 2004
- Bates, Daisy, *Aboriginal Perth: Bibbulmun Biographies and Legends*, Carlisle: Hesperian Press 1992
- Bell, John Stuart, "On the Einstein Podolsky Rosen Paradox", *Physics* vol. 1, no 3, 1964
- Bell, John Stuart, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, New York 2004
- Benson, Elizabeth, *Death and the Afterlife in Pre-Columbian America*, Washington: Dumbarton Oaks Research Library and Collection 1975
- Berger, Peter, Luckmann, Thomas, *Kunskapssociologi*, Stockholm: Wahlström & Widstrand 2007
- Berkeley, George, *An Essay Towards a New Theory of Vision*, Dublin: George Berkeley 1709
- Berkeley, George, *Principles of Human Knowledge and Three Dialogues*, Oxford University Press, New York: 1996
- Bleek, W.H.I., Lloyd, Lucy, Halstead, John (ed.), *Specimens of Bushman Folk-Lore*, Sandhurst: Abela Publishing 2009
- Blinder, S. M., *Introduction to Quantum Mechanics: In Chemistry, Materials Science and Biology*, Burlington: Elsevier Academic Press 2004
- Bohr, Niels, "Discussion with Einstein on Epistemological Problems in Atomic Physics", in Schilpp, P.A (ed.), *Albert Einstein: Philosopher-Scientist*, New York: Fine Communications 2001
- Borchert, Donald M., (ed.) *Encyclopedia of Philosophy*, vol. 6, Thomson Gale 2006
- Bowley, R., Sánchez, M., *Introductory Statistical Mechanics*, Clarendon Press: Oxford, 1999
- Boyd, Robert, Richerson, Peter, *Culture and the Evolutionary Process*, Chicago: University of Chicago Press 1985
- Bradley, Ben, Feldman, Fred, Johansson, Jens (eds.), *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012
- Brower, Jeffrey E., *Aquinas's Ontology of the Material World*, Oxford: Oxford University Press 2014
- Buechner, Jeff, "Not Even Computing Machines Can Follow Rules", in Berger, Alan (ed.), *Saul Kripke*, New York: Cambridge University Press 2011
- Bunnin, Nicholas, Yu, Jiyuan (eds.), *The Blackwell Dictionary of Western Philosophy*, Oxford: Wiley-Blackwell 2004
- Cabello, Adán, "Interpretations of quantum theory: A map of madness" <http://arxiv.org/abs/1509.04711v1>, 2015, retrieved 2016-07-20
- Campbell, Joseph, *The Masks of God: Primitive mythology*, New York: Viking Press 1959
- Cartwright, Nancy, "Causation: One Word, Many Things", *Philosophy of Science*, Vol. 71, No. 5 2004
- Casati, R., "Hilary Putnam", in Vattimo, Gianni (ed.) *Enciclopedia Garzanti della Filosofia*, Milan: Garzanti Editori 2004
- Chalmers, David, "Facing up to the Problem of Consciousness", in *Journal of Consciousness Studies* 2: 200-19, 1995

- Chalmers, David, *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press, 1996
- Chesterton, G. K., *Orthodoxy*, Stornoway: Reformation Press 2002
- Chesterton, G. K., *What's Wrong With the World*, London: Cassell and Company 1910
- Clauser, John F., "Early History of Bell's Theorem", in Bigelow, N. P., Eberly, J. H. et al (eds.), *Coherence and Quantum Optics VIII: Proceedings of the Eight Rochester Conference on Coherence and Quantum Optics*, New York: Kluwer Academic/Plenum Publishers 2003
- Clottes, Jean, *Chauvet Cave: The Art of Earliest Times*, Salt Lake City: University of Utah Press 2003
- Cocchiarella, Nino B., "Property Theory", in Craig, Edward (ed.) *Routledge Encyclopedia of Philosophy*, New York: Routledge 1998
- Collins English Dictionary – Complete & Unabridged 10th Edition*, "substance". Retrieved March 11 2014 from <http://dictionary.reference.com/browse/substance>
- Condemni, Silvana, Weniger, Gerd-Christian (eds.) *Continuity and Discontinuity in the Peopling of Europe*, New York: Springer-Verlag 2011
- Conee, Earl, Feldman, Richard, *Evidentialism: Essays in Epistemology*, Oxford: Clarendon Press, 2004
- Cotterell, Arthur Flagg, *A Dictionary of World Mythology*, Oxford: Oxford University Press 1986
- Cramer, John, "The Transactional Interpretation of Quantum Mechanics", *Reviews of Modern Physics* 58, 1986
- Craig, Edward (ed.), *Concise Routledge Encyclopedia of Philosophy*, New York: Taylor & Francis Routledge 2000
- Crumley, Jack, *A Brief Introduction to the Philosophy of Mind*, Oxford: Rowman & Littlefield Publishers Inc. 2006
- Cummings, Edward Estlin, "Seven Poems" in *The Dial*, Vol. 68, No. 1, January 1920
- Cummings, Edward Estlin, G. J. Firmage (ed.), *Complete Poems, 1904 – 1962*, New York: Liveright Publishing Corp. 1992
- Davies, Paul, Brown, Julian (eds.), *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics*, Cambridge University Press: Cambridge 1986
- Davidson, Donald, *Essays on Actions and Events*, Oxford: Clarendon Press 1980
- Demallie, Raymond J., Parks, Douglas R. (eds.) *Sioux Indian Religion*, Norman: University of Oklahoma Press 1987
- Deveaud, Benoît, Quattropiani, Antonio, Schwendimann, Paolo, *Quantum Coherence in Solid State Systems*, Società Italiana di Fisica: Bologna 2009
- Descartes, René, *Les passions de l'âme*, Amsterdam: Lodewijk 1649
- Descartes, René, *Meditations on First Philosophy*, Oxford: Oxford University Press 2013
- Dixon, R. M. W., Koch, Grace, *Dyirbal song poetry, the oral literature of an Australian rainforest people*, Brisbane: University of Queensland Press 1996
- Doolittle, Justus, *Social Life of the Chinese*, New York: Harper & Brothers 1876
- Dunand, Françoise, Lichtenberg, Roger, *Mummies and Death in Egypt*, Ithaca: Cornell University Press 2006
- Duncan, A. J., "Two-Photon Decay of Atomic Hydrogen", in Narumi, H., Shimamura, I. (eds.), *Atomic Physics 10*, New York: Elsevier Science Publishing Company 1987
- Durkheim, Emile, *The Rules of Sociological Method*, The Free Press: New York 1982

- Dürr, D., Goldstein, S., Zanghi, N., "Quantum Equilibrium and the Role of Operators as Observables in Quantum Theory", *Journal of Statistical Physics* 134, 2009
- Editors of Encyclopædia Britannica, The, "the Dreaming" *Encyclopædia Britannica, Encyclopædia Britannica Online Academic Edition*, Encyclopædia Britannica Inc. 2013, retrieved from <http://www.britannica.com/EBchecked/topic/171252/the-Dreaming> 2013-04-15
- Eibenberger, Sandra et al, "Matter-wave interference with particles selected from a molecular library with masses exceeding 10000 amu", *Physical Chemistry Chemical Physics* 15
- Engels, Friedrich, Marx, Karl, *Das Kapital* vol. I, Hamburg: Otto Meisner Verlag 1867
- Evans, C, *Pocket Dictionary of Apologetics & Philosophy of Religion*, Downers Grove: Intervarsity Press: 2002
- Fair, David, "Causation and the Flow of Energy", *Erkenntnis*, 14:215-219, 1979
- Falkenburg, Brigitte, *Particle Metaphysics: A Critical Account of Subatomic Reality*, Leipzig: Springer-Verlag: 2003
- Feser, Edward, *Philosophy of Mind*, Oxford: Oneworld 2006
- Feser, Edward, *The Last Superstition*, South Bend: St. Augustine Press 2008
- Feser, Edward, "Hayek, Popper, and the Causal Theory of the Mind", in Marsh, Leslie (ed.), *Hayek in Mind: Hayek's Philosophical Psychology (Advances in Austrian Economics, Volume 15)* Boston: Emerald Group Publishing Limited 2011
- Fischer, John Martin, "Results of funded research projects", The SPT Immortality Project, July 2012, retrieved from <http://www.sptimmortalityproject.com/wp-content/uploads/2012/07/RESULTS-OF-FUNDED-RESEARCH-PROJECTS.pdf> 2016-11-11
- Flew, Anthony G. N., *Merely Mortal: Can You Survive Your Own Death?*, Amherst: Prometheus Books 2000
- Flew, Antony G. N., *God and Philosophy*, London: Hutchinson 1966
- Fodor, Jerry, *Psychological Explanation*, New York: Random House 1968
- Forrest, Peter, "The Identity of Indiscernibles", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/fall2008/entries/identity-indiscernible/> 2016-11-14
- Foster, John, *The Immaterial Self*, New York: Routledge 1991
- Freedman, S. J., Clauser, John F., "Experimental test of local hidden-variable theories", *Physical Review Letters* 28, 938, 1972
- Frobenius, Leo, *Kulturgeschichte Afrikas: Prolegomena zu einer historischen ge-staltlehre*, Paderborn: Salzwasser-Verlag 2012
- Gamut, L. T. F. (Johan van Benthem, Jeroen Groenendijk et al), *Logic, Language and Meaning*, Chicago: University of Chicago Press 1991
- Gerry, Christopher, Bruno, Kimberly, *The Quantum Divide*, Oxford: Oxford University Press, 2013
- Gilmore, Cody, "When Do Things Die?" in Bradley, Ben, Feldman, Fred, Johanson, Jens (eds.), *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012
- Giustina, Marissa, Mech, Alexandra et al, "Bell violation using entangled photons without the fair-sampling assumption", *Nature*, 497 2013
- Glare, P.G.W., Edwards, W. M. et al (eds.), *Oxford Latin Dictionary*, Oxford: Oxford University Press 1968
- Good, I. J. (ed.), *The Scientist Speculates*, London: Heinemann 1961



- Gordon, Andrew H., Schwabe, Calvin W., *The Quick and the Dead: Biomedical Theory in Ancient Egypt*, Leiden: Brill/Styx 2004
- Gordon, Hayim, Tamari, Shlomit, *Maurice Merleau-Ponty's Phenomenology of Perception*, Westport: Praeger 2004
- Gouesbet, Gerard, *Hidden Worlds in Quantum Physics*, Dover Publications, New York, 2001
- Göcke, Benedikt Paul, *After Physicalism*, Notre Dame: The University of Notre Dame Press 2012
- Gribbin, John, Gribbin, Mary, Gribbin, Jonathan, *Q is for Quantum: An Encyclopedia of Particle Physics*, Simon & Schuster: New York 2000
- Grigolini, Paolo, *Quantum Mechanical Irreversibility and Measurement*, New Jersey: World Scientific Publishing Co. 1993
- Grof, Stanislav, *Beyond the Brain: Birth, Death and Transcendence in Psychotherapy*, New York: State University of New York Press 1985
- Hackett, Stuart, *The Resurrection of Theism*, Austin: Wipf and Stock 1957
- Hall, Ned, "Two Concepts of Causation", in Collins, J., Hall, N. (eds.), *Causation and Counterfactuals*, Cambridge: MIT Press 2004
- Hammond, Paula, *The Atlas of the World's Most Dangerous Animals*, Tarrytown: Marshall Cavendish 2010
- Hasker, William, *Metaphysics*, Downers Grove: InterVarsity Press 1983
- Hasker, William, *The Emergent Self*, Ithaca: Cornell University Press 1999
- Hasker, William, "Materialism and the Resurrection: Are the Prospects Improving?", *European Journal for the Philosophy of Religion*, 3:1, 2011
- Heidegger, Martin, *Wegmarken*, Frankfurt am Main: Vittorio Klostermann 2004
- Heidel, Alexander, *The Gilgamesh Epic and Old Testament Parallels*, Chicago: University of Chicago Press 1949
- Hemmick, Douglas, Shakur, Asif, *Bell's Theorem and Quantum Realism*, Heidelberg: Springer 2002
- Herbert, Nick, *Quantum Reality*, New York: Anchor Books 1985
- Hick, John, *Death and Eternal Life*, Louisville: Westminster John Knox Press 1994
- Hick, John, *An Interpretation of Religion*, New York: Palgrave MacMillan 2004
- Hiley, B.J., "Some Remarks on the Evolution of Bohm's Proposals for an Alternative to Standard Quantum Mechanics", preprint, 2010, retrieved from: [www.bbkc.ac.uk/tpu/BasilHiley/History\\_of\\_Bohm\\_s\\_QT.pdf](http://www.bbkc.ac.uk/tpu/BasilHiley/History_of_Bohm_s_QT.pdf) 2014-11-24
- Hiscock, Peter, *The Archaeology of Ancient Australia*, New York: Routledge 2008
- Home, Dipankar, Whitaker, Andrew, *Einstein's Struggles with Quantum Theory: A Reappraisal*, Springer Science: New York 2010
- Hooft, Gerard 't, "The Free-Will Postulate in Quantum Mechanics", *arXiv*, 2007, retrieved from: [arxiv.org/abs/quant-ph/0701097](http://arxiv.org/abs/quant-ph/0701097), 2014-12-07
- Hovers, Erella, *The Lithic Assemblages of Qafzeh Cave*, Oxford: Oxford University Press 2009
- Hovers, Erella, Kuhn, Steven, *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, New York: Springer 2010
- Horkheimer, Max, Adorno, Theodor, *Dialectic of Enlightenment*, Stanford: Stanford University Press 2002
- Hume, David, *A Treatise of Human Nature*, Oxford: Oxford University Press 1960
- Hume, David, *An Enquiry Concerning Human Understanding*, Oxford: Oxford University Press 2007
- Husserl, Edmund, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, The Hague: Martinus Nijhoff Publishers 1983

- van Inwagen, Peter, "The Possibility of Resurrection", *International Journal for the Philosophy of Religion*, 9 1978
- van Inwagen, Peter, "The Possibility of Resurrection", in van Inwagen, Peter, *The Possibility of Resurrection and Other Essays in Christian Apologetics*, Boulder: Westview Press 1998
- Jackson, Frank, "Epiphenomenal Qualia", *Philosophical Quarterly*, 32 1982
- Jaworsky, William, *Philosophy of Mind*, Malden: Blackwell Publishing 2011
- Jurmain, Robert, Kilgore, Lynn, Trevathan, Wenda, *Essentials of Physical Anthropology*, Belmont: Wadsworth 2009
- Kagan, Shelley, *Death*, New Haven: Yale University Press 2010
- Kastner, Ruth, *The Transactional Interpretation of Quantum Mechanics*, New York: Cambridge University Press 2013
- Keightly, David, "Art, Ancestors and the Origins of Writing in China", *Representations*, No. 56, Special Issue: The New Erudition 1996
- Kim, Jaegwon, *Supervenience and Mind*, Cambridge: Cambridge University Press, 1995
- Klawonn, Erich, *Mind and Death: A Metaphysical Investigation*, Odense: University Press of Southern Denmark 2009
- Kriegel, Uriah, *Subjective Consciousness: A Self-Representational Theory*, Oxford: Oxford University Press, 2009
- Kripke, Saul, *Wittgenstein on Rules and Private Language*, Cambridge: Harvard University Press 1984
- Kuhn, Thomas, *Black-Body Theory and the Quantum Discontinuity, 1894-1912*, Oxford University Press, Oxford 1978
- Lazarou, Dimitris, "Interpretation of Quantum Theory – An overview", <http://arXiv:0712.3466v2>, 2009, retrieved 2016-07-20
- Lenin, Vladimir, *Материализм и эмпириокритицизм*, Moscow: Foreign Languages Publishing House 1959
- Leroi-Gourhan, Andre, Michelson, Annette, "The Religion of the Caves: Magic or Metaphysics?", *October* vol. 37, The MIT Press
- Lewin, Roger; Foley, Robert, *Principles of Human Evolution*, Malden: Blackwell Publishing 2004
- Lewis, C. S., *Letters to Malcolm: Chiefly on Prayer*, New York: Harcourt, Brace & World Inc. 1964
- Lewis, David, *On the Plurality of Worlds*, Oxford: Blackwell 1986
- Lieberman, Philip, *Uniquely Human*, Cambridge: Harvard University Press 1991
- Lowenstein, Tom; Asatchaq; Tukummiq, *The Things That Were Said of Them: Shaman Stories and Oral Histories of the Tikigaq People*, Oakland: University of California Press 1992
- Lycan, William, (ed.) *Mind and Cognition*, Oxford: Blackwell 1990
- Lycan, William, "Is property dualism better off than substance dualism?", *Philosophical Studies*, 164:533-542, 2013
- Mackie, J. L., *The Cement of the Universe*, Oxford: Oxford University Press 1980
- Mahlin, Shimon, "The Collapse of Quantum States: A New Interpretation", *Foundations of Physics*, vol. 23, issue 26, June 1993
- Mainzer, Klaus, *Thinking in Complexity: The Complex Dynamics of Matter, Mind and Mankind*, Heidelberg: Springer-Verlag 1994
- Malcolm, Robert, Dixon, Ward, *The Dyrbal Language of North Queensland*, Cambridge: Cambridge University Press 1972
- Marcuse, Herbert, *One-Dimensional Man: Studies in the ideology of advanced industrial society*, New York: Routledge 2002

- Marean, C. W., Bar-Matthews, M., Bernatchez, J. et al, "Early human use of Marine resources and pigment in South Africa during the Middle Pleistocene", *Nature* 2007
- Martinez, Daniel E., "Mortality patterns suggest lack of senescence in Hydra", *Experimental Gerontology* 33, 1998
- Melnyk, Andrew, *A Physicalist Manifesto*, New York: Cambridge 2003
- Menzies, Peter, "Probabilistic Causation and the Pre-emption Problem", *Mind*, 105 1996
- Menzies, Peter, "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy* (Spring 2014 ed.), Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual> 2014-09-20
- Merricks, Trenton, "The Resurrection of the Body", in Flint, Thomas P., Rea, Michael C. (eds.), *The Oxford Handbook of Philosophical Theology*, Oxford: Oxford University Press 2009
- Metzinger, Thomas, *Subjekt Und Selbstmodell: Die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*, Paderborn: Mentis 1993
- Metzinger, Thomas, *Being No One. The Self-Model Theory of Subjectivity*, Cambridge: MIT Press 2011
- Meyers, Rusty L., *The Basics of Physics*, Westport: Greenwood Press 2006
- Miller, Seumas, "Social Institutions", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Zalta, Edward N. (ed.), retrieved from <http://plato.stanford.edu/archives/win2012/entries/davidson/> 2013-04-09
- Mitchell, Basil, *The Justification of Religious Belief*, London: The Macmillan Press 1973
- Moore, George Edward, "The Refutation of Idealism", *Mind*, 12, 1903
- Moore, George Edward, "Proof of an external world" *Proceedings of the British Academy* 25, 1939
- Moreland, J. P., Habermas, Gary, *Beyond Death: Exploring the Evidence for Immortality*, Eugene: Wipf and Stock Publishers 2004
- Morgenbesser, S., Suppes, P., and White, M., *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, New York: St. Martin's Press 1969
- Muirhead, J. H. (ed.) *Contemporary British Philosophy*, London: Allen & Unwin 1925
- Mure, Geoffrey, *Retreat from Truth*, Oxford: Oxford Basil Blackwell 1958
- Murphey, Nancey, *Bodies and Souls, or Spirited Bodies?* Cambridge: Cambridge University Press 2006
- Nadeau, Randall (red.) *The Wiley-Blackwell Companion to Chinese Religions*, Malden: Blackwell Publishing Ltd. 2012
- Nagel, Thomas, "What is it like to be a bat", *The Philosophical Review*, vol. 83, no. 4, 1974
- von Neumann, John, *Mathematical Foundations of Quantum Mechanics*, Princeton: Princeton University Press, 1955
- Noonan, Harold, *Personal Identity*, New York: Routledge 2003
- Norsen, Travis, Nelson, Sarah, "Yet Another Snapshot of Foundational Attitudes Toward Quantum Mechanics", <http://arxiv.org/abs/1306.4646v2> 2013, retrieved 2016-07-20
- Oderberg, David S., *The Metaphysics of Identity Over Time*, London: Macmillan Press 1993
- Oderberg, David S., "Hylemorphic Dualism", in *Social Philosophy and Policy* 22, 2005

- OED Online, June 2017, Oxford University Press, retrieved from: <http://www.oed.com/view/Entry/29140?redirectedFrom=causation&> 2017-08-12
- Parker, Michael A., *Solid State and Quantum Theory for Optoelectronics*, Boca Raton: CRC Press 2010
- Pearle, Philip, "Quasirelativistic quasilocal finite wave-function collapse model", *Physical Review A*, 71, 2005
- Penrose, Roger, *The Emperor's New Mind*, Oxford University Press: Oxford 1999
- Penrose, Roger, Hameroff, Stuart, "Consciousness in the universe: A review of the Orch-OR-theory", *Physics of Life Reviews* 11, 2014
- Peoples, Hervey C., Duda, Pavel, Marlowe, Frank W., "Hunter-Gatherers and the Origins of Religion" *Human Nature*, vol. 27, 2016
- Petitt, Paul, *The Paleolithic Origins of Human Burial*, Taylor & Francis 2011
- Piccinini, Gualtiero, "The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism", *Philosophy and Phenomenological Research*, vol. 81, no. 2, September 2010
- Planck, Max, "Über das Gesetz der Energieverteilung im Normalspektrum", *Annalen der Physik*, vol. 309, issue 3, 1901
- Plantinga, Alvin, *Warranted Christian Belief*, Oxford: Oxford University Press, 2000
- Pooley, Oliver, "Substantialist and Relationalist Approaches to Spacetime", in Batterman, Robert (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press 2013
- Popper, Karl, Eccles, John, *The Self and Its Brain*, New York: Springer Verlag 1977
- Postrel, Steven, Feser, Edward, "Reality Principles: An Interview With John R. Searle", *Reason*, Feb. 2000
- Price, Huw, "Does Time-Symmetry Imply Retrocausality? How the Quantum World Says 'Maybe'", *Studies in History and Philosophy of Modern Physics* 43, 2012
- Pust, Joel, "Intuition", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/intuition/>, retrieved 2013-04-11
- Putnam, Hilary, "Psychophysical Predicates", in Capitan, W., Merrill, D. (eds.), *Art, Mind, and Religion*, Pittsburgh: University of Pittsburgh Press 1967
- Putnam, Hilary, *Mind, Language and Reality. Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press, 1975
- Putnam, Hilary, *Representation and Reality*, MIT Press 1991
- Putnam, Hilary, "A Philosopher Looks at Quantum Mechanics (Again)", *British Journal for the Philosophy of Science*, Vol. 56, Issue 4, 2005
- Radke-Yarrow, Marian, Belmont, Barbara et al, "Young Children's Self-Conceptions", in Cicchetti, Dante, Beeghly, Marjorie (eds.) *The Self In Transition: Infancy to Childhood*, Chicago: University of Chicago Press 1990
- Rainforest Conservation Society of Queensland (RCSQ), *Tropical Rainforests of North Queensland: Their Conservation Significance: A Report to the Australian Heritage Commission by the Rainforest Conservation Society of Queensland*, Special Australian Heritage Publication Series No. 3., Canberra: Australian Government Publishing Service 1986
- Ramsey, William, "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2016/entries/materialism-eliminative/> 2016-10-11

- Random House, "Intuition", *Random House Webster's Unabridged Dictionary*, Random House 2013
- Reichenbach, Bruce, "Monism and the Possibility of Life After Death", *Religious Studies*, Vol. 14, No. 1 (Mar., 1978)
- Reichenbach, Hans, *The Direction of Time*, Oakland: University of California Press 1956
- Reppert, Victor, "The Argument from Reason", in Craig, William Lane, Moreland, J. P. (eds.), *The Blackwell Companion to Natural Theology*, Oxford: Blackwell Publishing Ltd. 2009
- Riches, David, "Shamanism: The Key to Religion", *Man*, New Series, Vol. 29, No. 2, 1994
- Rickles, Dean, *Symmetry, Structure and Spacetime*, Amsterdam: Elsevier 2008
- Robinson, Howard, "Substance", *The Stanford Encyclopedia of Philosophy*, Spring 2014 Edition, Edward N. Zalta (ed.). Retrieved 2014-03-31 from <http://plato.stanford.edu/archives/spr2014/entries/substance/>
- Rosenberg, Alex, *The Atheist's Guide to Reality: Enjoying Life Without Illusions*, New York: W. W. Norton & Company 2011
- Rosenthal, David, *Consciousness and Mind*, Oxford: Clarendon Press 2000
- Ross, James F., "The Fate of the Analysts: Aristotle's Revenge\*: Software Everywhere" *Proceedings of American Catholic Philosophical Association* 1990
- Rovelli, Carlo, "Relational quantum mechanics", *International Journal of Theoretical Physics* 35 1996
- Rowe, M. A., Kielpinski, D. et al, "Experimental violation of a Bell's inequality with efficient detection, *Nature* 409, 2001
- Russell, Bertrand, *Analysis of Mind*, New York: Cosimo Inc., 2004
- Russell, Bertrand, *Principles of Mathematics*, New York: Routledge, 2010
- Ryle, Gilbert, *The Concept of Mind*, Chicago: University of Chicago Press 2000
- Schlosshauer, Maximilian, Kofler, Johannes, Zeilinger, Anton, "A Snapshot of Foundational Attitudes Toward Quantum Mechanics", *Studies in the History and Philosophy of Modern Physics* 44, 2013
- Schwitzgebel, Eric, "Belief", *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/belief/>, retrieved 2013-04-10
- Searle, John, "Minds, Brains and Programs", *Behavioral and Brain Sciences*, 3 1980
- Searle, John, "Is the Brain a Digital Computer?", *Proceedings and Addresses of the American Philosophical Association*, vol. 64, no. 3 1990
- Serway, Raymond, Jewett, John, *Physics for Scientists and Engineers, Vol 1*, Belmont: Brooks/Cole, CENGAGE Learning 2010
- Shay, J. W., Wright, W. E., "Hayflick, his limit, and cellular ageing", *Nature Reviews Molecular Cell Biology* 1, 2000
- Shoemaker, Sydney, "Causality and Properties", in van Inwagen, Peter (ed.), *Time and Cause: Essays Presented to Richard Taylor*, Dordrecht: D. Reidel Publishing 1980
- Shushan, Gregory, *Conceptions of the Afterlife in Early Civilizations*, London: Continuum Publishing 2009
- Skinner, B. F., "The Operational Analysis of Psychological Terms", *Psychological Review* 52, 4, 1945
- Slate, <http://www.slate.com/bigideas/is-there-life-after-death/essays-and-opinions> (SPT Immortality Project description and overview), accessed 2016-11-11
- Smart, J. J. C., "Materialism", *Journal of Philosophy* 60, October 1963
- Solecki, Ralph, *Shanidar: The First Flower People*, New York: Alfred Knopf Inc. 1971

- Solecki, Ralph, "Shanidar IV, a Neanderthal Flower Burial in Northern Iraq", *Science*, vol. 190, iss. 4217, 1975
- Sommer, Christoph, "Another Survey of Foundational Attitudes Towards Quantum Mechanics", <http://arxiv.org/pdf/1303.2719v1> 2013, retrieved 2016-07-20
- Spengler, Oswald, *The Decline of the West*, Oxford: Oxford University Press 1932
- Sprigge, Timothy, "Final Causes", *Proceedings of the Aristotelian Society*, Suppl. Vol. 45, 1971
- Sprigge, Timothy, McHenry, Leemon B. (ed.) *The Importance of Subjectivity*, New York: Oxford University Press 2011
- Stanner, William Edward Hanley, *The Dreaming and Other Essays*, Carlton: Black Inc. Agenda 2009
- Stapp, Henry P., *Mindful Universe: Quantum Mechanics and the Participating Observer*, Berlin: Springer-Verlag, 2011
- Stenmark, Mikael, *Rationality in Science, Religion and Everyday Life*, Notre Dame: University of Notre Dame Press 1995
- Stenmark, Mikael, "What is scientism", *Religious Studies*, v. 33, no. 1, 1997
- Stoljar, Daniel, "Physicalism", *The Stanford Encyclopedia of Philosophy*, Spring 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/spr2016/entries/physicalism/> 2016-11-08
- Strong, Michael J. (ed.), *Amyotrophic Lateral Sclerosis and the Frontotemporal Dementias*, Oxford: Oxford University Press 2012
- Sudduth, Michael, *A Philosophical Critique of Empirical Arguments for Postmortem Survival*, New York: Palgrave MacMillan 2016
- Swinburne, Richard, *The Existence of God*, Oxford: Oxford University Press 2004
- Swinburne, Richard, *Mind, Brain and Free Will*, Oxford: Oxford University Press 2013
- Taliaferro, Charles, *Consciousness and the Mind of God*, Cambridge: Cambridge University Press 1994
- Tegmark, Max, "The Interpretation of Quantum Mechanics: Many Worlds or Many Words?" *Fortschritte der Physik*, 46, 1998
- Tiwari, Kedar Nath, *Comparative Religion*, Motilal Banarsidass & Shri Jainendra Press 1983
- University of California Riverside, <http://www.sptimmortalityproject.com> (SPT Immortality Project website), accessed 2016-11-11
- Universität Tübingen, "Ivory Venus Figurine From The Swabian Jura Rewrites Prehistory", *Science Daily*, May 14 2009, retrieved from <http://www.sciencedaily.com/releases/2009/05/090514084126.htm> 2013-04-02
- Vanhaeran, Mariah, d'Errico, Francesco, Stringer, Chris, James, Sarah L., Todd, Jonathan A., Mienis, Henk K., "Middle Paleolithic Shell Beads in Israel and Algeria", *Science* 312 (5781): doi:10.1126/science.1128139. PMID 16794076, retrieved 2013-04-02
- Wallace, David, *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*, Oxford: Oxford University Press 2012
- Walter, Mariko Namba, Friedman, Eva Jane Neumann (eds.) *Shamanism – An Encyclopedia of World Beliefs, Practices and Culture*, ABC-CLIO 2004
- Watanabe, Satosi, "Symmetry of physical laws. Part III. Prediction and retrodiction." *Reviews of Modern Physics* 27.2, 1955
- William, Ramsey, "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), retrieved from <http://plato.stanford.edu/archives/win2016/entries/materialism-eliminative/> 2016-10-11

- Williams, Bernard, *Problems of the Self*, Cambridge: Cambridge University Press 1973
- Wilson, Nigel, (ed.) *Encyclopedia of Ancient Greece*, Routledge, New York 2010
- von Wright, G. H. *Explanation and Understanding*, Ithaca: Cornell University Press 1971
- Wynn, Thomas, "Did Homo erectus speak?" *Cambridge Archaeological Journal* 8 (1) 1998
- Yün-Hua, Jan, "Problems of Tao and 'Tao Te Ching'", *Numen*, Vol. 22, 3 1975
- Zeh, H. Dieter, "On the interpretation of measurement in quantum theory", *Foundations of Physics* 1, 1970
- Zerzan, John, "The Catastrophe of Postmodernism", in *The Anarchist Library*, (1990s, exact date unknown) retrieved from <https://theanarchistlibrary.org/library/john-zerzan-the-catastrophe-of-postmodernism.pdf> 2016-11-05
- Zimmerman, Dean, "Personal Identity and the Survival of Death", in Bradley, Ben, Feldman, Fred, Johansson, Jens (eds.), *The Oxford Handbook of Philosophy of Death*, Oxford: Oxford University Press 2012

ACTA UNIVERSITATIS UPSALIENSIS  
Uppsala Studies in Philosophy of Religion  
Editor: Mikael Stenmark

1. Maria Klasson Sundin, Barnets religionsfrihet, en villkorad rättighet?: en filosofisk undersökning utifrån barnkonventionen. 2016.
2. Lotta Knutsson Bråkenhielm, Religion – evolutionens missfoster eller kärleksbarn? Kognitionsvetenskaplig religionsforskning och dess relevans för religiösa trosföreställningars rationalitet. 2016.
3. Johan Eddebo, Death and the Self: A Metaphysical Investigation of the Rationality of Afterlife Beliefs in the Contemporary Intellectual Climate. 2017.

\*\*\*

Prior to 2002, studies in philosophy of religion from Uppsala university were published in a joint series with Lund university: *Studia Philosophiae Religionis*. The following books were published in this series:

1. Ulf Hanson, "Religious experience": en semantisk studie. 1973.
2. Carl-Reinhold Bråkenhielm, How philosophy shapes theories of religion: an analysis of contemporary philosophies of religion with special regard to the thought of John Wilson, John Hick and D. Z. Phillips. 1975.
3. Bo Hanson, Application of rules in new situations: a hermeneutical study. 1977.
4. Lars Haikola, Religion as language-game: a critical study with special regard to D. Z. Phillips. 1977.
5. Ulf Görman, A good God?: a logical and semantical analysis of the problem of evil. 1977.
6. Eberhard Herrmann, Der religionsphilosophische Standpunkt Bernard Bolzanos unter Berücksichtigung seiner Semantik, Wissenschaftstheorie und Moralphilosophie. 1977.
7. Eberhard Herrmann, Die logische Stellung des ontologischen Gottesbeweises in Charles Hartshornes Prozesstheologie und neoklassischer Metaphysik. 1980.
8. António Barbosa da Silva, The phenomenology of religion as a philosophical problem: an analysis of the theoretical background of the phenomenology of religion, in general, and of M. Eliade's phenomenological approach, in particular. 1982.
9. John A. Sealey, Religion in schools: a philosophical examination. 1982.
10. Jan Löfberg, Spiritual or human value?: an evaluation-systematical reconstruction and analysis of the preaching of Jesus in the synoptical gospels. 1982.
11. Håkan Thorsén, Peak-Experience, religion and knowledge: a philosophical inquiry into some main themes in the writings of Abraham H. Maslow. 1983.
12. Eberhard Herrmann: Erkenntnisansprüche: eine orientierende Erkenntnistheoretische Untersuchung über Fragen zum Verhältnis zwischen Religion und Wissenschaft. 1984.



13. Christer Norrman, *Mystical experiences and scientific method: a study of the possibility of identifying a "mystical" experience by a scientific method, with special reference to the theory of Walter T. Stace*. 1986.
14. Jari Ristiniemi, *Experiential dialectics: an inquiry into the epistemological status and the methodological role of the experiential core in Paul Tillich's systematic thought*. 1987.
15. Olof Franck, *The criteriologic problem: a critical study with special regard to theories presented by Anthony Flew, D. Z. Phillips, John Hick, Basil Mitchell, Anders Jeffner and Hans Hof*. 1989.
16. Martin Holmberg, *Narrative, transcendence and meaning: an essay on the question about the meaning of life*. 1994.
17. Anders Nordgren, *evolutionary thinking: an analysis of rationality, morality and religion from an evolutionary perspective*. 1994.
18. Stefan Andersson, *In quest of certainty: Bertrand Russell's search for certainty in religion and mathematics up to The Principles of Mathematics (1903)*. 1994
19. Dag Hedin, *Phenomenology and the making of the world*. 1997.
20. Olof Franck, *Tro och transcendens i Ulf Ekmans och Kristina Wennergrens författarskap: om teologisk realism och referentiell identifikationsteori i två samtida trosuppfattningar*. 1998.
21. Ulf Zackariasson, *Forces by which we live: religion and religious experience from the perspective of a pragmatic philosophical anthropology*. 2002.
22. Olof Franck, *Förtryckets grundvalar: norm och avvikelse i argument om homoseksuellas, invandrares och kvinnors rättigheter*. 2002.
23. Kirsten Grønlien Zetterqvist, *Att vara kroppssubjekt: ett fenomenologiskt bidrag till feministisk teori och religionsfilosofi*. 2002.

