

The semantic approach as an anti-physicalist renewal of the
explanatory gap problem in contemporary philosophy of mind

by

Adrienne Canning
Bachelor of Arts, Wilfrid Laurier University, 2007

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Arts

in the Department of Philosophy

© Adrienne Canning, 2013
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

Supervisory Committee

The semantic approach as an anti-physicalist renewal of the
explanatory gap problem in contemporary philosophy of mind

by

Adrienne Canning
Bachelor of Arts, Wilfrid Laurier University, 2007

Supervisory Committee

Dr. Jeffrey Foss, Department of Philosophy
Supervisor

Dr. Margaret Cameron, Department of Philosophy
Departmental Member

Abstract

Supervisory Committee

Dr. Jeffrey Foss, Department of Philosophy

Supervisor

Dr. Margaret Cameron, Department of Philosophy

Departmental Member

Contemporary philosopher, Joseph Levine, has argued that human phenomenological experience cannot be explained solely through the resources of neuroscience, and that a significant ‘explanatory gap’ exists between the rich features of human experience and scientific explanations of the mind. This thesis examines Guiseppina D’Oro’s novel suggestion that the gap exists, but that it is a semantic rather than an empirical problem. D’Oro argues that the ‘gap’ is a persistent philosophical problem because of its semantic nature, and that advances in neuroscience will fail to resolve the gap because its source is a conceptual distinction that is not marked by empirical difference. In the thesis I will discuss some virtues and difficulties with D’Oro’s thesis, and the implications her claim has more broadly for philosophers of mind.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	iv
Acknowledgments.....	v
Dedication	vi
Chapter One: An Introduction to the Explanatory Gap	1
Introduction.....	1
Materialism and Anti-Materialist Intuition.....	2
Materialism and Explanation	5
Kripke’s Anti-Materialism.....	8
The Explanatory Gap	11
Preliminary Critique.....	16
Chapter Two: An Introduction to the Semantic Gap	18
Introduction.....	18
The Semantic Gap.....	19
Conceptual Idealism and Anti-Materialist Intuition	23
Mental Causation in an Anti-Materialist Metaphysics	29
Chapter Three: A Critical Comparison of Semantic and Epistemic Approaches to the Gap	33
Introduction.....	33
All Kinds of Gaps	33
Metaphysical Baggage and the Gap.....	36
Explanation in a Theory of Mind.....	42
Mental Causation in a Theory of Mind.....	49
Concluding Remarks.....	56
Chapter 4: What the Semantic Approach Means to the Future of the Explanatory Gap ..	59
Bibliography	62

Acknowledgments

With sincere thanks to my supervisor, Dr. Jeffrey Foss, and Dr. Margaret Cameron, for your thoughtful support and criticism. This thesis would not have been possible without your patience and encouragement. I would also like to express my great appreciation to the Social Sciences and Humanities Research Council of Canada for helping to fund my project, the University of Victoria Graduate Student Society, and the Department of Philosophy for your help and guidance throughout my research.

Dedication

To my family

Chapter One: An Introduction to the Explanatory Gap

Introduction

The “explanatory gap” is a phrase coined by Joseph Levine in his 1983 paper, “Materialism and Qualia: The Explanatory Gap.” The phrase and title capture the idea that materialism is required to explain mental causation, but that it leaves the connection between the macroproperties and microproperties of our mental states unexplained. The explanatory gap presupposes that there exist qualitative properties of our mental states that are left unaccounted for by physicalist theories. Levine uses Kripke’s metaphysical argument against materialism to draw out an intuitive difference between identity claims that employ our concept of qualia and those that employ other rigidly designated concepts. Levine critiques Kripke’s use of this intuition of difference as a basis for drawing metaphysical conclusions, and presents what he considers to be a weaker epistemological thesis. Levine rejects Kripke’s claim that psycho-physical identity statements must be false because they are conceivably false in favour of the weaker claim that psycho-physical identity statements leave a significant explanatory gap. The gap thus attempts to shift the explanatory burden to the materialist to explain the mind-body connection, by claiming that materialism, if successful, would yield a realization theory of mind¹.

In this chapter I first outline the version of materialism that Levine endorses and explain the motivation behind his commitment to physicalism. I show that Levine’s endorses the ‘standard view’ that physicalism is required in order to secure a causal role for the mind,

¹ Levine defines a “realization theory” for a mental property as “one that shows us explicitly how the property in question is physically realized” (Purple Haze 16). He expands on the requirement later, stating that a crucial element is to identify the property with a causal role, and then to show that the property in question fulfills that role (Purple Haze 45).

and that his version of materialism requires the possibility of explanatory reduction. Secondly, I present Kripke's argument against materialism, which Levine uses to underpin his weaker epistemological charge that materialism has a significant explanatory shortcoming. Levine reiterates Kripke's view that the felt contingency of psycho-physical identity statements demonstrates a failure of materialism, but argues that this is only because of the epistemological failure to produce a realization theory of mind. Thirdly, I explain in detail the explanatory gap as a thesis against materialism, focussing in particular on the role of our concept of qualia in the perceived success or failure of explanatory reductions. We will see how Levine moves from the intuition about an explanatory difference in two types of identity statements to the more substantive charge that physicalism is inadequate to account for our concept of mind. Finally, I make some preliminary comments regarding Levine's failure to overcome the modal rationalist presuppositions for which he criticised Kripke, and demonstrate that, problematically, it is only by presupposing physicalism that Levine's arguments against physicalism gain intuitive traction in the first place.

Materialism and Anti-Materialist Intuition

Materialism is typically regarded as a thesis regarding the ultimate constituents of the universe, which states that the fundamental properties of physics (such as matter and energy) are the basis of all other properties. Philosophers are often drawn to this view for its ability to reconcile our beliefs that minds seem to make a difference in the world of matter, and that it is only if mental properties are somehow connected to physical properties that they can play this role. In *Purple Haze: The Puzzle of Consciousness*, Levine justifies his support for materialism on the grounds that "only if mental

phenomena are somehow constructible from, or constituted by, the physical phenomena that serve as the ultimate basis for all change in the distribution of matter and energy does it seem possible to make sense out of mental-physical causal interaction” (5). Of course, we could abandon our interactionist intuitions regarding the mind or endorse some version of epiphenomenalism², but as Levine puts it, “it seems overwhelmingly obvious that mental phenomena are both causes and effects of non-mental, physical phenomena” (Purple Haze: The Puzzle of Consciousness [PH] 5). Despite the apparent ability of physicalism to provide a causal role for mental properties, Levine thinks that the existence of an explanatory gap is a strong motivation not to endorse materialism (PH 76). For those who believe in the gap, the development of the physical story (through the advancement of physics and neurophysiology for example) does little to expand our understanding of subjectivity and consciousness because when it comes to the mind, there is something more that requires explanation than just the physical story.

Through the nineteen seventies, eighties, and again in the nineties, contemporary philosophers of mind were presented with various thought experiments meant to demonstrate the failure of materialism to explain mental phenomena. Beginning with Thomas Nagel’s “What is it like to be a Bat” in 1974, materialists were faced with the proposal that subjectivity, and in particular the qualitative features of subjective experience, could not be reduced to, nor be derived from, physical facts known from a third person perspective. Then, again in 1982, Frank Jackson’s “Epiphenomenal Qualia” challenged materialists with the prospect that there are facts about experience that are not physical facts, and properties of experience that can only be acquired from a first-person

² Levine defines ‘epiphenomenalism’ as “the doctrine that mental properties play no causal role in the production of behavior” (Purple Haze 22).

perspective. These thought experiments challenged the view that knowledge acquired from a third-person perspective, such as that derived from physics and chemistry, could adequately account for the phenomena of first-person experience. For it seems that given the whole physical story about one's qualitative experience, there are additional questions left unanswered, such as whether the physical facts could be accompanied by different qualitative experiences than they are, and whether any qualitative experience could exist in the absence of some particular physical base.

Perhaps the most well known and most cited anti-materialist argument in contemporary philosophy of mind is David Chalmers's zombie argument, as detailed in his 1996 work, *The Conscious Mind*. This reversal of the classic Cartesian thought experiment challenges the materialist to imagine a complete physical duplicate of herself, who despite being identical in all physical ways, does not have any qualitative experience. The 'phenomenological zombie' would act and react in all the expected ways when confronted with various stimuli, but is supposed not to feel anything at all. The apparent conceivability of a physical duplicate without qualitative states is meant to persuade us that it is possible that there are qualitative features of experience that are not subsumed under the realm of the physical sciences, because "if the physical or functional story really explained the qualitative character, it would not be so clearly imaginable that the qualia should be missing" (On Leaving out What it is Like [LO] 129). Thus, we are left with what Levine refers to as a 'Kantian antinomy of conscious experience': that we have good reason to believe both that qualitative features of experience are a species of physical phenomena and that conscious experience can't be captured in physical terms (PH 10). If the mind does not fall into the realm of the physical, it is unclear how it can

play a role in the production of behaviour and belief, but when reduced to the physical, an explanation of mind does not capture the subjectivity and rich qualitative features of experience.

Despite these challenges, Levine thinks that the fact that the physical story reveals little about our concept of mind should not deter our endorsement of materialism. In fact, he thinks “we don’t need a clear conception of the physical to formulate materialism. All we need is a clear, or even not-so-clear, conception of the mental” (PH 20). For Levine, the belief in a rich and determinate qualitative character of experience is irresistible, and that this feature is left unaccounted for by materialist theories seems like a serious epistemological challenge. However, it is not at all apparent that a referent to our concept of qualia even exists, let alone that it would constitute a failure of materialism. In *Purple Haze*, for example, Levine says both that we have a substantive and determinate conception of qualia and that we do not. On the one hand, “Qualia present a problem for reductive explanation precisely because there is a real content to our idea of a quale and not...because it is merely ostensive” (84) while on the other hand, when it comes to characterising qualia all Levine can do “is point at the phenomenon, using hand-wavy terms like ‘what it’s like...’” (132). For qualia to really pose a challenge to materialism one must justify the existence of qualia in a more substantive way than simply waving their hands and saying there is something that it’s like.

Materialism and Explanation

Levine ultimately embraces materialism because he finds its alternatives unpalatable.

Epiphenomenalism counters our most basic belief that the mind plays a causal role in a

material world, and functionalism³ has a tendency to trivialize the ‘substantive’ and intimate connection we have to qualia. Most importantly, “so long as we take mental properties to be causally relevant to the production of physical behaviour, and accept the principle that the fundamental physical properties provide the only causal bases there are for changes in physical properties, we have reason to believe that [materialism] must be true” (PH 16). The physicalist commitment to a closed causal system often carries with it the expectation that any part of the system is sufficiently connected to its other parts such that its occurrence can (at least in principle) be made intelligible by reference to the others. This epistemological demand is shared by Levine, for whom materialism minimally implies that every phenomenon is either describable in the terms of physics or, if not, can be made intelligible by some mechanism that is. For, if materialism is true, “what we should expect is bottom up necessity, both metaphysically and epistemologically” even with regard to qualia (PH 76).

Levine must be careful not to too closely tie ontological facts about a domain to the way we conceive of it, but he does concentrate on the link between materialism and epistemology to the extent that “the substantive nature of our conception [of qualia] provides that material for the substantive nature of our explanatory demand” (PH 8-9). Levine’s expectation is that it is reasonable to demand an explanation of the connection between qualia and their underlying physical states because ultimately consciousness needs to be explained at the microscopic level for it to be consistent with materialism, and materialism is the only way to secure an active role for qualia. For Levine, the high expectation of epistemic bottom-up necessity, derivability, and not just metaphysical

³ Levine defines functionalism as “the view that mental states, including qualia, are definable in terms of their causal roles – their causal relations with stimuli, behavior, and other mental states” (Purple Haze 96).

necessity, falls naturally out of the explanatory demands on materialism, since “if materialism is true we have reason to expect that any phenomenon can be explained by reference to the physical laws and principles that govern nature as a whole” (PH 76). Not only should we expect an adequate explanation to provide an understanding of the ontic features of a system, such as the relations of cause and effect responsible for an event, but it should also render the explanandum expectable in light of the explanans. That is, a good explanation should be able to explain not just events or regularities, but also the instantiation of properties such that their presence, absence and characters are rendered expectable. A completed theory of physics, for example, should also provide a theory of mind that explains the mind-body relation.

Since Levine thinks that consciousness arises at the macroscopic level, it would seem quite odd for him to treat it as a primitive fact or as something not in need of further explanation. In his seminal work on the gap, “Materialism and Qualia,” he states:

Materialism...implies explanatory reduction of at least this minimal sort: that for every phenomenon not describable in terms of the fundamental physical magnitudes... there is a mechanism that is describable in terms of the fundamental physical magnitudes such that occurrences of the former are intelligible in terms of occurrences of the latter... brute facts... will not arise in the domain of theories like psychology. (358-9)

It is this requirement of a realization theory that Levine reiterates in *Purple Haze*, where he describes the need for a theory that shows explicitly how a property is physically realised. If materialism is true, he argues, a realization theory ought to be at least in principle possible, and minimally, we should construct theories of mind such as to eventually yield a realization theory. Thus while Levine attempts to avoid the modal rationalist assumptions that underpin Kripke’s anti-materialist argument, he thinks that

the fact that we do not currently have a realization theory for the mind shows that materialism leaves an explanatory gap.

Kripke's Anti-Materialism

Kripke's argument against materialism is related to the above thought experiments in an important way, but shifts the focus from our intuitions about imagined scenarios to our intuitions toward theoretical identifications involving rigid designators. According to Kripke, there is an important distinction between psycho-physical identities and other theoretical identities. The difference is based on our ability to explain away the apparent contingency of the latter, but not the former, and has important consequences for materialism. Kripke's conclusion depends on two premises regarding identity statements: first, that all identity statements using rigid designators on both sides of the identity sign are, if true at all, true in all possible worlds where the terms refer; second, that psycho-physical identity statements are conceivably false, and are therefore, by the first claim, actually false (Materialism and Qualia: The Explanatory Gap [MQ] 354). Presuming for the moment that some version of externalism is true and that qualitative state types are rigidly designated, there does seem to be a tension between necessary a posteriori identities and the intuition that things could have turned out otherwise.

Consider that for Kripke it is conceivable (that is, apparently contingent) that heat is not identical to the motion of molecules (i.e. that the identity heat = molecular motion is false) because we can imagine something that affects our senses the way heat does that is not the motion of molecules. Kripke's explanation for this is that when we imagine heat without the motion of molecules, we are really imagining the phenomenon we experience as the effects of heat (such as the rising of thermometers and the sensation of warmth)

and not heat itself as being caused by something other than the motion of molecules. This same separation of cause from effect is not possible in the case of pain however, because according to Kripke, to imagine what we experience as pain is just to imagine pain itself. And, since there is no explanation of why it seems contingent that pain is the firing of c-fibres (of the sort offered in the theoretical reduction of heat to molecular motion), the identity statement is false.

One would assume that both heat and pain would analogously relate to their correlates in their respective identity statements since both are (according to Kripke) rigid designators. Rigid designation occurs when a term picks out an object in all possible worlds and picks out nothing in worlds where that object does not exist. For example, “since ‘heat’ and ‘molecular motion’ are both rigid designators, the identification of the phenomena they name is necessary” (Naming and Necessity [NN] 148) and similarly, since “‘pain’ is a rigid designator of the type, or phenomenon, it designates: if something is a pain it is essentially so, and it seems absurd to suppose that pain could have been some phenomenon other than the one it is” (NN 148-9). However, when we take a closer look at the apparent contingency of each identity, the analogy begins to fall apart. In the case of heat and molecular motion, the illusion of contingency can be explained by appealing to how the reference of the designator is determined. That is, if the referent and designator coincide only contingently, that explains why the identity statement (mistakenly) seems contingent (NN 150). According to Kripke, when someone says (inaccurately) that heat might not have been identical to molecular motion, what is true of what he says is that someone could have sensed a phenomenon in the same way we sense heat, by means of its production of the sensation we call ‘heat’, without molecular

motion. An analogous situation is not possible for the case of pain and c-fibers because while:

someone can be in the same epistemic situation as he would be if there were heat, even in the absence of heat...No such possibility exists in the case of pain and other mental phenomena. To be in the same epistemic situation that would obtain if one had a pain is to have a pain; to be in the same epistemic situation that would obtain in the absence of a pain is not to have a pain. (NN 152)

The possibility that stimulation of C fibres could not be accompanied by the feeling of pain is a flat out contradiction with the supposed necessary identity of pain with the corresponding physical state. The disanalogy is meant to persuade us that “in the case of molecular motion and heat there is something, namely, the sensation of heat, which is an intermediary between the external phenomenon and the observer. In the mental –physical case no such intermediary is possible, since here the physical phenomenon is supposed to be identical with the internal phenomenon itself” (NN 152).

We can summarize the above as follows: for Kripke, if an identity statement is conceivably false, it is actually false. Both identity statements, “Heat = Molecular motion” and “Pain = the firing of C-fibres” are conceivably false, but there is an important difference between theoretical and psycho-physical identity statements. What we really conceive of in the former case, and thus what is truly contingent, is not the identity of heat and molecular motion, but the identity of the phenomena we experience as heat and the motion of molecules. The statement “heat is molecular motion” is then not truly contingent, and we can explain why it seems contingent by separating the phenomena we experience *as* heat from heat itself. On the other hand, the statement “pain is the firing of c-fibres” is really contingent, because the phenomena we experience as pain *is* pain, and “we cannot make the distinction here, as we can with heat, between the

way it appears to us and the phenomenon itself' (MQ 355). Since we have no story about why it seems contingent to us that pain is the firing of c-fibres, it must really be contingent, and thus the identity must be false. It is the intuition that there is a substantive difference between psycho-physical identity statements and other theoretical reductions that is used to support Kripke's metaphysical thesis. My goal here is of course not to critique or defend Kripke, but to show how his theory laid the groundwork for Levine's proposal of an explanatory gap.

The Explanatory Gap

Levine's critique of Kripke's argument targets the underlying assumption that our intuitions, or for that matter, any of our rational thoughts, can and do track the way things really are. In particular, Levine points out that Kripke's second premise relies problematically on a 'particular intuition regarding conscious experience': that whereas we can distinguish between macro properties and the micro properties that constitute them to explain the apparent contingency in normal theoretical reductions (i. e. we can conceptually separate the phenomenon of heat from the motion of molecules), we cannot similarly distinguish conscious experiences and the constitutive brain states to explain the apparent contingency of psycho-physical identity statements (i.e. we cannot conceptually separate pain from the sensation of pain). The intuition is that because psycho-physical identity statements are contingent in a different way than other theoretical reductions, psycho-physical identities are different than other theoretical identities.

Levine thinks that although the intuition exploited by Kripke cannot be used to support his metaphysical conclusion, it can be used to support the weaker epistemological thesis that there is a gap in the explanatory import of psycho-physical identity statements. It is

from this 'gap' in the explanatory import of the latter type of statement that the phrase 'explanatory gap' is derived as an epistemological thesis. Levine thinks that our inability to explain away the contingency of psycho-physical identity statements like 'pain = the firing of c-fibres' highlights an important epistemological distinction between two types of identity statements. On the one hand, we have 'non-gappy', theoretical identity statements, which are 'fully explanatory, with nothing crucial left out' and on the other hand, we have 'gappy', psycho-physical identity statements, which seem to leave something crucial unexplained.

That same 'something crucial' is also not explained by a functionalist account that captures whatever fills the causal role of the mental property (e.g. pain). As the story goes, what is left out is an explanation of the qualitative character associated with pain, and an explanation of why the pain feels the way it does. This is because, Levine argues, it will still seem reasonable to ask, once the functionalist story is in, why a functional state like C-fibres firing is accompanied by *this* rather than *that* qualitative experience. He claims that "if there is nothing we can determine about C-fiber firing that explains why having one's C-fibers fire has the qualitative character that it does...it immediately becomes imaginable that there be C-fibers firings without the feeling of pain, and *vice versa*" (MQ 359). The possibility of missing or inverted qualia is taken by Levine to show that pain is not identical to its functional role. Levine identifies that psycho-physical identity statements and functional explanations leave unexplained both why pain feels the way it does and the connection between the qualitative side of pain and its base. In the case of psycho-physical identities, "there seems to be no discernible connection between the physical description and the mental one, and thus no explanation of the latter

in terms of the former” (PH 77). It is this felt contingency that Levine thinks makes psycho-physical identity statements subject to Kripkean style objections and creates “gappy identities”: those identity claims that admit of an intelligible demand for explanation.

We can spell out the difference between successful and unsuccessful explanatory reductions more clearly using the psycho-physical identity statement “Pain = the firing of C-fibres” and the theoretical identity statement “Heat = the motion of molecules” as examples. Levine takes the statement “Heat=molecular motion” as a sufficiently explanatory statement. He names a number of features of the identity claim that make it successful: first, that it tells us by what mechanism the causal functions we associate with heat are affected, and second, that along with the relevant background knowledge, it makes intelligible how something like the motion of molecules can play the causal role we associate with heat. The success of the explanation is largely based on that (according to Levine) even antecedent to our discovery of heat’s essential nature, our concept of heat is exhausted by the phenomenon’s causal role. That is, we conceptualize heat as that thing that raises the mercury level in the thermometer and causes water to boil. He argues that it would seem absurd to ask for a further explanation of why heat feels the way it does after given the causal story, because we can easily come to realise that the way heat appears to us is explained by its microphysical properties. So, though it seems that identity statements involving heat and molecular motion are contingent “one can become disabused of this notion by noting that instead of imagining *heat* without the motion of molecules, one is really imagining there being some phenomenon that affects our senses in the way heat in fact does, but is not the motion of molecules” (MQ 355). For Levine,

as for Kripke, it is our ability to separate the appearances from the underlying phenomenon that allows us to sufficiently explain how molecular motion epistemically and metaphysically necessitates heat.

In “Leaving Out What it’s Like” Levine states that similarly to the story about heat, an adequate explanatory reduction of qualia needs (1) a property being reduced, and (2) a property or set by which (1) is normally picked out. This is difficult in the case of qualia, argues Levine, because in the case of qualia the distinction between (1) and (2) is collapsed. That is, qualitative states serve as their own modes of presentation, or have ‘substantive’ as opposed to ‘thin’ modes of presentation. For example, “heat” is a thin concept because it merely labels a substance/phenomenon in the world, whereas “pain” has a substantive or “thick” mode of presentation that is not exhausted by an ostensive definition (PH 83-4). A thin mode of presentation is one whose superficial properties exhaust the nature of the thing in question. To the extent that there is a non-causal element (i.e. qualia) in our concept of qualitative character, it will escape the explanatory net of physicalistic reduction because, according to Levine, “our concepts of qualitative character do not represent... causal roles” (LO 134). However, a successful reduction of the mental must still account for its causal role unless we are willing to accept epiphenomenalism. Therefore, to adequately account for the causal role of the mental in a materialist theory, we must first reduce it (or explain how it is related) to its microscopic physical properties. What stands in our way of accomplishing this according to Levine, is that “unlike other macro domains, when it comes to qualia, we are not lacking merely enough detail to provide the requisite explanation, but any idea of how such a theory might go” (PH 69).

So what is it precisely that our concept of qualia refers to if not just causal role? For Levine, there are two interrelated features of conscious experience that both resist explanatory reduction to the physical: subjectivity and qualitative character. In particular, these two features present themselves as challenges to explaining how it is that consciousness, or more specifically, qualia, are realized at the physical level. The first feature picks out the intuition that when qualia are instantiated, they are related in such a way to the subject so as to generate an explanatory gap. The second feature poses a challenge to explain why qualia, as properties of experience, don't appear to be constructible out of basic physical properties. For Levine, an appeal to causal/nomological relations doesn't capture the account of "cognitive intimacy" we need for an account of conscious cognition (PH 173-174). But as discussed above, we haven't been given a robust justification for our belief in qualia other than the disputable claim that there is 'something that it's like' to occupy perceptual states.

It is here that Levine's strategy seems dangerously close to that which he critiqued Kripke for employing. A significant factor in Levine's insistence on the failure of explanatory reduction is based on his appeal to two modes of presentation via which, he believes we discover the distinct properties of the mind (its causal versus its qualitative aspects). He states that "The point is that an intelligible request for explanation seems to entail a distinction in properties somewhere. If it isn't to be located in the properties of the one thing we're representing on both sides of the identity sign, then it must be that the terms flanking the identity sign themselves represent different things" (PH 87). Levine thinks there must be two distinct properties corresponding to our two concepts of the mental because a purely conceptual or linguistic distinction between the mental as causal

and the mental as qualitative fails to account for the substantive nature of our rich and intimate inner lives. For Levine, phenomenal concepts have a special nature, one that cannot be explained in physicalistic terms because of their ‘thick’ substantive character. But Levine is careful not to treat this as a death-blow for materialism, and instead treats gappy identities as motivation to believe in distinct properties that create an explanatory gap. However, in distinguishing the qualitative from the causal concepts of mind, Levine admits he finds the temptation to think there is a genuine distinction of properties corresponding to each representation “cognitively irresistible.” Ultimately, he decides that the question remains open whether the distinct concepts we have correspond to distinct properties of the mental.

Preliminary Critique

Levine, through his transformation of Kripke’s metaphysical argument into an epistemological claim, attempts to present a metaphysically neutral argument against psycho-physical identity based on the failure for psycho-physical identity statements to explain certain features of our experience. These features, often referred to as qualia, subjectivity or phenomenal consciousness are flatly rejected by some scholars and clung to by others, and are not uncontroversial in themselves. However, in giving Levine the benefit of the doubt that we indeed are faced with two concepts corresponding to our mental lives, we find ourselves back in the practice of basing our epistemic demands on criteria that make sense only from within a particular metaphysical framework. Only when we presume that mental concepts are ontologically connected to their referents in a way that is different than non-mental concepts are, is there something unexplained by their reduced concepts, and only if we presuppose that materialism is correct to begin

with that this is problematic at all. To get past the fairly insignificant charge of an explanatory gap, to the claim that it indicates something more than a peculiar linguistic feature of our conceptual framework, Levine inevitably falls victim to his metaphysical intuitions and ultimately, to his own critique. For only by introducing qualia as irreducible properties of a causal mind does the gap even begin to propose a challenge to materialism. We must then question why the charge of an explanatory gap has gained so much traction in the literature and why the charge of an explanatory gap is so compelling.

Chapter Two: An Introduction to the Semantic Gap

Introduction

The idea that an explanatory gap exists between the mind as immaterial and mind as causal has prompted significant debate in contemporary philosophy of mind. It is a matter of debate whether a gap exists, and if it does, what its source might be, and whether we are likely to find its solution. The suggestion that the gap may have a semantic source is a novel approach to the problem, outlined by Giuseppina D'Oro in her article, "The Gap is Semantic, not Epistemic". The article, though brief, mounts an interesting and not-uncontroversial critique of the typical approaches taken toward studying the mind-body problem. In particular, D'Oro argues that the explanatory gap is an essentially semantic problem, rather than a metaphysical or an epistemological one, and that its persistence and persuasiveness can be attributed to its essentially semantic nature. Unlike Joseph Levine, who claims that our future discovery of a psychophysical link will resolve the gap problem, D'Oro claims that the gap's essentially semantic nature means that future advances in the sciences are irrelevant to its resolution.

In this chapter, after a brief review of the explanatory gap, I first introduce D'Oro's direct reply to Levine in which she says that there is a gap, but it does not emerge because of the way we access internal mental states. D'Oro thinks the gap cannot be resolved by scientific advances because the source of the gap is a semantic distinction that applies where no empirical distinction exists. In the second section, I present D'Oro's defence of conceptual idealism, which claims that ontology is contextually dependent and that because the distinction between mind and body is semantic, there is no psychophysical link and no possibility of mind-body reduction. In the final section, I discuss D'Oro's

claim that physicalism is not required to secure a causal role for the mind. Whereas contemporary philosophers such as Levine see physicalism as the only way to make sense of mental causation, D'Oro shows that idealism is able to accommodate our intuitions of mental causation without the necessity of mental to physical reduction. The reader will see that D'Oro's conceptual idealism has the consequence that both explanatory reduction and mind-body interaction are not possible, and that not only might the gap not exist because of the epistemological challenges identified by Levine, but that if D'Oro is correct, the gap also has no chance of being resolved.

The Semantic Gap

In our introduction to the explanatory gap in Chapter one, we saw that Levine thinks materialism is required in order to secure a causal role for the mind in nature. He believes that because physicalism must be true we are reasonable to expect the possibility of explanatory reduction, and that the current lack of such an explanation results in a significant explanatory gap. Though he attempts to avoid the modal rationalist presuppositions for which he critiques Kripke, it is only after presupposing materialism that Levine makes his case against it. On the one hand, he presumes that physicalism is the only way to accommodate our intuitions of mental causation, but on the other hand, argues against it for its failure to explain qualia. Ultimately, Levine presents a weaker epistemological case against physicalism because it is unable to explain the relation between these subjective and objective features of our mental lives.

D'Oro enters the debate from a non-materialist background, and offers a fresh set of eyes on a recurrent philosophical conundrum. Rather than approach the gap as an ontological challenge to reduce mind to body or as an epistemic challenge that results from how we

access two different types of properties, D'Oro understands the explanatory gap as the challenge "to explain what it is that makes statements about mind-body identity more problematic than scientific identity statements" (The Gap is Semantic, Not Epistemological [TGS] 168). This characterization is unproblematic, since both the metaphysical and epistemological arguments for a gap are motivated by the failure to capture phenomenological features of the mental in statements about physical processes of the brain. One might expect mind-brain identity statements to provide successful reductions because of successful reductions that have been carried out in the sciences, for example, in the explanatory reduction of water to H₂O or of heat to molecular motion. But, as all proponents of the gap have argued, unlike scientific identity statements, psycho-physical identity statements leave their apparent contingency unexplained. To be successful, a theory of mind is expected to account for both the causal efficacy of the mental and the autonomy of the mental within a single theoretical framework.

As we saw in chapter one, at least two types of response have been offered to explain the contingency of psycho-physical identity statements. D'Oro labels these 'the metaphysical argument' and 'the epistemological argument'. The metaphysical argument is that which is offered by Kripke and Jackson, who combine the apparent contingency of psycho-physical identity statements with the additional assumption of conceivability as a guide to possibility to support the conclusion that mind and body are metaphysically distinct (TGS 168). D'Oro's semantic approach sees this technique to be fundamentally flawed not because of its use of conceivability arguments per se, but because of the methodological

assumption being made about the relationship between statements about the mind and statements about the body⁴.

D'Oro characterizes the epistemological interpretation of the gap offered by Nagel and Levine as the claim that “even if physicalism were the whole story metaphysically speaking, it cannot be the whole story epistemically speaking because introspection yields...phenomenology that is unaccountable from within the objective perspective” (TGS 169). The epistemological thesis is evidently, not a full blown rejection of physicalism (i.e. not a claim about the truth of physicalism), but rather a claim about the physicalist's inability to explain the phenomenal character of experience. However, like the metaphysical argument, it also presupposes a view of metaphysics as an ‘under-labourer of science,’ in which ultimately, metaphysical questions are whittled down to questions about the physical constituents of reality and then described from an objectivist perspective (TGS 170). D'Oro discusses this phenomenon in “Collingwood's ‘Solution’ to the Problem of Mind-Body Dualism,” where she argues that the claim that the mind-body problem after the 1950s and 1960s deals with the metaphysical problem (rather than the conceptual or linguistic problem) is a shift from a debate that recognized the presuppositions of each science to one that did not (359). This interesting historical point is mentioned to shed light on the development of the contemporary mind-body debate, in which there is little consensus on what problem the mind-body debate seeks to resolve or on what the proper methods are for undertaking the challenge.

For D'Oro, the idea that the mind-body debate moved from asking conceptual questions to asking ontological questions marks an important failure to recognize that the mental

⁴ As discussed in chapter one these approaches prioritize the ontology of natural science despite a belief in methodological pluralism. We might also consider here that Collingwood distinguishes thoughts not by spatio-temporal properties, but by conceptual content.

causation debate is not about mind-body interaction, but rather one about two types of explanation. This explains D'Oro's staggering reply to Levine that "The mind-body gap is neither ontological nor epistemological, but semantic" (TGS 168). By this she means that the problem of mind-body dualism, and the related problems of mental causation and the explanatory gap, do not arise because of a substantial or metaphysical distinction between minds and bodies, but because of a semantic distinction applied where there is no empirical distinction. D'Oro doesn't disagree that there is an epistemic gap at issue in the mind-body debate, but she distances herself from this view because she doesn't think that the gap emerges because of a failure to find the psycho-physical link that connects first and third modes of epistemic access.

As discussed above in response to the metaphysical arguments of Kripke and Jackson, the semantic thesis does not adhere to a view of metaphysics in which mind and matter can be related in such a way that there can be a psycho-physical link⁵. And, as we will see in the next section, D'Oro argues that because mind and matter exist as two ways of looking at the same thing, explanatory reduction is impossible because it is not possible for a concept employed by a theory of mind to retain its semantic content in a purely physical theory. We cannot simultaneously explain an action mentally (e.g. as intentional action) and physically (e.g. as a causal chain of events) because there is no empirical distinction when the intension of each description is altered to be compatible with the alternate theory. Since there is no empirical distinction in question, to change the semantic content of the terms involved in the reduced and reducing theories is to abandon a distinction

⁵ Though Collingwood (Method, 52) says that the various elements must somehow be interrelated, but it would be a significant task to pull out precisely what this connection might be. Due to interpretational conflicts in reading Collingwood and the enormity of such a task, I will consider this beyond the scope of my current project.

altogether. This is not to say that an action cannot have both a complete and truthful mental and a physical explanation, it just means that if one uses the vocabulary of the physical sciences they are not dealing with the same type of thing as when they use mental descriptions.

Conceptual Idealism and Anti-Materialist Intuition

D'Oro's view of metaphysics is borrowed from R.G. Collingwood, who, in *The New Leviathan*, makes two clear statements that ontology is dependent on methodological and conceptual context. First, in 1.83-4, he says that man as body is whatever the sciences of body say he is, and that man as mind is whatever he is conscious of being. Second, in 2.44, that all of man is body insofar as he approaches the problem of self-knowledge by the methods of natural science, all mind insofar as he approaches the problem of self-knowledge by expanding and clarifying the data of reflection. This is not just to say that there are multiple descriptions of his being, but that he has *being* in two different ways. Collingwood's brand of idealism suggests that since certain concepts exist only from within a particular methodological framework, and since there are at least two distinct frameworks from which we can investigate man's nature, there are at least two distinct ways for man to "be". These are not mere descriptions, but rather concepts that are constitutive of existing in some way⁶. As a result, these concepts cannot be redescribed in the vocabulary of an alternate framework and retain their meaning because the methods and presuppositions of the framework are a precondition of their existence.

⁶ D'Oro points out that this is closely tied up with our willingness to accept the synthetic/analytic distinction and that there are non-empirical concepts.

To persuade the reader that mind and body are the products of two types of explanation, D'Oro considers the distinction between an event when viewed on the one hand, using the presuppositions of physics, and on the other hand, using the presuppositions of psychology. Take for example a simple event involving the motion of your arm toward a window frame and the subsequent opening of the window. There is a complete description of this scene available from the domain of physics, involving a causal chain of events, which can be traced from the activity of firing neurons, to the activation of the nerves and muscles in the arm, to the application of force to the window and the movement of the window upon the application of a certain amount of pressure, etc. In the domain of physics, this (or some similar yet more complicated description) would suffice to determine what happened in this situation. For example, one would not be pressed to explain your being warm or the belief that opening the window will cool you down when asked for an event description in the lab, because the study is undertaken with the goal of producing physical event descriptions.

Despite that this causal chain offers a complete description of the event, there is an alternative account of what happened here which is compatible with, but neither explanatorily tied to, nor reducible to, the physical description. From the perspective of folk psychology, an explanation of what happened here has little to do with the chemical processes and conditions satisfied in various neural networks, and everything to do with the intentions of the actor as determined by a reflective observer. That is, what happened was that our subject wanted to open the window because they were warm, and opened it because they believed a breeze through an open window would relieve them from discomfort. This type of description, involving reasons and beliefs, has no interest in the

physical causal processes underlying the purposeful action because “causal explanations of the emergence of consciousness, even if they could be found, are simply not explanations of mind” (TGS 178). The semantic thesis is committed to the claim that physico-chemical processes are irrelevant to the folk psychological description of what is going on.

Many human actions have both a rational and causal explanation – consider for example, my opening the car window while I drive. One might say that what happened here was that I was warm, and believed that by opening the window I would find relief. The rational explanation is true and complete, and doesn’t require any input from scientists to tell me what happened. Yet alternatively, one might say that what happened when I opened the window was that there was a causal chain initiating with electrical impulses, nerve excitation and muscle activation, terminating in the application of force to an electric switch, which lowered the window. This scientific description is also complete, and doesn’t require any explanation of my intentions or reasons for acting. The point is that the difference between the two descriptions is not based on the empirical facts, which are the same, but on the concepts we apply a priori, that is, whether we look at it as a mental or physical phenomenon. This is why Collingwood and D’Oro treat epistemology as first science – because in order to engage in any investigation, we must first have determined the criteria we will use to establish truth.

It is important to recognize that for Collingwood, each type of explanation on this view is a full explanation of the event in question, because the considerations offered by its alternative are independent of, and irrelevant to, the description on offer. In D’Oro’s words, “explanations in different sciences are not rival explanations because they do not

explain the same (sort of) thing... they provide complementary rather than mutually exclusive accounts” (Collingwood’s “Solution” to the Problem of Mind-Body Dualism [CSP] 355). The reason this is possible is that there is no empirical difference between the two states of affairs, while there are important methodological and semantic distinctions in play. When applied to the explanatory gap debate, the entailment of this claim is that “knowledge of the psycho-physical relation is not so much beyond our cognitive grasp as irrelevant to our understanding of mind... because our concept of mind is logically required in order to identify any given particular as mental” (TGS 177)⁷. Even if we presume that there is a real epistemic gap caused by how we access mental states (e.g. we introspect our reasons for acting or attribute those states to others) and how we access empirical states (e.g. we open the brain and see neural pathways firing as part of the decision making process), if we attempt to reduce one state to the other we will always feel that something is lost in the process. That is because the semantic distinction between mind and body doesn’t depend on whether we access the states from a third or first person perspective as much as it depends on the a priori distinctions we make between what is mind and what is body. Even in assessing my own first person experience I have the choice to view myself as a collection of physical bits and processes or as a rational being persisting through a series of meaningful events. It is because philosophical classifications are concerned with concepts from the point of view of their intension not their extension, that “it is not possible to reveal the distinction between such concepts empirically” (Collingwood on Philosophical Knowledge and the Enduring Nature of Philosophical Problems [CPK] 103).

⁷ Brie Gertler (2002) defends dualism on similar grounds.

The idea that philosophy is concerned with conceptual distinctions without empirical difference is an important feature of Collingwood's metaphysics that D'Oro borrows in her articulation of the semantic thesis. To understand this claim, we can first contrast the main activity of philosophy from the main activity of science. According to D'Oro, the physical sciences are concerned with discovering empirical classifications in which "the coordinate species of an empirical genus form mutually exclusive classes," and "cannot be extensionally equivalent" (TGS 172). For example, within the genus "mammal" there are no objects that are both goats and cows, no objects in which 'goat' and 'cow' are jointly instantiated because there is no overlap between species distinctions within a genus. This is because science makes conceptual distinctions that are intended to cut nature at the joints, unlike philosophical distinctions which allow for extensional overlap, and can coexist in their instances⁸.

To illustrate a philosophical distinction, D'Oro gives two examples. First, the example of a Bob Dylan song that is at once both an instance of poetry and an instance of music, and secondly, the action of a shop keeper who returns his customer's overpayment, that at once illustrates the principle of duty and the principle of utility. In the first case, being music does not exclude a song from being poetry because it is possible for both concepts to be illustrated by a single activity. In the second case, the shopkeeper's action being dutiful doesn't exclude it from being useful. This is precisely because the aesthetic concepts, 'music' and 'poetry' and the philosophical concepts, 'duty' and 'utility' are not empirical classifications, they are semantic ones. Philosophical distinctions allow for extensional overlap, because they are intended as semantic distinctions to which

⁸ This doctrine is closely related to Aristotle's doctrine of the overlap of classes, in which two concepts can be jointly instantiated in a single instance.

correspond no empirical classifications (TGS 172-3, CPK 102). To assume that philosophical distinctions are classificatory in the same way as scientific ones is a fallacy for Collingwood, and it is precisely when we forget this distinction between philosophical and non-philosophical employment of concepts that philosophical problems arise (CPK 105).

Though the idea that philosophy is concerned with classes which overlap in their instances traces back at least as far as Aristotle, D'Oro does consider two possible objections to the suggestion that the overlap of classes is a distinctive feature of philosophical concepts. First, she acknowledges that it may be argued that some empirical concepts do in fact allow for extensional overlap. But, she argues, this is easily refuted if we consider that empirical concepts that seem to overlap, such as "has a beak" and "has feathers" in birds, are not coordinate species of the same genus (TGS 173). Second, it may be suggested that the existence of hybrids seems to suggest that the overlap of classes does occur in empirical genera. For example, in the classification of T-Rex into the class of predator or scavenger, the dispute was ultimately settled by saying that T-Rex is both, and a hybrid classification created to accommodate the discovery (TGS 175). But, argues D'Oro, this doesn't cast doubt on the claim that philosophical classifications differ from empirical classifications because philosophical concepts simply are not jointly instantiated in the same way in which empirical concepts are. Philosophers don't assume that when they find an overlap of classes that they need to restructure their classification systems, because "Philosophical disputes address semantic questions...and opposing positions in philosophical debates...do not carve up reality along the lines of empirical kinds, but along the lines of distinctions of reason" (TGS 177). It is

this belief in the possibility of extensional identity despite intensional difference that D'Oro then applies to the explanatory gap debate.

Mental Causation in an Anti-Materialist Metaphysics

The semantic thesis, as a meta-philosophical view, does not see the mind-body problem as a special kind of problem involving the different ways we access mental and physical states, nor a metaphysical problem of how mind and body interact. It regards the mind-body distinction as “similar to other purely intensional distinctions found in other branches of philosophy such as ethics and aesthetics” that has been treated according to the mistaken assumption that mind and body exist as parts of a layered ontology like is presumed by supervenience and anomalous monism (TGS 177). But we should not be led to conclude that there is no ontological difference between motives and causes simply because there is no empirical difference underlying the semantic distinction, nor that motives will eventually be explained by a theory of physical causes.

One of the most pressing challenges for philosophers engaged in the mind-body debate is to explain how human actors can exert purposeful actions on the physical world, and how this causal activity can be reconciled with our understanding of causes in the domain of physics. The suggestion that there are multiple senses of the word ‘cause’ is perhaps one of the more interesting claims underlying the semantic thesis, and offers one way of reconciling these two drastically different concepts of causation. Borrowed directly from Collingwood, but applied by D'Oro to the contemporary mind-body debate, the proposal that there are at least three, mutually exclusive, meanings of the word ‘cause’ gives a surprisingly intuitive perspective on why different disciplines speak of causes in such incompatible ways. These three senses each offers a different type of explanation of

events, which correspond to each science's investigative goals. For the philosopher or folk psychologist, for example, causes are teleological descriptions that pull actions under causal descriptions insofar as they are rationalised. Reasons are causes, but not in the same sense as in the physical sciences. For the scientist, causes either are those things which we have no power to control (e.g. in the domain of physics) or those things which we have power to produce or prevent (e.g. in medicine and engineering). D'Oro argues that philosophers must accept that for reasons to be understood as causes "the meaning of cause must be redefined as logical ground rather than temporal antecedent" (Collingwood and the *Metaphysics of Experience* [CME]). Reasons just are not the type of thing that can be empirically isolated or stand in order of spatio-temporal priority to other reasons; they are not Humean causes, they are logical causes.

Since the distinction between mental and physical events requires one to use two different semantics (or to use Collingwood's phrase: phases) for the same word, 'cause', it is no wonder that philosophers have been confused at the failure of physical descriptions to capture the sense provided by mental descriptions. In fact, if we accept that there are different senses of 'cause' it becomes clear that an explanatory reduction of mental causes to physical causes is impossible. In a mental description of a cause, what is meant by cause is something like 'motive' and this is not intended to capture the same type of causal process as indicated by the physicists term 'cause', meaning the necessary and sufficient conditions to bring about some state of affairs. According to D'Oro, it is precisely because "we mean very different things when we speak about the mind and when we speak about the body" that the mind-body problem is unlikely to wither as our

scientific knowledge advances (TGS 169). It is when we forget that causal descriptions in one domain are relative to its methods and presuppositions that we run into difficulty.

The idea that ontology is dependent on methodology has at least two important consequences for the explanatory gap problem. First, it requires a rejection of the assumption that at bottom, a physicalist ontology underlies all metaphysical claims, since it would mean that what man is depends on how we approach the question of his being. If we accept this proposal, the result is that one cannot reduce mind to body, as each exists only as part of independent methodological frameworks. That is, to question man's existence qua bodies, we must first engage the pursuit with the methodological presumptions that allow us to study his being that type – empirical study and a belief in the regularity of nature, for example. To question man's existence qua minds – we presume rationality and a logical connection between man's thoughts and actions. For Collingwood, man's body and man's mind are not two different things, but “one and the same thing, man himself, as known in two different ways” (The New Leviathan [TNL] 2.43). Thus, to attempt a reduction of mind to body is to mistakenly assume that mind and body are related hierarchically rather than as the subject matter of independent methods on equal ontological footing.

The second consequence of this view of metaphysics is that there is no meaningful question about mind-body interaction aside from the question of how theories of mind and theories of body interact/sect. D'Oro thus sees herself as disposing of the popular presupposition that mind and body are primarily ontological categories rather than explanatory categories⁹. In “Collingwood's Solution to Mind-Body Dualism,” D'Oro

⁹ Collingwood claims mental and physical explanations are genera are the same species: explanation

claims that modern philosophers of mind, in attempting a ‘metaphysical’ solution to the problem of mind-body interaction, have failed to acknowledge that the distinction between the mental and the physical is a distinction between types of explanation, not one between explanatory and metaphysical questions. This is why, in D’Oro’s view, the question of the relation between the mind and body is a question for philosophers, not scientists. “The ‘indirect relation between body and mind’ is the relation between the sciences of body, or natural sciences, and the sciences of mind; that is the relation inquiry into which ought to be substituted for the make-believe inquiry into the make-believe problem of ‘the relation between body and mind’” (TNL 2.49). The gap emerges because ‘mind’ means ‘that which can be explained under the laws of reason’, and ‘body’ means ‘that which can be explained under the laws of science/physics’ even though mind and body are extensionally equivalent. If we accept this version of the nature of the explanatory gap, the consequence is that the gap is not something that can be resolved. In particular, it cannot be resolved in the way anticipated by the epistemological argument because the concepts of mind and body are not differentiated primarily by two modes of access and cannot retain their distinct meanings if one is reduced to the other, or through advances in neurophysiology. In Collingwood’s words, the problem is that “What is assumed is that man is partly body and partly mind. On this assumption questions arise about the relations between the two parts; and these prove unanswerable” (TNL 2.42). Thus, the semantic thesis is not an attempt to solve the mind-body problem empirically, nor a deflationist attempt to solve it a priori. It is an attempt to show why it is here to stay (TGS 178). In the following chapter we will discuss the benefits and disadvantages of D’Oro’s semantic approach.

Chapter Three: A Critical Comparison of Semantic and Epistemic Approaches to the Gap

Introduction

In the first two chapters we have been introduced to two unique approaches to the ‘gap’: the epistemic and the semantic. We now may consider whether these different viewpoints, each with its own method and aims, gives any future hope of resolving the explanatory gap problem, or at least clarifying for philosophers what kind of problem the gap is, and how we ought to approach our study of it. In this chapter I first review various approaches that have been taken to studying the gap. I concede that to engage in the debate, we must accept that there is a gap to be accounted for, but that a variance in theories gives us reason to suspect that the problem is not entirely clear. In the second section I then look more closely at D’Oro’s and Levine’s approaches to the gap. In particular, I highlight that their divergent views on the source of the explanatory gap is the direct result of their different metaphysical views, and that the conceptual idealism that underpins the semantic approach to the gap results in its being irresolvable, while a physicalist position entails the possibility of a reductive resolution. In the final sections I look more closely at the role that explanation and mental causation play for each author in the development of their versions of the gap. Ultimately, I conclude that while D’Oro’s semantic gap provides numerous explanatory advantages, it does not adequately account for what role qualia should play in a philosophical theory of mind.

All Kinds of Gaps

It will be obvious to the reader from the preceding chapters that there is no generally accepted methodology for studying the gap between first person phenomenal statements

of experience and third person descriptions of human experience. The gap is studied by the neuroscientist as an empirical problem of how to find and interpret functioning regions of the brain. It is studied in philosophy metaphysically, as a question about how minds and bodies interact; epistemically, as a question of how to provide a deductive, predictable science of mental events; and semantically, as a study of how philosophical theories of mind and scientific theories of mind fit together. In each field researchers attempt to study the mind under their respective methods and background assumptions, and take as their subject matter that which is accepted within their discipline as 'mental'. Not only are there differing methodologies for studying the gap, but divergent aims as well. In the case of the metaphysician, exemplified by Kripke and Jackson, a science of mind needs, but fails to explain, the 'what it's like' of firsthand phenomenal experience with the same clarity and precision offered by scientific reductions. An investigation into the matter revolves around discovering the mind-body connection, and in particular, explaining what types of things minds must be in order to have such an intimate and rich interaction with the physical world. In the case of the epistemologist, exemplified by Levine, a science of mind must, but fails to, explain the qualitative features of mind that seem missing from a description of all the physical facts. For the epistemically-focused philosopher such as Levine, what is needed is a theory of not only how mental properties and physical properties interact, but how their interaction can be explained in basic terms of the physical sciences. But for the conceptual idealist, exemplified by D'Oro's semantic approach, a study of the gap requires a closer look at what we suppose to fall under the heading 'mind' and what methods are appropriate for studying it. Her metaphilosophical view tells us to step back and ask if our aims for a theory of mind are suited to the subject

matter with which we are presented, and whether our subject matter is of the type that can be a candidate for either epistemic or ontological reduction. This is certainly not an exhaustive list of approaches that have been taken to studying the explanatory gap¹⁰, but the variance demonstrated already between Levine's and D'Oro's approaches provides enough motivation to look more closely at what we presume to be the source of an explanatory gap in philosophy of mind.

To engage in the debate, let us first concede that there is some kind of explanatory gap. The most prominent historical and contemporary philosophers, from Rene Descartes to David Chalmers, have argued in favour of one type of gap or another, so in order to advance the debate, let us provisionally assume that there is an unexplained feature to be accounted for within philosophy of mind. Minimally, we can assume that the gap can be studied and explained from metaphysical, epistemic, or semantic backgrounds, but hinges at its most basic on the failure of philosophers to provide a theory that explains the relationship between mind and body. What the relata are specifically will depend on from which view we are addressing the challenge; for example, as one between two ontological categories, mind and body, one between mental properties and empirical properties, or one between statements made within discrete methodological contexts of inquiry. One can then add additional challenges, such as the requirement that this theory be able to predict mental events, fit a particular model of explanation, connect theories of mind and body in a particular way, or fit particular intuitions.

¹⁰ See for example Gertler (2002), Scheier (2008), Papineau (1998)

Metaphysical Baggage and the Gap

Narrowing our focus now to the epistemic and semantic views discussed in the previous chapters, we will recall that for Levine, the gap emerges because of a failure of scientific theories to explain mental properties – in particular, our first person qualitative experiences, such as what it's like to see red, or taste a ripe orange. Because for Levine these phenomenal properties are too closely known to us to be denied, his goal is not to prove that a phenomenal property X exists, but to expand our scientific understanding to encapsulate properties of type X. His optimism in the future refinement of our scientific knowledge, coupled with his physicalist tendencies explain why he thinks that, in order to resolve the gap, we ought to employ a methodology similar to that currently employed by neuroscientists. The expectation is that with the right amount and type of searching in the brain, something will turn up that explains X. And not just explains X, but makes the occurrence of X predictable and deductible from within the accepted theory.

D'Oro's semantic approach, by contrast, shifts the gap from a problem of explaining the relation between mental properties and physical properties to a question about the semantic relationship between theories of mind and theories of body. For example, D'Oro might ask not where the mind is to be found in our grey matter, but what the relation is between folk psychology and neuroscience, or between ethics and descriptive history. For D'Oro, in order for us to study the gap, we must first do away with the mistaken assumption that mind and body are related in such a way to allow for empirical study of the mind. It is only when we mistake the metaphysical conditions that create the explanatory gap that we would approach it as a scientific question of interaction between mental and physical properties. Thus, it is not simply that D'Oro has no faith in science to

advance to a level of detail that can capture phenomenal properties of experience, but that she thinks empirical science is altogether unsuited to investigate the mind.

Levine's and D'Oro's beliefs in different metaphysical systems create an obvious difficulty for those who might attempt to synthesize their views. These fundamental differences in their theories create a situation where they agree that there is a gap, and then talk past each other because they disagree about the nature of the gap itself. While they both agree that there is a gap to be studied, for Levine the required research will take place in the lab and for D'Oro we can engage in such research from the comfort of our armchairs. Because it changes the focus of the explanatory gap to statements and theories, D'Oro's project requires a method more suited to semantic analysis. Certainly if mental properties are semantic attributes that cannot be discovered empirically, we are wasting our time poking grey matter and expecting to find resolution. Instead, conceptual analysis seems more apt to uncover why we speak of reasons causing actions, and ethics motivating behaviour. But what about those pesky phenomenological features with which Levine claims we are so closely acquainted?

In setting out the semantic approach, D'Oro does not directly address the problem of explaining phenomenological properties. She instead looks to mental theories and the descriptions and meanings behind things mental, and in that sense does not offer a direct reply to the problems highlighted by Levine's discussion of an epistemic gap. For Levine, qualia are the primary reason that functionalist theories don't explain the mind, yet D'Oro hardly discusses them. In Collingwood, we find two hints as to why D'Oro might not find it necessary to address phenomenological experiences, if these can be understood as 'feelings'. Collingwood believed that thought is not spatio-temporally located, and as

such, that it could not be the subject of empirical study. Because mental events such as thoughts are individuated qualitatively rather than by their spatio-temporal properties, to attempt to study thoughts or beliefs empirically by searching in the head, for D'Oro, would be absurd. Feelings, on the other hand, are spatio-temporally located according to Collingwood. The correlates of feeling may be discoverable in the body, but, as D'Oro and Collingwood would have it, just are not of particular interest to philosophers. Recall that for the conceptual idealist, philosophical concepts are just those things to which correspond no empirical distinctions, and what empirical facts there are have little to no bearing on our willingness to retain or abandon philosophical distinctions.

This may seem less controversial if we consider that the same principle of empirical sameness with conceptual distinction is employed in the popular concept of phenomenological zombies. This intuition pump, canonized by David Chalmers, in *The Conscious Mind*, is what is meant to inspire our belief in non-physical qualitative mental properties. The thought experiment simply requires that we imagine an exact physical duplicate of ourselves, who despite acting and seeming in all ways just like ourselves, has no phenomenological experience going on behind the scenes. It is our ability to entertain the conceptual possibility of such a being that is meant to show that, while there would be no empirical property distinction between a normal person P and a phenomenological zombie Z, there are other distinctions we are willing to make in order to distinguish P from Z. We are led to believe from this thought experiment that the mind is something above and beyond the body. For Levine, zombies are the principal manifestation of the explanatory gap (PH 79).

But this is precisely where Levine and D'Oro's views diverge. For Levine, the goal of a theory of mind is to eventually discover an empirical distinction between types P and Z, perhaps in a completed theory of physics. But for the conceptual idealist, to hold out for an empirical distinction would be to mistakenly assume that the missing properties are of the same ontological category as physical properties, i.e. classified empirically. For D'Oro, there are additional properties in question, but they are not additional empirical properties that obtain of P. Instead, they are semantic properties attributed to P by whoever entertains the thought experiment. Of course it is not easy for all parties to agree to D'Oros further point that the distinction that does obtain (assuming one does) between P and Z is most importantly a semantic one. Sure, Levine might agree that there are semantic differences in play, but he does not think that a 'mere' semantic distinction is what accounts for the important difference between mental and physical properties. What Levine would likely say is that while there may be conceptual or semantic distinctions in play between P and Z, it is the metaphysical distinctions that really matter in determining their differences, not the semantic facts that we use to justify and explain them.

The idea that there is non-physical mental stuff seems to play as strong a role now for Levine as it did for Descartes in the 17th century. Levine may not be arguing for the same version of substance dualism, but he still expresses a resistance to physicalism based on the failure of scientific explanation to explain mental stuff, such as what it's like and why it's like *this* and not *that*. Someone like Chalmers, for example, might even agree that mental stuff only makes an empirical/sensible difference to the subject, but from a third person perspective, is ineffable. The conceptual possibility of phenomenological zombies is supposed to convince us that there isn't an empirical difference even when we suppose

there to be a phenomenological difference, so the existence of mental stuff does make a difference to him, just not one that can be determined by anyone but the subject of experience. This is troublesome because it presumes that a subject has special access to their own (otherwise undetectable) mental states via some process of introspection or reflection, without providing any basis for believing that mental states exist, or how they could exist such that they can only be seen 'from the inside'. For both Levine and Chalmers, we need mental stuff to account for our beliefs that our mental states play a causal role in our lives, and that there is something that explains our phenomenological lives that isn't merely grey brain matter (i.e. giving our actions a rational basis). So for these two thinkers the ontological status of mind must be such that it can fulfill these two roles: being only determinable by the subject of the experience (i.e. isn't merely a rearrangement of brain matter) and being causally efficacious.

Levine's project is motivated by his belief that there is something that falls outside the scope of science which requires reduction¹¹, and D'Oro depends on the controversial claim that the mind-body distinction is most importantly semantic. In the case of the former, the explanatory gap emerges because of a feature of the subject of experiences, whereas for the latter the gap emerges because of the a priori concepts with which we engage with the subject. Of course, both authors have considerable work to do to defend their metaphysical baggage that could itself constitute a separate study, but we can at least determine that since they do not agree on what causes the gap, they are unlikely to agree on whether it can be overcome, what the gap really tells us about philosophy of mind or how to best investigate it. For D'Oro, the a priori is central to the mind body

¹¹ As shown in earlier chapters, his claims begin by arguing for an epistemic something, but toward the end of Purple Haze, his defence of the epistemic 'something' is burdened by the heavy metaphysical baggage of an indescribable metaphysical something.

debate, because borrowing from Collingwood, she thinks that the distinctions we make when looking at the world are based on the assumptions that we bring to the table¹². So, in the case of the mind-body debate, she thinks that when all the empirical facts are in, we will have still said nothing that is of interest to a philosophical theory of mind. As philosophers, we may consider what empirical features might arise in an advanced physical theory that would satisfy our intuitions about the special nature of mental properties, or whether empirical study is altogether incapable of distinguishing mental events. Levine thinks that science will eventually find success in explaining mental properties, and that the goal of philosophy of mind is closely tied to the ability of scientific theories to predict mental events and explain the relationship that holds between the phenomenological and physical aspects of experience. Throughout *Purple Haze*, Levine outlines what his expectations are for a reductive theory of consciousness. Roughly, he thinks a scientific theory of the mind ought to be able to imply a theory of the mind deductively and make mental events predictable, which is not something that even a scientific theory clearly can do for physical ‘stuff’. This requirement makes it all the less likely that we’ll ever succeed in closing the ‘explanatory gap’ that Levine labelled in 1983, despite the theoretical possibility. Because the gap is primarily epistemological for Levine, it is expected that an explanation of ‘how the physical gives rise to the mental’ would eliminate the intuition of difference and make inconceivable the possibility that the mind not be identical to the body. But of course we need to be more precise about the nature of the properties we’re dealing with¹³. If we take Levine on his

¹² The reader is welcomed to review D’Oro’s works on Collingwood, and Collingwood’s original texts for a full discussion on the role of the a priori in conceptual idealism

¹³ I do think that there is an interesting area of study in investigating metaphors like ‘gives rise’ that make it unclear exactly what is at issue in a theory of mind. It also bears consideration that the development of

own terms and expect an explanation suited to the deductive nomological model, it is hard to imagine science ever getting to the point of being able to provide us one.

Explanation in a Theory of Mind

Both Levine and D'Oro agree on the existence of an explanatory gap, but their understanding of the role of explanation and justification in a theory of mind is as different as their metaphysical commitments. For Levine, the explanatory gap exists because we are missing an explanation of how and why our conscious, subjective experience should be like *this* and not like *that*. Why is our seeing an orange accompanied by *this* sensation, and not by the one which accompanies the experience of seeing a red apple? How is it that an orange produces such a vivid phenomenological experience in us at all? As we saw in chapter one, for Levine a successful theory of mind will need to explain why our conscious experience feels the way it does, and why it should be conceivable that even given all the physical facts, there is something more to explain. But for D'Oro, the gap is not a problem with explaining *this* or *that* particular subjective mental event; the gap is a question about the relationship between our theories of mind and body, and it emerges because mental concepts are not susceptible to being explained with the concepts of the physical sciences. Recall from chapter two, that on the semantic view, the explanatory gap is primarily a question of why mind-body identity statements are different from scientific identity statements, and in particular, why the mind cannot be reduced to body in the same way as other scientific reductions. Given

quantum physics in the past few decades has largely devastated the hope of deductive and predictable theories of particles, let alone mind

these two different theories about the nature of the gap, it is no wonder that in each theory the role of explanation in a theory of mind differs so greatly.

One of the most obvious differences between the role of explanation in Levine's epistemic approach and D'Oro's semantic approach is due to their underlying metaphysics. For Levine, materialism requires the possibility of explanatory reduction from the macrophysical facts to their microphysical components. He makes the point when he says that "if materialism is true we have reason to expect that any phenomenon can be explained by reference to the physical laws and principles that govern nature as a whole...what we should expect is bottom-up necessity, both metaphysically and epistemologically" (PH 76). In the context of the gap, this means that even complex concepts such as subjectivity and phenomenal experience, should at least in principle be reducible to basic physical facts¹⁴. The concept of qualia is meant to show that materialism currently fails at explanatory reduction, though Levine thinks that we ought not give up hope that a satisfactory realisation theory will be provided by a more advanced future science, even if at present "there are good reasons for thinking that, unlike other macro domains, when it comes to qualia, we are not lacking merely enough detail to provide the requisite explanation, but any idea of how such a theory might go" (PH 69).

For D'Oro, explanation, like ontology –is– relative to the aims and method of our investigation and for this reason she thinks that mental to physical explanatory reduction is impossible. D'Oro thinks that mind and body are the products of looking at the world with a set of background beliefs that produce objects suited to our methodologies and

¹⁴ Neil Campbell, in "Why We Should Lower Our Expectations about the Explanatory Gap" suggests that the explanatory gap makes excessive demands on physical theory, and should not be of concern to physicalists.

aims. For example, a physical explanation may be available to us when we describe an action, but it is only relevant to the explanation if we are seeking a physical cause. More than likely, when we describe human actions as philosophers, the types of explanations we seek are those that draw on beliefs, rationality and justifiable behaviour. In the semantic view, there is no such thing as explanation that succeeds independently of our preconceived notions, background beliefs and investigative aims. Thus, a microphysical explanation of the mind does not fail to explain the mind because our scientific theories are not far enough advanced to explain phenomenal experience, but because it is conceptually and ontologically impossible to reduce mind to body. If mind is that which exists when we engage the world with the presupposition that events are intentionally related, and body is that which exists when we investigate the world with the goals of scientific inquiry, the difference between them is one of ontological *type*, not one of level of detail. D'Oro accepts that different types of explanation exist for understanding human experience, but does not see these as *levels* of descriptions which describe the same entity with varying degrees of specificity. One of D'Oro's primary arguments is precisely that because mind and matter exist as two distinct ways of looking at the same empirical thing, explanatory reduction is impossible because a concept employed by a theory of mind cannot retain its semantic content in a purely physical theory.

In contrast, Levine begins his investigation with the presumption that a successful theory of mind is one that can be explanatorily grounded in the concepts of science. So whereas D'Oro's and Collingwood's conceptual idealism would say that a theory of mind looks at the human experience as a series of rationally connected events that require rational explanations, Levine's physicalist ontology relates mind and body hierarchically and

presumes they are related in a way suited to being both ontologically and epistemically reduced. For Levine it is the detail of our theories rather than a mistake in our methodology that prevents us from understanding the mind-body connection. That is, we would expect to carry out a successful mind-body reduction using the same methods that have been successful in other theoretical reductions. For example, the identity statement 'heat = molecular motion' is successful because it tells us how the causal functions we associate with the concept are performed, and makes intelligible how the microphysical properties play the causal roles we associate with the macrophysical phenomenon. So, from the relevant micro-level facts about molecular motion, we are able to deduce all of the causal properties of the macro-level phenomenon we call heat. But the case of qualia is pre-emptively excluded from success using this method because for Levine, "our concepts of qualitative character do not represent... causal roles" (LO 134). So, to the extent that our a priori concept of qualia is not exhausted by causal function or other superficial properties, its theoretical reduction will fail to be fully explanatory. In fact, even if the causal or functional roles of mental states were identified with certain microphysical properties, Levine would still argue for an explanatory gap on the basis that the conceivability argument shows that things could have been otherwise (Conceivability, Identity, and the Explanatory Gap [CIE]). So long as any reasonable demand for explanation remains when given an identity claim, the explanatory gap remains.

A similar conclusion follows from a comparison of mind-brain identity statements to another identity statement that Levine thinks is fully explanatory: 'water = H²O'. This identity is fully explanatory because it explains how the meanings of the terms on either

side of the identity sign are bridged and provides us with a deductive path from the facts about water to its behaviour, such as its boiling point at sea level or its tendency to freeze at zero degrees centigrade. Each bridge principle in the reduction acts as “a premise that identifies the phenomenon to be explained with some phenomenon describable in the relevant micro-vocabulary” because their content is analytic and the identity must follow based on the meaning of the terms (CIE). Regardless of the level of detail that accompanies a similar reduction of mental states to brain states via bridge principles, for Levine there still remains a viable question of why the identity holds or how a certain qualitative state should accompany a certain microphysical state because for Levine it is neither analytically true nor conceptually necessary that brain state B should accompany, cause, or be caused by, qualitative state A.

Neither Levine nor D’Oro think that brain states conceptually entail mental states, but only D’Oro fully acknowledges the important role of the a priori in this failed attempt to explain mental states according to their underlying brain states. If, as Levine admits, it is conceptual entailment that is required to explain mental states qua their micro-physical components, semantics come to the forefront of successful explanatory reduction. Since there are no empirical distinctions at play in identity claims (since to be numerically identical is to be empirically the same) perhaps what we are really discussing are semantic distinctions. Levine would deny this of course, since to him a mere semantic distinction can’t account for such a substantive property distinction between soggy grey matter and vivid qualitative experience, but D’Oro’s argument does carry some force. For, to suggest that there is a reasonable demand for explanation left over when all the physical facts are in is just to say that there is something in the meaning of the to-be-

reduced term that is not provided by the reduced concept. This crucially relies on our a priori concepts, as even Levine will admit:

What normally justifie[s] an identification of a common property/substance... with a scientifically discovered property is the use to which the identity statement can be put in constructing explanatory arguments whose conclusions express the commonly held beliefs about it. So an identification which contradicted every single one of these commonly held beliefs would be hard pressed to find any justification. (PH 63)

If every belief we held about qualitative state A were contradicted by its reduction to brain state B, we would be unjustified in accepting the reduction. Gappy identities will always result if an identity is not an analytic truth or a direct conceptual entailment, according to the epistemic view.

The force of the semantic view is strong when it comes to an analysis of why explanatory reductions fail, but is limited when it comes to the utilization of neuroscientific data. For example, we seem to be making some progress toward understanding the mind with recent advances in neurophysiology and cognitive science. Neuroscientists have identified the Clinton gene, which accompanies any experience of the image of Bill Clinton, and the existence of mirror neurons which accompany the experience of empathy. We also have fairly good reason to be confident that we can explain why we see yellow and not blue when we look at a lemon, and that there is a particular region or cell in the brain that coincides with every possible experience. Furthermore, we can control these phenomena by stimulating brain regions to produce predictable events. What Levine's explanatory gap is meant to show is that despite the prodding around in our grey matter, the subjective experience itself seems to defy capture in terms suited to describe grey matter. And this is not just a failure of our linguistic acumen, but the inability for the scientist to understand the necessary connection that is believed to hold

between the soggy grey matter and my experience of it. D'Oro's novel suggestion is that if we cannot even determine how we might use our current tools to answer a problem, perhaps we ought to reconsider our choice of tools, and ask whether they are suited to the problem at hand.

Despite that we may have identified functional regions in the brain that correspond to predictable mental events, D'Oro asks whether we really consider this type of information to give a philosophically satisfactory explanation of mind. On the one hand she agrees with Levine that the description of a subject qua person cannot be derived from a microphysical description, but on the other hand, D'Oro's suggestion that scientific study is not suited to describe the mind severely limits her ability to incorporate scientific knowledge to the philosophical study of persons. So long as there exists a sense of 'cause' that is used to describe a rational connection between two events, according to some teleological system in which there is held to be a meaningful connection between the action, the actor and the outcome, any description that utilizes 'cause' in the sense of physical causation will not describe the mind. D'Oro might say that advances in science are adding more to the scientific explanation of physical causes and effects in the brain. These are not explanations of mind unless we are willing to suspend the metaphysical assumptions that uphold the scientific enterprise (e.g. that entities are connected causally rather than rationally, etc.), even if only temporarily, to investigate the subject matter from a different framework. It is only once we look at the events from the desired outcome of a rational explanation that they become explanations of mind.

Mental Causation in a Theory of Mind

D'Oro's and Levine's divergent views on the efficacy of the mental are the natural extension of their incompatible metaphysical views. One of the primary challenges that contemporary philosophers of mind face is to reconcile the mind with our physical theories, while at the same time being sure to retain the autonomy of mental properties, not overdetermine effects by their causes, and at all costs, avoid epiphenomenalism. But as we saw above, these are problems posed from the perspective of physicalism, and D'Oro's project forces us to consider whether a reconciliation of mind with body in a way that secures mental causation must necessarily happen in a way consistent with the methodology of empirical science.

In chapter one, we saw that one of Levine's primary motivations for endorsing a physicalist theory was to secure a causal role for the mind. For Levine, mental properties need to be sufficiently connected to their physical bases in order to both cause, and be caused by, physical events. His presumption is that "only if mental phenomena are somehow constructible from, or constituted by, the physical phenomena that serve as the ultimate causal basis for all changes in the distribution of matter and energy does it seem possible to make sense out of mental-physical causal interaction" (PH 5). To Levine, it is not just intuitive, but "very well confirmed" that "what determines the distribution of matter and energy is exclusively determined (to the extent that there is determination) by non-mental physical forces" (PH 5). Given this backdrop, it is no wonder that Levine is committed to uncovering an empirical basis for mental properties – for without that we might as well abandon the idea that mental properties produce behaviour. Unfortunately for Levine, so long as the mental cannot be linked to the causal explanations of the

sciences in the same way that the causal role of water can be derived from its chemical structure there will remain an explanatory gap.

In chapter two we saw that, luckily for D'Oro, a theory of mind is not necessarily threatened by the inability of mental properties to fit within a physicalist theory, because her understanding of how the mental is causal is understood in a rational, not a physical sense. For D'Oro, the relationship between cause and effect in a mentalistic explanation cannot be one of spatial or temporal contiguity (CSP 354) "since spatio-temporal properties cannot be legitimately ascribed to thought" (CME). Given that D'Oro individuates thoughts and beliefs qualitatively, she is therefore not required to reduce mind to matter in order for mind to have meaningful force in the causal world. That is, because she rejects that spatio-temporal properties can be applied to thought, she ensures that thoughts are ontologically incapable of fulfilling the concept of cause employed by the physical sciences. Rather than prioritise the sense of 'cause' that means temporal antecedent with sufficient physical connection, for D'Oro, to identify a mental cause means to provide an event explanation in which the agent is given a rational motive for their behaviours. That is not to say that beliefs are not causes, but that physical explanation is only one of a number of ways to describe an action. We can look for empirical connections, in which case we use the causal descriptions of natural sciences, or instead look for non-empirical connections, and employ explanations that appeal to the logical grounds rather than the antecedent states of an event. For D'Oro and Collingwood, the contemporary problem of mental causation only has force if we presume that mental causation needs to be explained from within a physicalist metaphysics.

D'Oro does not think that the immateriality of the mind requires us to give up mental causation, and Levine at one point almost seems to be in agreement. Levine acknowledges that there are different ways to secure a causal role for the mind that are not necessarily tied to being the basic physical constituents of a materialist theory:

First, we have to be satisfied with perhaps a lesser grade of causal efficacy than we might want... If materialism is true, then all causal efficacy is constituted ultimately by the basic physical properties... So if by "causal efficacy" one means the kind of role that, according to materialism, only basic physical properties can play... then of course it will turn out that mental properties, along with all other non-basic physical properties, are not causally efficacious. But so long as we also recognize another sense of "causal efficacy," a sense that applies not only by virtue of being the ultimate ground of all causal transactions, then there will be a sense in which mental properties are causally efficacious. (PH 28)

In this case, what grounds causal claims is that they support counterfactuals, and are confirmed by their instances. And if what we really care about when making causal claims is control and explanation, then it seems that rational explanations are in fact, causal explanations. Despite this, Levine is careful to separate the justification of psychological explanation from the metaphysical facts that make them true. He thinks that while it's fine to describe thoughts and beliefs as causes, they are not real causes in the same sense that heat causes a thermometer to rise or my foot on the gas pedal causes a car to move. For Levine "what warrants psychological explanations, and what makes them true are two separate issues. What we mean by calling mental states 'causally explanatory' is independent of the nature of the mechanisms by which their causal work is accomplished" (PH 33). Sure, we might call beliefs 'causes' in a psychologically justifiable sort of way, but beliefs are only truly causal insofar as they interact with the physical stuff. To say that beliefs do any causal work as psychological posits, would be to

over-determine their effects, since for Levine, it is the physical stuff doing the real causal work.

As a committed physicalist, Levine insists that all of our epistemological questions about mental causation are secondary to the metaphysical fact that only physical properties causally affect the world. And while he argues that the justification of causal claims is a separate issue from the metaphysical facts that ground causal claims, he does acknowledge that to claim that a mental property is causally efficacious without any explanation of how the property is physically realized is problematic (PH 183). If a physicalist insists that mental properties are causally efficacious then the onus will be on them to show how mental properties can affect physical properties. Even if this challenge were met with success by Levine, there still remains a daunting task to explain subjectivity and qualia. This is why Levine thinks functionalism has failed to explain the mind: sure, it can explain or recreate causal functions of the mind, but it leaves out the detail required to make sense of the hard problems of subjectivity and phenomenological experience that accompany human consciousness. As the reader will recall from Chapter one, for Levine, while being causal is an important feature of mental properties, the mind is not exhausted by its causal role in the production of behaviour.

Perhaps one of the most significant roadblocks that D'Oro faces in the mental causation debate is to persuade physicalism-loving philosophers that ontology is only secondary to epistemology. In the context of the challenge of mental causation, this would mean abandoning the empirical question 'how is mental causation possible' in favour of the epistemological question of what logical commitments our theories of mind require. Because for D'Oro, mental properties such as beliefs are not ontological candidates for

physical interaction, and exist independently of the a posteriori discovery of physical facts about the brain, “the so-called problem of mind-body interaction is a pseudo-problem that arises out of a failure to understand the aforementioned relationship” (CME). What is more philosophically interesting to D’Oro, is what we must presume about subjects in order to explain their actions as the consequences of their thought processes. Moreover, if, as even Levine admits, mentalistic explanations are able to support counterfactuals and provide a suitable level of control over events (independently of how they are grounded metaphysically), why presume beliefs must be further reduced to empirical causes?

For D’Oro it makes little sense to require that belief explanations be compatible with the methodology and ontology of the physical sciences. If we begin our inquiry from within the conceptual domain of a particular science, we should be careful to proceed using the tools suited to that discipline and not change horses midstream. For Levine, it is entirely appropriate to apply the physical model of causation to mental properties because he thinks that anything else would provide a “lesser degree” of causation that ultimately must be grounded in physical facts anyway. It is here that Levine might benefit from a recall of Collingwood’s lesson on the phases of the word ‘cause’: that the different senses of the word ‘cause’ do not themselves have meaning outside of the semantic presumptions and methodological tools of particular sciences. This is why D’Oro thinks that it is wrong to apply the language of the physical sciences to the mental causation debate; because the language of the physical sciences is only suited to describe and explain subjects as physical objects and not as rational beings. For the conceptual idealist, the mental is causally efficacious, just not in the sense used by the physical sciences.

In “Collingwood’s ‘Solution’ to the Problem of Mind-Body Dualism” D’Oro reminds us that the distinction between rational and causal explanations (or more accurately, between historical and naturalistic explanations) does not track the distinction between explanatory (i.e. epistemological) and ontological (i.e. metaphysical) questions. She states that “the distinction between reasons and causes, when properly understood, is a distinction between two kinds of explanation, rational and causal explanations, rather than a distinction between explanatory and ontological or metaphysical questions” (CSP 360). So whereas Levine states that it is not perverse to apply the “physical model of causation” to mental events if what this means is that we assume mental events have spatio-temporal location and “make a difference” to the distribution of mass-energy in space-time (PH 38), D’Oro’s point is that mental properties do not need to be physically individuated in order to make a difference. There is no requirement to show that qualia interact with physical properties because for D’Oro, feelings and sensations do not play a predominant role in a theory of mind. For philosophers, thoughts and beliefs are what are enacted in rational explanations. In fact, as discussed earlier, feelings, which have spatio-temporal location and can be studied empirically, are not of much interest to philosophers. For D’Oro, both rational and physical causes exist, and there is an ontological difference between them, but the goal is in no way to attempt to reduce one to the other.

For philosophers it will be disappointing to learn from D’Oro that the mind-body connection is not something special, but similar to other intensional distinctions that utilise multiple systems to understand a single extensionally equivalent thing. I demonstrated in chapter two that it was non-problematic for both D’Oro and Levine to

accept extensional equivalence between zombies and ‘real’ feeling people, and this same principle applied to mental causation debate implies that the purported problem of explaining mental-physical interaction is not really a problem at all. Since there are not two things, mind and body, whose causal relations with one another need to be explained, but instead one thing, which is described according to two distinct systems of causation, it is really no problem to explain how the mind can be at once causally efficacious while still autonomous from physical causes. One system describes the connection between the movements of the body’s physical parts, while another describes the logical connection between its states when viewed as rational components. This gives D’Oro the significant theoretical advantage that she is able to explain the autonomy of the mental in a physical world without resorting to substance dualism or deflationist approaches that don’t account for our intuition that beliefs really do cause behaviour¹⁵.

Despite being able to simultaneously explain the mind as causal yet autonomous, D’Oro does face the dilemma of explaining what connection, if any, does hold between physical

¹⁵ On first reflection, Collingwood’s conceptual idealism may appear similar to the well-known attempt by Donald Davidson’s to give a non-reductive explanation of how the mind can play a causal role in a physical world. Davidson’s anomalous monism, like conceptual idealism, does attempt to fit mind and body into a monistic metaphysics, and distinguish the sciences of the mind from the sciences of the body by a normative element. But the views and their motivations are quite different, as conceptual idealism does not begin with the presupposition that everything is ultimately physical, and does not endorse a layered view of reality in which the mental supervenes on an ultimately physical ontology. Both anomalous monism and conceptual idealism share the view that there is a normative element to action descriptions that are absent from simple events, and both Davidson and D’Oro agree that one cannot reduce normative concepts to descriptive ones, as something will be lost in the reduction. But D’Oro’s description of conceptual idealism on all other fronts is drastically opposed to the core tenets of Davidson’s view. For example, take the three key premises on which anomalous is based: that there is causal interaction between mental and physical events, that events related by cause and effect are covered by strict laws, and that mental events are not covered by strict laws. Conceptual idealism disagrees with each of these. Conceptual Idealism also disagrees that rationality is constitutive of mental events (it is not a degree of rationality that is required to be considered rational, since there are agents who could be said not to exhibit a degree of rationality exceeding a doorknob’s, but it is rather our a priori willingness to attribute rationality, to any degree, that constitutes mind). Those interested are directed to read D’Oro’s thorough comparison of conceptual idealism and anomalous monism in “Idealism and the Philosophy of Mind” wherein she claims that conceptual idealism is better equipped to handle the challenge of epiphenomenalism, and is truer to Kant’s transcendental idealism, from which Davidson borrows methodological notes.

and rational explanation. On the one hand, she has the advantage that it is not a problem for her theory that the mind might not fit within a physical theory, while on the other hand there is a challenge to reconcile these two positions –the first, a normative description that appeals to standards of rationality and logical thought, and secondly, a quantitative and impersonal description of events. Despite this monumental challenge that deserves its own independent study, D’Oro has offered a fresh perspective for philosophers to consider in their attempts to explain mental causation. Rather than presume that the debate must be resolved using the tools of physical science, D’Oro asks us to consider what it is that grounds causal claims, and how significantly our metaphysical presuppositions influence our views on mental causation. What for many physicalist inclined philosophers may have seemed like a straightforward project of locating the physical bases of thoughts and beliefs could, if D’Oro is correct, turn out to be a futile project that will not advance our philosophical understanding of the mind at all.

Concluding Remarks

Though we remain just as far as ever from a solution to the contemporary mind-body problem, the reader will by now have gained a greater understanding of why the problem persists. As demonstrated above, while philosophers are happy to agree that there is a gap to be accounted for in philosophy of mind, there is little consensus with regard to what the source of the problem is or how we ought to resolve it. With an aim to better understand the perennial recurrence of the gap problem, in this chapter I first looked at how Levine’s and D’Oro’s metaphysical commitments provided the key indication of how they would approach the gap problem. I demonstrated that a physicalist, such as

Levine, is committed to the methodology of the empirical sciences and therefore must treat the gap as a problem with an empirical solution. This proved problematic when it came time to reconcile his preferred methodology with his claims that mental states such as qualia are ultimately not empirically verifiable and in fact, are only observable by the introspection of the subject of experience. On top of this, we learned that for Levine, if materialism is true (which he thinks it is) mental states need to be physical in order to bear causally on physical states, or else the mind plays no role in the production of behaviour and belief.

One of the most important consequences of D'Oro's shift toward studying theories rather than properties is that it shifts Levine's project back to its original intentions. Recall that Levine sets out to correct Kripke's mistake of using epistemological criteria to draw metaphysical conclusions, and instead presents the explanatory gap as a weaker claim about the underlying cause of the distinction between scientific and theoretical reductions. But Levine's defence of an explanatory gap consistently falls back on the presumption that some property exists that is unaccounted for in our current empirical theories. So, where Kripke was criticized for drawing metaphysical conclusions from his purely epistemic evidence, Levine can be critiqued for failing to acknowledge the metaphysical presuppositions that underlie his defence of the explanatory gap.

Ultimately, these internal conceptual tensions mean that Levine is unlikely to secure a physicalist solution to the explanatory gap, at least as long as he retains his demand for a realization theory.

By putting the mind-body debate into the context of a distinction without difference, D'Oro's semantic approach has the advantage of being able to explain how the mind can

have causal bearing without physical contact, and why the tools of empirical science will never be able to reduce mental states to their physical components. Recall that for D'Oro, if some version of conceptual dualism is true, and mind and body are explanatory categories used to describe the same empirical state of affairs, of course it will be impossible to explanatorily reduce mind to body –even if they are, empirically, one and the same thing. I utilised the thought experiment of phenomenological zombies to show that we cannot discover empirically that which makes no empirical difference. If mind is a non-empirical ontological category, it will also be impossible to explanatorily reduce mind to body because the a priori distinctions, concepts and language employed in a description of mind are not suited to describe objects such as bodies, which are of a different ontological type. If we cannot foresee how to resolve the gap even with all the physical information available to us, D'Oro is likely right to suggest that we need to reconsider our strategy. In the final section that follows I will explore some of the broader implications this shift in strategy might have on the explanatory gap debate.

Chapter 4: What the Semantic Approach Means to the Future of the Explanatory Gap

In previous chapters I introduced the explanatory gap as an epistemological spin on a problem identified by Kripke. The problem was how to explain the intuitive difference between reductions in the physical domain and reductions involving mental properties. Kripke resolved the dilemma with the claim that the contingency of psycho-physical identity claims was underpinned by an actual difference in properties, since for an identity that involves rigid designators to be conceivably false, it must be actually false. Levine critiqued Kripke's claim that our beliefs track the way things really are, and attempted to revise the problem as merely an explanatory failure of physicalism to provide a realization theory of how mental properties, such as qualia and subjectivity, are connected to their physical bases. Despite having no idea how this resolution might play out, Levine did not question the truth of physicalism (since he believed it to be required to secure mental causation), but instead targeted its explanatory shortcomings, and professed hope that a more detailed future physical theory would one day resolve the gap. If Levine is correct, it would seem that philosophers are beholden to advances in science to advance the explanatory gap debate, if not the mind-body problem entirely. But, Levine's expected solution seems improbable to achieve when we take a closer look at how he established the problem. He said that the conceivability of phenomenological zombies is a prime example of the explanatory gap, yet the force of the zombie thought experiment comes from the fact that there is supposed to be no empirical difference between physical states that are, or are not, accompanied by phenomenology. Furthermore, we learned that qualia are only determinable by the subject of experience, which gives us little reason to expect that third-person empirical study will ever discover

them -since it would certainly seem odd to discover empirically that which makes no empirical difference. Needless to say, based on Levine's explanation, we have plenty of motivation to question whether the explanatory gap is really an empirical problem. Given that Levine defines qualia as empirically undetectable and determinable only by the subject of experience, D'Oro has an easy avenue from which to critique his commitment to them. But instead of shifting the onus to Levine to defend the existence of qualia, D'Oro hardly mentions qualia at all. Rather, she challenges Levine's unquestioned commitment to physicalism, and with it, his claim that the explanatory gap is an empirical problem that requires an empirical solution. D'Oro's semantic thesis proposes that the explanatory gap is not an empirical problem at all, but rather a semantic problem that arises because of a distinction in meaning applied to empirically indistinguishable events. According to D'Oro, contemporary philosophers have largely failed to recognize that physicalism is only one of a number of ontological systems that can explain mental causation. What D'Oro demonstrates is that there are multiple senses of the word 'cause', each of which describes real causes, and each of which is relative to their field of study's metaphysical presuppositions and theoretical aims. D'Oro's metaphilosophical project thus explains why empirical processes are irrelevant to philosophical facts about the mind, and demarcates a distinctly philosophical task of discovering the a priori conceptual commitments we make when we describe human actions.

Some philosophers may welcome the fact that D'Oro sees an indispensable role for armchair philosophy in the mind-body debate, but her argument also entails that the explanatory gap cannot be resolved. The semantic view requires that explanations in the physical sciences operate independently of the rational explanations used in philosophy,

and consequently that there will never be a reductive link, or realisation theory between them. Their explanations are not mutually exclusive, but rather complementary, since they each offer a unique perspective on the human experience that is neither necessarily contradictory nor relevant to advances in the other fields. And while it will be the philosopher and not the scientist who will be interested to tease out the different ontological commitments that our explanations of mind and body uncover, D'Oro doesn't indicate to what extent philosophers ought to incorporate knowledge from the empirical sciences. D'Oro's insistence that qualia are of no philosophical value is likely to be a hard sell. In a future study, one might play out the wider implications of the idea, or test her claim that philosophical study does, or ought to, deal only with distinctions to which no empirical distinctions apply.

The semantic approach to the explanatory gap gets to the heart of the mind body problem. D'Oro's argument reminds us that our use of theories, background information and a priori semantic markers are not the facts themselves, but interpretations of the empirical data we encounter. Physicalism is only one of many possible ways to understand the world, and may be of use if we are interested to study feelings, but is of little value for studying rational human action. Ultimately, the effect of D'Oro's reply to Levine will depend on our willingness to approach the mind-body problem from outside the confines of physicalism. If we can agree with D'Oro that ontology is secondary to epistemology, the future of the explanatory gap becomes less about physically explaining the mind, and entirely about uncovering the conceptual and semantic distinctions that hold between physical and rational theories of mind. Fortunately, this leaves philosophers with plenty of work to do.

Bibliography

- Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press (1996).
- Collingwood, R. G. *The New Leviathan*. Oxford: Clarendon Press (1942).
- D'Oro, Giuseppina. "The Gap is Semantic, Not Epistemological." *Ratio*, 20 (2007): 168–178.
- . "Collingwood's "Solution" to the Problem of Mind-Body Dualism." *Philosophia*, 32, 1-4 (2005): 349-368.
- . "Collingwood on Philosophical Knowledge and the Enduring Nature of Philosophical Problems." *The British Journal for the History of Philosophy*. 12, 1 (2004): 93-109.
- . *Collingwood and the Metaphysics of Experience*. New York: Routledge (2002).
- Gertler, Brie. "Explanatory Reduction, Conceptual Analysis, and Conceivability Arguments about the Mind." *Noûs* 36 (2002): 22-49.
- Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly*, 32 (1982): 127–36.
- Kripke, Saul. *Naming and Necessity*. Blackwell (1980).
- Levine, Joseph. *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press (2001).
- . "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly*, 64 (1983): 354-61.
- . "On Leaving out What it's Like." In Martin Davies & Glyn W. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays*. Blackwell (1993)

---. "Conceivability, Identity, and the Explanatory Gap."

<http://cognet.mit.edu/posters/TUCSON3/Levine.html> Web. 15 Sept. 2013.

Nagel, Thomas. "What is it Like to be a Bat." *The Philosophical Review*, 83, 4 (1974):
435-450.