The London School of Economics and Political Science

AIDS Denialism in South Africa: A case study in the rationality and ethics of science policy

Katherine Furman

A thesis submitted to the Department of Philosophy, Logic and Scientific Method of the London School of Economics for the degree of Doctor of Philosophy, London, September 2016.

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

Abstract

From 28 October 1999 to 26 September 2000 Mbeki publically endorsed the position of 'denialist' AIDS scientists – a marginal group who oppose the claim that HIV causes AIDS – and used their views as the basis for a policy of not providing ARVs (antiretrovirals – the treatment that prevents HIV from replicating) via the public health system. This policy persisted until 2004, with severe consequences – best estimates indicate that it resulted in 171, 000 avoidable new infections and 343,000 deaths over the 1999 – 2002 period. I use this case to address two questions. First, is it reasonable for policy makers to consult non-mainstream scientists in the process of policy development? Second, can they be held personally morally responsible for the consequences of having done so when things go very badly wrong?

I begin by providing a motivation for why philosophers should be interested in real-world cases. Having justified the philosophical "methodology" of this thesis, I move on to describing the specific case of South African AIDS denialism in the early 2000s. I then take a chronological step back in order to assess the rationality of accepting HIV as the sole cause of AIDS in 1984, when the virus was first identified. I argue that it was rational, but that some explanatory power was lost when other competing accounts of the disease's aetiology were discarded. I argue that this explanatory loss can be accounted for by re-considering the way causation is understood in biomedicine and epidemiology. Having settled the scientific issues of the case, I then move on to the question of moral responsibility. I specifically look at when an agent can be held morally responsible for their ignorance, and the role of suppressed disagreement in the production of that ignorance.

For Richard, who always believed in this thesis.

Acknowledgments

A great many people have made this thesis possible, and I am grateful for all the love and support I have received.

My supervisors, Luc Bovens and Katie Steele, have been extremely generous with their time, and have patiently helped me through this project. I cannot fully express all the gratitude I feel toward them both.

I have been lucky to have many smart and patient friends, who have talked through the problems in this thesis with me, and have read and commented on its various component parts. Special thanks must go to the following people for all the help you have provided in bringing this document into being: Asbjørn Schmidt, Mary Carman, Joe Roussos, Laura Sochas, Stuart Theobald, Stefan Schubert, Rebecca Swartz and Gordon Inggs. You have given me many of your good ideas, and saved me from many of my bad ones. Very special thanks must go to Deirdre Morgan, who offered me support every day during the final months of my thesis – thank you, you saved me.

Thanks to the LSE and the Oppenheimer Memorial Trust for funding this work. And thanks to the LSE Mobility Scholarship for funding my research at the University of Cape Town in 2015. This was an extremely productive and creative time for me, and many of the ideas that I stumbled upon while there are in this thesis.

Thanks to my family – Roslyn Furman, Elaine McDermid, and Elaine Stupart for not giving up on me. I am especially grateful to Elaine McDermid for all the prayers she sends to St Jude (the patron saint of lost causes) on my behalf.

Most importantly, thanks to Richard Stupart. Without him this thesis would not exist. His support and cooking and proof-reading all made this work possible. And his endless sense of good cheer and humour kept me sane. I am so grateful that I get to spend my time with you.

Table of Contents

Introduction

From 28 October 1999 to 26 September 2000 the South African President, Thabo Mbeki, endorsed the position of 'denialist' AIDS scientists – a marginal group who oppose the claim that HIV causes AIDS.

Throughout this thesis the term 'AIDS denialist' will refer to anyone who adopts the position that HIV does not cause AIDS, even though they might accept both the existence of AIDS and the existence of HIV, as Peter Duesberg, the most prominent of the AIDS denialists, does (more will be said about this in Chapter 3). Individuals who believe that HIV does not cause AIDS prefer to refer to themselves as 'AIDS dissidents', rather than 'denialists' (Kalichman, 2009, p. 12). Using either term is value laden. Referring to those who accept this position as 'dissidents' casts them as valiant heroes, standing up for their beliefs against the oppressive force of mainstream science, and draws on associations with political and religious dissidence. The term 'denialist' implies that they deny obvious and well-established pieces of evidence, similar to Holocaust denialists (Kalichman, 2009, p. 11). Using either term implies the rightness or wrongness of the position from the outset. However, in order to write intelligibly, I need to call those who support the position something. I could introduce a third more neutral term, but to do so would be a departure from the current nomenclature and literature on the subject. I anticipate that as this thesis progresses, readers will be in increasing agreement that 'denialist' is the more appropriate term.

Two claims constitute Mbeki's particular brand of denialism: 1) ARVs (antiretrovirals – the treatment that prevents HIV from replicating), specifically AZT (azidothymidine), are toxic; and 2) HIV is not a necessary condition for AIDS (Cherry, 2009, p. 16). On the basis of these beliefs, Mbeki adopted a policy of not providing ARVs via the public health system. It is important that he held both beliefs (that HIV is not a necessary condition for AIDS, and that ARVs are toxic) at the same time. If he had just believed that HIV is not a necessary condition for AIDS, but had not held the belief that AZT is toxic, he may still have concluded that ARVs should be publically available, perhaps guided by something like the

Precautionary Principle. Similarly, if he had held the belief that AZT has severe (or "toxic") side effects, but had accepted that HIV causes AIDS, he may also have made ARVs available, based on the calculus that the risks associated with the treatment are outweighed by the benefits. This policy persisted until 2004, with severe consequences. Best estimates indicate that it resulted in 171,000 avoidable new infections and 343,000 deaths over the 1999–2002 period (Nattrass, 2008, p. 157).

Thabo Mbeki's support for these scientists and the resulting policies he adopted in the 1999-2002 period are puzzling. Mbeki built his leadership reputation as a rational intellectual, often to his detriment – this aspect of his public persona also made him appear aloof, and ultimately led to a sense that he was disconnected from ordinary citizens. Some argue that this was the cause of his loss of support within the ANC (African National Congress – the South African ruling party) in 2008 and his replacement by Jacob Zuma, who is taken to be the less intellectual and more charismatic leader (Lodge, 2009). In the context of Mbeki's intellectualism, his support for denialist AIDS science seems incongruous. Posel asks:

> Why did an intelligent and politically shrewd President, with no expertise or training in matters of health, weigh in so directly and heavily on the question of AIDS, risking all the national and international turbulence provoked by his interventions? (Posel, 2008, p. 18)

This incongruity has generated a range of literature on Mbeki's AIDS denialism from a wide variety of perspectives: economic (Nattrass, 2004); psychological (Kalichman, 2009); anthropological (Fassin, 2007); and political (Butler, 2005). Little has been said about the case from a philosophical perspective, with the notable exception of Ward Jones's short paper, "Dissident versus Loyalist: Which Scientists Should We Trust?" (2002), written in the midst of Mbeki's denialism, he warns against consulting non-mainstream scientists in the process of policy development. Robert Kowalenko, in a more recent paper, "Thabo Mbeki, postmodernism, and the consequences" (2015), uses the case of AIDS denialism as a counter-example to postmodern/post-colonial theories of science. Both Jones and Kowalenko focus on the case in the context of compartmentalized

issues in philosophy of science – the relationship between mainstream and marginal science (Jones), and the status of postmodern/post-colonial scientific approaches (Kowalenko). Neither provides a detailed analysis of the case. This thesis offers the first thoroughgoing analysis of AIDS denialism from a philosophical perspective.

There are good reasons, both practical and philosophical, for philosophers to pay attention this case. Practically, the harms of Mbeki's policy decision were severe, both in terms of avoidable loss of life and new infections. Further, his public statements about HIV and AIDS introduced an element of uncertainty about AIDS into the public sphere in South Africa, which continue to hinder public health education efforts to tackle the disease (Grebe & Nattrass, 2012, pp. 770-771). Due to these harms, there has been a call from members of South African civil society to assess whether Mbeki can be held morally responsible for his actions (see, for example, Geffen (2009)). Philosophical reflection on this topic is thus practically useful. AIDS denialism is also worth considering for its own sake due to the variety of philosophical questions that it raises. Providing a full assessment means addressing a wide range of philosophical topics; spanning from issues of disease causation through to epistemic considerations in attributions of moral responsibility. This case is philosophically rich.

In my analysis I consider two main questions. First, was it reasonable for Mbeki to consult non-mainstream scientists? Second, can he be held personally morally responsible for the consequences of having done so?

I begin, in Chapter 1, by providing a motivation for why philosophers should be interested in real-world cases. In particular, I argue that the thought experiments that are the central methodological tool of contemporary analytic philosophy run the risk of excluding philosophically relevant material in the streamlining process that is required for their production. Real-world cases avoid this problem, and can provide a check that this has not occurred in philosopher's streamlined hypothetical cases. Furthermore, real-world cases can help to ensure that the results of philosophical study are transferrable to novel scenarios (i.e.

they can help with external validity), and they provide a 'context of discovery' for generating new philosophical hypotheses.

Having justified the methodology employed in this thesis, I move on to the case. I provide a brief history of South African AIDS denialism in chapter 2 and then I take a chronological step back in chapter 3 to consider early AIDS science in the 1980s. Contemporary AIDS denialists claim that they were unfairly and prematurely cut out of the debate when HIV was accepted as the definitive cause of AIDS. I take their complaint seriously. If they are correct then we might be mistaken in some aspects of current AIDS science, and Mbeki might have been acting to remedy an epistemic injustice by including members of an unfairly marginalized scientific group on his advisory panel. I use Lakatos's theory of scientific research programmes to assess the rationality of accepting the viral theory of AIDS in the 1980s; concluding that the acceptance of the viral explanation *was* rational, but that corroboration came late, and that some explanatory power was lost in the move from multiple theories of the disease's aetiology to the strictly viral account. For instance, it became difficult to explain the role that socio-economic factors played in aetiology of AIDS, which in turn made it difficult to account for the disproportionate prevalence of AIDS in southern Africa.

This was the state of AIDS science when Mbeki entered the debate in the late 1990s. HIV was accepted as the definitive cause of AIDS, and there was no room for non-viral (socio-economic) factors in the account. In chapter 4 I argue that one plausible (albeit charitable) reading of the Mbeki case is that he became aware of this explanatory gap and began to consult non-mainstream AIDS scientists in an effort to remedy it. Non-mainstream AIDS scientists were able to include socio-economic factors in their account of the disease, and were thus able to explain the disproportionate prevalence of AIDS in southern Africa, however they incorrectly excluded the causal role of the virus in the process. I argue that underpinning this problem is too strict a divide in biomedicine and epidemiology between mono-causal and multi-causal theories of disease explanation: accounts of disease that favour single, typically microbial, sources of disease, and those

that permit multiple causes, but fail to distinguish the causal salience of microbial factors. I suggest that these accounts can be integrated by thinking of the issue as a causal selection problem. The task then becomes one of distinguishing the 'causes' from the 'background conditions', and Woodward's (2011) criteria of causal selection provides guidance on how to do this with some rigour.

Having a thorough grip on the scientific and political issues that cumulated in Mbeki's AIDS denialism, in chapter 5 I move on to questions of moral responsibility and blame. In particular, I assess whether Mbeki can be held personally morally responsible for the consequences of his policy decision. I explore the possibility that Mbeki might have an excuse from ignorance available to him for his harmful action – he might (hypothetically) argue that had he really known that HIV causes AIDS, and that ARVs prevent the virus from replicating, then he would have made the treatment available via the public health system. However, excuses from ignorance are only acceptable if the ignorance itself is not culpable. One can avoid culpable ignorance by satisfying one's 'procedural epistemic obligations' – that is, one is required to take care when forming beliefs that inform actions with potentially harmful outcomes. One way to take care is to pay attention when one's epistemic peers – those with similar reasoning abilities and similar access to evidence – disagree. In the Mbeki case, it is plausible that his epistemic peers were other members of his political party. However, when we look at the state of debate around the issue of AIDS in his political party in the late 1990s and early 2000s, there is surprisingly little disagreement with Mbeki's view. Commentators on the case argue that members of his party did disagree with Mbeki, but that their disagreement was suppressed because they feared him. I explore the issue of suppressed disagreement in the context of culpable ignorance assessments, concluding that ignorance resulting from suppressed disagreement typically cannot provide an excuse for wrongful action. Furthermore, I argue that if the agent who draws on ignorance as an excuse in situations of suppressed disagreement is the same agent who suppressed the disagreement, then they are additionally blameworthy for the harms associated with silencing their disagreeing interlocutors.

Ultimately, I conclude that Mbeki was not wholly unreasonable in his consultation of non-mainstream AIDS scientists. There were explanatory gaps created by the move from multiple theories of AIDS's aetiology to the strictly viral account. Too strict a divide between mono-causal and multi-causal theories of disease worsened the situation. However, his decision to follow the AIDS denialist scientists in their rejection of the virus was incorrect and had harmful consequences. Further, he silenced those who might have been able to prevent him from making the wrong decision. It will thus be concluded that his ignorance is culpable, and he is morally responsible for the tragedy of AIDS denialism.

1. What use are real-world cases for philosophers?

1.1. Introduction

This thesis focuses on the detailed study of a real-world case from a philosophical perspective. This is a relatively unusual project in contemporary analytic philosophy, where the emphasis is on abstracting away from the particularities of the world. If a purported piece of philosophical writing strays too far in the direction of the practical, its credibility as a work of philosophy is at risk (Srinivasan, 2016, p. 1).[1] To see how philosophical concern is typically focussed on abstracting away from particulars, note that Cohen argues that facts about the world are irrelevant to developing an account of justice:

> [F]acts are irrelevant in the determination of fundamental principles of justice. Facts of human nature and human society of course (1) make a difference to what justice tells us to do in specific terms; they also (2) tell us how much justice we can get; and they (3) bear on how much we should compromise with justice, but, so I believe, they make no difference to the very nature of justice itself (Cohen, 2008, p. 285).[2]

To see that if one strays too far in the direction of the practical, there is a risk that one's work will be dismissed as not philosophical, note that Srinivasan's first reaction to reading Nancy Bauer's *How to Do Things with Pornography* (2015), which deals with a range of real-world concerns related to sexuality, was: "Is this philosophy?" (Srinivasan, 2016, p. 1). Srinivasan indicates that she is hopeful that the answer to that question is "yes". Less optimistically, Jenkins reports that on telling a fellow philosopher about her work on the applications of metaphysics to gender, she was met with the response: "That's not philosophy" (Jenkins, 2014, p. 262).

[1] Srinivasan does not defend this view of philosophy. She merely notes (with some regret) that this is typically the case.
[2] I initially encountered this quote in Wilson (2014, p.19).

It is not entirely clear why this stance persists, given that criticism against this isolationist vision of philosophy has been put forward for some time. Williams, for instance, in *Morality: An Introduction to Ethics* (1972) is famously critical of approaches to philosophy that neglect the real world, claiming that although "[m]ost moral philosophy at most times has been empty and boring… contemporary moral philosophy has found an original way of being boring… by not discussing moral issues at all" (Williams, 1972, p. 9).[3] In a similar vein, he argues in the preface to *Moral Luck* (1981) that moral theorising needs to be attached to moral experience. Further, as Williams himself points out, this is a concern that substantially pre-dates his work, when he introduces an article on the role of history in philosophy with this 1878 quote from Nietzsche: "Lack of a historical sense is the hereditary defect of philosophers… So what is needed now is historical philosophising, and with the virtue of modesty" (Williams, 2002, p. 9 quoting Nietzsche). Regardless of the persistence of this issue, it is clear from the discussion above – Cohen arguing in favour of the philosophical armchair on the one end of the spectrum, and Srinivasan and Jenkins lamenting philosophical detachment from the real world on the other – that making philosophical use of material from the world still requires a defence if it is to be taken seriously. In this chapter, I aim to provide such a defence.

To start, it is fairly uncontroversial that the clarity and rigour provided by philosophical thinking can be useful for achieving better understanding in certain real-world cases (Srinivasan, 2016, p. 2; Kamm, 2009, pp. 19-20). For instance, in the sciences, philosophers often perform important clarificatory work when conceptual confusion arises (Kitcher, 2011, p. 253). Somewhat more controversially, it might be argued that philosophers have an obligation, either professionally or ethically, to spend a portion of their work-time devoted to real-world cases in order to assist with this important clarificatory work (Jones,

[3] In an interview in 1983 he claims that between 1969, when he initially wrote *Morality,* and the interview being conducted, that this problem had begun to resolve itself and that his criticism had become out of date. Instead, he argues that a new problem entered philosophy in the form of "applied philosophy", which is that philosophers had begun to develop something of a cottage industry around quickly applying pre-existing philosophical theories to practical problems, which risks turning philosophy into a boring "quasi-legal" enterprise. However, in 2002, in a piece in the *London Review of Books*, he once again laments the lack of historical concern in contemporary analytic philosophy.

2006). I will not address either of these claims in this chapter – namely, that philosophical rigour is likely to be of benefit to better understanding real-world cases, and that philosophers might be obliged to assist in this regard. In this chapter I will not attempt to produce a defence of real-world philosophical engagement as a kind of professional public service.[4]

Instead, my target in this chapter is the naysayers – those who believe that no philosophical gains can be achieved by considering the real world. Taking this as my target, I focus on defending the claim that *philosophical* benefits can be achieved by paying careful attention to real-world cases: that is, over and above the thought experiments that are already a standard part of the philosophical practice, or the more detailed fictional cases that are often advocated for when thought experiments are found to be insufficient (see, for instance, Nussbaum (1990)). My intention is not to argue that thought experiments or fictional cases should be excluded from philosophical methodology, just that real-world cases ought to be included. That is, material from the real-world should be available to philosophers as a legitimate part of their philosophical toolkits – no philosopher is required to make use of material from the real-world, but it should be available to those whose practice would benefit from doing so. In particular, I argue that if thought experiments and cases from fiction were the only kinds of cases available in philosophical methodology, there would be a gap.

The main line of argument that will be pursued is as follows. Thought experiments (or 'streamlined hypothetical cases') are effective because they allow for potential 'confounders' to be removed. In the sciences, a confounder is a factor that interferes with the relationship between the cause under study and its purported effect.[5] Typically, in the scientific context, this interference occurs when there is unaccounted for difference between the test case (the case in

[4] However, I do believe that this is a valuable way for philosophers to spend their time, and much of this thesis will be focussed in that direction – on using the tools and techniques of philosophy to provide greater clarity to a real-world case. For instance, in Chaper 3, I use Lakatos's theory of scientific research programmes to assess the transition from multiple theories of the aetiology of AIDS to the strictly viral account.

[5] There is also a use of the term 'confounder' which refers to a common cause that undelies a spurious correlation – such as the correlation between yellowed fingers and lung cancer, where smoking is the confounder. This is not the intended use here.

which the experimental intervention occurs) and the control case (the case in which there is no experimental intervention). For instance, to borrow a hypothetical example from the medical sciences, if I were studying the effects of a new headache medication, but the individual receiving the treatment was very dehydrated at the time, and dehydration has an independent effect on the rate of recovery, then this is a confounder because it interferes with one's ability to track the causal relationship between the experimental treatment and the relief from the headache.

In philosophy, a confounder is similarly taken to be an interfering factor; in particular, a confounder is anything that interferes with the relationship between one's intuition and the philosophical principle under consideration.[6] The term 'confounder', when used in the philosophical context, is much broader than how it is typically used in the sciences – anything that disrupts the pathway from intuition to principle is going to count as a confounder in philosophy, not just unaccounted for differences between tests and controls. I take 'intuition' here to mean what Kamm describes as a "judgement about a case", where that judgement is reason-driven and not merely a gut-feel emotional response (Kamm, 2009, p. 23).[7] By '*interfere* with one's intuitions' I mean any factor that distracts one's intuitions from the philosophical principle being assessed. For instance, this might occur in cases in which there is more than one morally relevant factor present, such that one's intuition gets drawn away from the particular issue that is being analysed. For example, note the role of malice in Rachels's (1978/1997) famous Bathtub Case, which is meant to provide a test for whether killing is morally worse than letting die. The case asks the respondent to imagine two scenarios, as follows:

> In the first, Smith stands to gain a large inheritance if anything should happen to his six-year old cousin. One evening while the child is taking his bath, Smith sneaks into the bathroom and drowns the child, and then arranges things so that it will look like an accident. In the second, Jones

[6] I am following Elgin (2014) and Wilson's (2016, forthcoming) use of terminology here.
[7] This use of the term 'intuition' is different to how psychologists use it. Psychologists use it to describe a response that is "automatic, quick, effortless, associative, and often emotionally charged… [and] not open to introspection" (Kahneman, 2009, p.72). The usage in this chapter tracks the standard use in moral philosophy.

> also stands to gain if anything should happen to his six-year-old cousin. Like Smith, Jones sneaks in planning to drown the child in his bath. However, just as he enters the bathroom Jones sees the child slip and hit his head, and fall face down in the water. Jones is delighted; he stands by, ready to push the child's head back under if it is necessary, but it is not necessary. With only a little thrashing about, the child drowns all by himself, "accidentally," as Jones watches and does nothing (Rachels, 1978/1997, p. 79).

The cases are intended to differ only in that the first involves killing, while the second involves letting die (Wilson, 2016, forthcoming, p. 2). Rachels takes the second case to be no worse than the first, concluding that there is no morally relevant distinction between killing and letting die. However, in both cases the protagonist has malicious intent (a morally relevant factor), and malice swamps any finer distinctions that might be drawn between the cases. If we were able to examine the distinction between killing and letting die without the influence of malicious intent, we might find one worse than the other. Rickless (2011) makes the same point (albeit in response to a different set of cases), arguing that cases such as these show that there is no distinction been maliciously causing harm and maliciously allowing (or enabling) harm, they do not show that there is no distinction between non-maliciously causing harm and non-maliciously allowing (or enabling) it (p. 71).

Excluding potential confounders is useful, but I argue that doing so comes at a risk. The risk is that in the process of streamlining out potential confounders, important 'support factors' might be unwittingly removed. A 'support factor' in the sciences is any factor that is required for the cause to achieve its effect – for instance, oxygen would be an important support factor in an experiment examining the causal relationship between matches and combustion (Cartwright & Hardie, 2012, p. 62). In philosophy, the 'support factors' are, strictly speaking, those factors that need to be present for the key factor of interest to 'do its work'. What this amounts to in the philosophical context is that support factors are those that are required in order for one's intuitions to latch on to the salient philosophical relationships: those factors that are required to keep one's

intuition on the target principle.[8] For instance, in the standard Trolley Problem (which I will discuss in more detail below), a support factor might be that the respondent needs to imagine herself some *distance* away from the person that they are considering sacrificing. Evidence suggests that once distance is taken out of the scenario, and the respondent is required to imagine themselves physically closer to the person that they might sacrifice (such as is required in the Footbridge Case[9]), they lose focus on the principle under consideration (whether it is permissible to sacrifice one to save five), and become distracted by how repulsive they find direct physical violence (Kahneman, 2009, p. 79; Singer, 2005). Distance, it turns out (if Kahneman and Singer are correct), is a support factor in the standard Trolley Problem.

A large part of the problem with support factors is that it hard to identify them just by looking at a case – if they were easily identifiable, we could just check that we hadn't factored them out in the process of streamlining and the problem would be solved. Support factors often only become visible once we have a detailed understanding of how the mechanism driving a particular effect works, or once a particular support factor has been removed and the effect ceases. For example, Cartwright and Hardie (2012) use the California Class Size Experiment (an example from the social sciences) to show that support factors may only become identifiable on the *ex post* analysis. In this case, there had been a successful educational intervention in Tennessee, in which class sizes had been reduced in order to improve school pupils' reading scores. The same policy was subsequently implemented in California and the intervention was unsuccessful. On the *ex post* analysis it became clear that in Tennessee there had been adequate space for the increased number of smaller classes and available well-trained teachers to take over the additional classes. These were both important

[8] Note, "support factors" are the inverse of "confounders". A confounder disrupts one's intuition from tracking the principle under consideration, by dragging one's intuition toward something else (like the race of the hypothetical participants, or additionally morally relevant factors), while a support factor is something that needs to be in place to keep one's intuition on track.
[9] The Footbridge Case asks the reader to imagine that there is an out-of-control trolley hurtling down the tracks, which if left un-intercepted will kill five people further down the track. You are watching the scene from a footbridge over the track. Next to you is a very fat man; so fat that he would stop the trolley if he were pushed in front of it. Do you push the fat man off the footbridge, which will kill him, in order to intercept the trolley and save the five (Thompson, 1985)? Most respondents say that it is impermissible to do so.

support factors for the success of the intervention in Tennessee and neither was present in California, which is why the policy failed there. That these were support factors only became clear once they were removed and the effect was not achieved (Cartwright & Hardie, 2012, pp. 65-66).[10]

The philosophical analogue is that we are typically unable to tell precisely what the support factors are in a particular case until we have a detailed understanding of the relationship between the intuition and the principle, or once we remove a relevant support factor and the intuition suddenly no longer holds. For instance, this is the case when distance was removed from the standard trolley problem and respondents no longer had the intuition that it was permissible to sacrifice one in order to save five.

Consulting real-world cases helps us to check that important support factors have not been inadvertently excluded in the process of streamlining out potential confounders. Given that fictional cases are extended thought experiments (a claim that I will defend later in this chapter), cases from fiction will not be able to fulfil the function of checking that support factors have not been excluded, because they run the same risk as more austere thought experiments – support factors might unknowingly be removed by mistake. Real-world cases are the only remaining place to look in order to make sure that all the relevant support factors have been included. There are two related risks to not consulting the real world: the first is that our philosophical methodology would not be as rigorous as it could be, the second is that we might miss out on important philosophical relationships. Given these risks, real-world cases fulfil an important philosophical function, as I will defend below, and should be part of philosophical methodology.

[10] However, Cartwright and Hardie (2012) argue that this mistake could have been avoided by devoting more time to uncovering the mechanism that had driven the success of the intervention in Tennessee, which would have made the support factors visible and would have allowed policy makers to ensure that they (or suitable substitutes) were present in the California case. Either way, the point is that support factors are unlikely to be immediately obvious just by looking at the case, which increases the risk that they might be accidentally removed.

In addition to the main line of argument presented in this chapter, I also discuss two further philosophical benefits that can be achieved by consulting real-world cases. These are: 1) ensuring that the findings from studying particular philosophically-relevant cases (be those standard thought experiments, fictional cases, or real-world cases) can be translated to novel scenarios (in the language of experimentation in the natural sciences, ensuring 'external validity'); and 2) consulting real-world cases is a useful way of generating philosophical hypotheses in the first place, and (relatedly) for helping to extend and adapt pre-existing philosophical concepts.

A few caveats before proceeding. First, the emphasis in this chapter is predominantly on thought experiments in moral and political philosophy, because this is an area of philosophy in which thought experiments are used to test out hypotheses and principles, while in other areas of philosophy thought experiments are often focussed on examining conceptual entailments (Wilson, 2016, forthcoming, p. 1). As a result, the arguments presented here might not be transferrable to other areas of philosophy, such as metaphysics or epistemology. Second, I focus on cases (thought experiments, fictional cases, real-world cases) in their capacity *as* experiments, by which I mean, as tests for hypotheses and principles (Wilson, 2014, p. 13). I ignore the other philosophical functions that cases might play – such as the illustrative role that they might serve at certain points in the exposition of a philosophical argument (Brown & Fehige, 2014) or the educational role that they might play in training philosophers to be sensitive to identifying that which is morally relevant (Nussbaum, 1990). These are undoubtedly useful functions for cases to play in our philosophical practice, but they will not be discussed here.

1.2. Starting with Thought Experiments

Williamson (2011) argues that one of the salient features of contemporary analytic philosophy is the importance of thought experiments in its methodology (p.215). I take thought experiments to be the major methodological rival to real-world cases, and so it is important to have an idea of what thought experiments

are and how they function before drawing comparisons. Also, given that my main line of argument in this chapter is that real-world cases fill a gap left behind by thought experiments and fictional cases, a clear description of that gap needs to be provided. Providing an account of thought experiments is the focus of this section.

Fischer describes thought experiments as "schematized hypothetical scenarios in which only a few details are filled in, and all the other details are left out" (Fischer, 1995, p. 4). Similarly, Wilson (2016, forthcoming) describes them as: "toy ethical cases that are designed to simplify an ethical problem along a number of dimensions, thus making the problem more philosophically tractable" (pp. 1-2). Relatedly, Elgin (2014) argues that this streamlining process in philosophical thought experiments should be understood as analogous to scientists' lab experiments, and she is echoing a relatively popular position in the literature (also see Wilkes (1993), Fischer (1995) and Wilson (2016, forthcoming)). Elgin describes this commonality between philosophical thought experiments and scientific lab experiments as follows:

> It is a controlled manipulation of events, designed and executed to make some particular phenomenon salient… Important properties and relations are often masked by the welter of complexities that embed them. In experimenting, a scientist isolates a phenomenon from many of the forces that typically impinge on it. To the extent possible, she eliminates confounding factors. She holds most ineliminable factors fixed, effectively consigning them to the cognitive background of things to be taken for granted. This enables the effect of the experimental intervention on the remaining variables to stand out. Through such a strategy, she casts into bold relief factors that might be typically hidden from view. (Elgin, 2014, p.222).

An illustrative example of streamlining out potential confounders in the sciences is that of a controlled trial (Elgin, 2014, p. 222 - 223). To see this, it is worth returning to the headache treatment example from the introduction. Imagine that we want to test a new headache medication and we know that various factors impact on how quickly individuals recover from headaches, independently of whether or not they receive any treatment. These factors may include things

such as age, sex, weight, whether the patient is a smoker, how many hours the patient sleeps a night, etc. In an ideal test of the treatment, we would want there to be two groups, an experimental group (the group that receives the treatment) and a control group (the group that does not receive the treatment), and the members of the two groups would be identical in terms of the things that are relevant to the effect – they would all be the same age, sex, weight, they would all be non-smokers (or smokers), and they would all sleep the same number of hours each night, etc. The only thing that should differ between the groups is whether or not they receive the treatment, and this allows for the causal relationship between the treatment and the effect to be isolated. At least, this is the case for a particular type of methodologist. A lesser requirement would be that the groups be balanced, but not necessarily homogenous, in which case the researcher would isolate the average causal effect on heterogeneous but identical groups of that sort. Either way, the point is to streamline out confounders and to focus just on the factors that are relevant for the hypothesis being tested.

Thought experiments in philosophy are meant to do something similar. The intention is to factor out potential confounders. I take the standard Trolley Problem to be a classic example of a thought experiment in moral philosophy. In this case, a runaway trolley is hurtling down the tracks, where it will kill five people. As a passer-by, you happen upon a switch, which allows you to divert the trolley down a neighbouring track, where it will only kill one person. The philosophical question is this: do you sacrifice the one in order to save the five (Foot, 1967)?[11] Importantly, obvious potential confounders have been factored out of the case. For instance, things that might trigger implicit biases have been excluded – such as the race, age and gender of the individuals on each of the tracks, etc.[12] The set of possible actions has also been significantly curtailed (Wilson, 2014, p. 14) – you can only allow the trolley to continue or you can divert it; the five cannot escape, nor can you warn the one on the neighbouring

[11] This is not quite Foot's original version of the Trolley Problem, in which the reader is asked to imagine herself in the position of the driver of the trolley, but it more clearly reflects what is taken to be the classic version of the problem, which is Thompson's (1985) version of the case.
[12] Again, this is not quite the same as Foot's original version of the problem, in which the five are male railway workers, but it does accurately reflect the current version of the standard trolley problem.

track.[13] Additionally, other potentially morally relevant factors that might distract one's intuition have been excluded; for instance, nobody involved in the case has malicious intent, and so there is no risk of malice obscuring the finer distinctions present in the case. By streamlining these extraneous factors out of the case, we are able to focus on just that which is relevant to the philosophical investigation at hand – whether it is permissible to sacrifice a few in order to save many.

For an even clearer analogue between controlled laboratory experiments and philosophical thought experiments, remember Rachels's Bathtub case from the introduction. The only thing that is meant to differ between the cases of Jones and Smith is that Jones kills the child while Smith merely allows the child to die – the intention being that the philosopher can focus exclusively on that distinction. This is analogous to the ideal test of the headache treatment, in that the only thing that is meant to differ between the test group and the control group is that the one receives the treatment and the other does not. Unfortunately, in the Rachels case the distinction is lost because of the role of malice, but the analogy with the control trial is still clear. Being able to streamline out potential confounders and focus just on that which is philosophically relevant seems to be obviously useful.

1.3. The smuggled intuition worry

One worry the reader might have at this point is that thought experiments might not actually produce 'clean' intuitions – that is, there might still be extraneous factors influencing our intuitions and distracting us from what is philosophically important. In particular, we might smuggle extraneous factors into our interpretation of cases and this might have an impact on our intuitions. There are at least two ways in which this might happen. First, in the way that we imaginatively fill in the details of sparsely described thought experiments.

[13] Anyone who has taught the trolley problems to undergraduate students will understand the importance of severely curtailing the range of available actions. Otherwise, questions like "why don't the five run away?" and "why don't you sacrifice yourself?" distract from the philosophical problem under investigation.

Second, we might unknowingly allow our intuitions about the real world to colour how we view the thought experiment. Each of these will be discussed in this section.

Wilson (2016, forthcoming) suggests that an unavoidable consequence of very austere thought experiments might be that respondents can fill in the details of the case however they choose. This creates a methodological problem in that different individuals might effectively be responding to different thought experiments, depending on how they personally fill in the details (Wilson, 2016, forthcoming, p.10). Williamson (2007) makes the same point about individuals filling in the details of very austere thought experiments when he states that:

> In philosophy, examples can almost never be described in complete detail. An extensive background must be taken for granted; it cannot all be explicitly stipulated. *Although many of the missing details are irrelevant to whatever philosophical issues are at play, not all of them are*. This applies not just to highly schematic descriptions of examples, such as the initial abstract Gettier schema, but even to much richer stories Gettier and other philosophers like to tell… Similarly, when moral philosophers assess imaginary examples, *one can almost always fill out the case with unintended but morally relevant additions that would reverse the verdict. Any humanly compiled list of such interfering factors is likely to be incomplete* (p. 185)[emphasis added].

Wilson (2016, forthcoming) uses Kamm's (2006) Reach Case to illustrate this point. In the Reach Case, the reader is asked to imagine an agent with arms that are so long that they reach from one side of India to the other, allowing the agent to rescue a child who is drowning in a pond on the other side of the country. The intuition is meant to be the same as if the child were physically close by: that is, the agent with very long arms is morally required to rescue the child drowning on the other side of the country the same as if the child were drowning in a pond right next to the agent. However, Wilson argues that there are various ways of filling out the details in this case (2016, forthcoming, p. 7). For instance, are we meant to imagine a normal human-sized agent with disproportionately long arms, or is the whole person supposed to be scaled up to the size of a giant, so that the long arms are in proportion to the rest of their body? When I asked a friend to imagine this case he told me that he pictured Mr Tickle (a cartoon

character, composed of a round orange body and very long arms), which gives an indication of how vast the interpretative range might be.

Wilson notes that the fact that there are various ways that one might fill in the details of the case is likely to be a problem, but he does not fully articulate what that problem is. It seems to me that the problem is that our intuitions about what is morally required of the agent in the case are likely to vary depending on how we imagine it. The principles that we come up with are more specific (i.e. pertinent to the particular type of scenario) than we initially think. If we imagine a normal-sized person with disproportionately long arms, then it is unlikely that our intuitions would be the same as for the case of a child drowning in a nearby pond. This is because the agent's head (along with their eyes and ears) is still on the other side of the country. The concern is that the agent might not know about the child drowning in the pond on the other side of India, and that even if they did know, they might not have enough information to intervene effectively, because they lack relevant sense data that they would have if the child were right next to them. Having arms that are long enough to rescue the child drowning on the other side of the country is not enough for us to agree that the agent has the same moral duties toward that child as the agent would have to the child drowning right next to them. We might have to imagine that the whole agent is scaled up to giant size, or that the agent's neck is also long enough that their head reaches the other side of the country, before we agree that the agent has the same duty to save the distant drowning child as they would to the nearby drowning child.

The second concern is that even if we are able to imagine thought experiments in the way that the author of the case intends, we might still bring our intuitions about the actual world into the world of the thought experiment. Fischer (1995) gestures toward this concern when he notes that intuitive responses are not strictly compartmentalised (p. 9). He argues that we spend the rest of our lives encountering moral problems in context, and that it might not be possible to turn off our contextualised reactions when we encounter streamlined hypothetical cases.

Thompson's (1985) Transplant Case helps to illustrate this concern. In this case, a brilliant surgeon has five patients, each with a different failing organ, and they will each die if they do not receive a transplant for the required organ. A healthy passer-by, who just happens to be a match for all five of the patients, drops by the surgeon's practice. We know that all of the transplants will be successful (because this is a thought experiment). Is it permissible for the surgeon to sacrifice the passer-by in order to save the five? Thompson notes that while almost everyone thinks it is permissible to sacrifice one in order to save five in the standard Trolley Problem, almost nobody thinks it is permissible for the surgeon to sacrifice the passer-by in order to save five in the Transplant Case (Thompson, 1985, p. 1396). This is puzzling, given that the underlying philosophical principle in both cases seems to be the same. It seems as though, if our intuitions were unimpinged, we would expect to have the same responses to both cases.

One reason we might have differing intuitive responses to the two cases is that we have smuggled in our intuitions about doctors and our trust in the medical establishment into the Transplant Case. This is both because it is difficult to accept the specification that *all* of the operations *will* be successful, given what we know about the riskiness of transplant surgery in the real-world, and because even if we could really accept the specifications of the thought experiment, we might find it difficult to endorse a moral principle that makes it permissible for doctors to distribute the organs of healthy bystanders to the ill (could we ever trust the motivations of our own doctors once we had accepted the moral permissibility of such a principle?) Our own experience of the medical establishment taints our reading of the Transplant Case, and distracts our intuition away from the principle being assessed: Is it permissible to sacrifice one in order to save the lives of five? If this is a correct reading of what is going on in the Transplant Case, then this shows that not all of the relevant confounders have been streamlined out of the case, because the introduction of medical details is distracting, and so it might not be a good thought experiment for testing the hypothesis. Alternatively, it might be a very good case for testing the hypothesis,

because it gives us a taste of what being committed to the principle that it is permissible to sacrifice one in order to save five would really amount to if we took it seriously. Either way, it seems plausible that our intuitions are responding to factors that have either been explicitly excluded from the case (that the surgeries might not succeed) or to factors that are external to the case (what if doctors were actually permitted to harvest the organs of healthy bystanders?), or to both.

Another example (albeit not an example from moral philosophy) of the smuggled intuition problem is provided by Tversky and Kahneman's (1983) Conjunction Fallacy. In this case respondents are presented with details about Linda's life and are asked to make a judgement about what she is most likely to be doing now. Respondents are told that: "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations" (Tversky & Kahneman, 1983). They are then provided with a range of options about her current occupation, and are asked to rank the choices in terms of likelihood. Amongst the available options are that she is a 'bank teller' and that she is a 'bank teller and is active in the feminist movement'. Overwhelmingly, respondents rank that she is a 'bank teller and is active in the feminist movement' over 'Linda is a bank teller', in terms of likelihood. This is taken to be an obvious error because the conjunction of two events is always less probable than a single event.

Tversky and Kahneman's explanation of why individuals get this case wrong so often is that they are making use of a misplaced heuristic. In this case, they are "replacing the attribute that was the target of the question (the relative probability of the descriptions' truth) with an attribute that comes more easily to mind (the relative resemblance of the description to the introductory statement about Linda)" (Voorhoeve, 2009, p. 68). Bovens and Hartmann (2003) argue that it would be too quick to judge the respondents as merely irrational, and they offer an alternative interpretation. On their reading of the case, the respondent does not assess the probability of the various statements as they stand; rather

they assess the probability of the statements as though they were being informed of the statement by a source (a newspaper, an acquaintance, etc.), which is how one would ordinarily receive such information in the real world. When the respondents assess the likelihood of the statement that Linda is a bank teller, they suspect that the informant is unreliable, given how different this statement is from the information they have already received about Linda from the vignette. When they are presented with the statement that Linda is a bank teller *and* active in the feminist movement, this makes the informant seem more reliable (and the statement thus more likely) because it coheres better with what the respondent already knows about Linda (Bovens & Hartmann, 2003, pp. 85-88). If Bovens and Hartmann are correct – that is, we respond not the statement itself, but to the statement as though it were presented by an informant (which is the way we would typically receive such statements) – this is provides a nice illustration of Fischer's point above about how it is difficult to turn off our real-world intuitive responses in the context of the thought experiment.

If it turns out that our intuitive responses to thought experiments are not as clean as we had initially hoped, then perhaps we should discard thought experiments entirely. The initial claim was that thought experiments are useful because they allow us to streamline out extraneous factors and to focus just on that which is philosophically relevant. If extraneous factors are interfering anyway, then perhaps we are better off just focussing on real-world cases. I am not going to pursue this line of argument here. It still seems preferable to be able to factor out some extraneous factors in order to better focus on certain things, even if this process is imperfect, than to be constantly confronted with the whole "blooming, buzzing confusion [of reality]" – or at least it would be useful to be able to do it some of the time (Elgin, 2014 , p. 222). At this stage it is just worth flagging that there are various ways in which our intuitions about thought experiments might not be as clean as we would have hoped. This should make us somewhat more open to allowing real-world cases into our philosophical methodology, given that a standard objection to consulting real-world cases is that they are too "messy" and that there is too much going on in the real world

that might distract our intuitions, but it seems as though this is also a problem with thought experiments.

1.4. Support Factors and Thought Experiments

In this section I describe what I take to be the central methodological problem with thought experiments: that they run the risk of inadvertently streamlining out support factors when confounders are removed.

So far I have argued that thought experiments play a useful role in philosophical methodology because they allow for extraneous factors that might otherwise interfere with philosophical intuitions to be streamlined out, even if this is only done imperfectly (due to the smuggled intuition problem). However, in the process of excluding potential confounders, required support factors may be inadvertently removed. To borrow (and somewhat adapt) an example from Fischer (1995) to illustrate this problem in the sciences, we can imagine a scientist who is very eager to study the relationship between matches and combustion. In their eagerness to be thorough, the scientist decides to perform the experiment in an airless vacuum to make sure that all potential confounders have been excluded. However, oxygen (an important support factor) has been inadvertently removed in the process, and so the scientist incorrectly concludes that there is no relationship between matches and combustion (Fischer, 1995, p. 10).

The 'Ticking Bomb" case makes it clear that the problem of unwittingly excluding support factors is a real risk for philosophical thought experiments, and not just for scientific laboratory experiments.[14] Walzer (1973) first introduces the Ticking Bomb case in his discussion of the dirty hands problem.[15] In the thought experiment, the reader is asked to imagine a scenario in which a terrorist has been captured. Authorities have good reason to suspect that the terrorist knows

[14] Many thanks to Andy Lamey who brought this case to my attention.
[15] There were precursors to the case, which make the same point about it being permissible to use torture when it would save the lives of a great number of people, but which do not invoke the case of the bomb directly. See, for instance, Gert (1969, p.623).

the location of a bomb (or a number of bombs) that will go off shortly. The question posed is whether it is permissible to torture the terrorist in order to get the information about the location of the bombs, and thus save the lives of innocent people who will die if the bombs are not located in time (Walzer, 1973, p. 173). The issue that the thought experiment highlights is that we typically think torture is unacceptable under *all* circumstances, but this case shows one scenario in which that intuition does not hold – overwhelmingly respondents say that it is permissible to torture the captured terrorist.

Bufacchi and Arigo (2006) argue that our intuitive response to the Ticking Bomb case (that it is permissible to torture the terrorist to save the innocents) rests on relevant details having been stripped away for the purposes of creating the thought experiment – that is, our intuitions are the result of the streamlining process and they are not tracking that which is actually salient. Their point is that once those details have been reintroduced, so that the case more closely resembles actually torturing a captured terrorist, our response to the case would be reversed and we would no longer accept that torture is permissible (p. 359).

To focus in on just one factor that Bufacchi and Arigo highlight as being salient and excluded, the case assumes that the captured terrorist will provide *accurate* information about the location of the bombs. That information will then allow the relevant authorities to locate and defuse the bombs, thus saving the lives of innocents, which makes the torture permissible. However, pre-existing evidence on coercive interrogation techniques show that torture leads to false confessions and inaccurate information being offered in the majority of cases. This is often because prisoners will say whatever they believe the interrogator wants to hear in order to put an end to the torture. Alternatively, savvy prisoners may intentionally give false information in order to mislead their captors and keep their plot intact. For instance, the Japanese captured a US fighter pilot in August 1945, and after "rough interrogation" the pilot told his captors that the US intended to drop atomic bombs on Kyoto and Tokyo (when the truth was Hiroshima and Nagasaki), thus misleading the Japanese and ensuring that the plan went forward unimpeded. In yet another alternative, in scenarios in which

members of organisations are likely to be captured and tortured, and they are aware that this is the case, those higher up within the organisation might intentionally give individuals in lower ranks incorrect information. The idea is that false testimony will then be offered to their enemies when they are predictably captured and tortured (Bufacchi & Arigo, 2006, pp. 361-362). Once we recognise the high probability of false testimony in torture scenarios, it no longer seems that the bombs will be located and that any lives will be saved. Thus, it is no longer permissible to torture the prisoner. The excluded details, once reintroduced, reverse our moral judgment of the case.

The principle that the Ticking Bomb case set out to test was whether torture is *always* impermissible. Assessing the case results in the conclusion that it is sometimes permissible. However, that conclusion is based on having removed information that we know to be true of torture – that it produces unreliable testimony. We only accept that it is permissible in the context of the thought experiment because of the implausible assumption that doing so will produce true testimony and thus saves lives. But what we are interested in is whether *torture* is ever permissible (not some other practice), and unreliable testimony seems to be very closely associated with torture (based on the evidence that we have about it). Removing salient information about the quality of testimony produced under conditions of torture thus means that our intuitions are no longer able to latch on to the relevant philosophical issue. Support factors in the context of thought experiments are those factors that are required to be in place for our intuition to latch on to the salient philosophical principle (see the introduction). The problem in the Ticking Bomb case, phrased slightly differently is thus: unreliable testimony is a support factor of the intuition that it is always morally impermissible to torture, once that support factor is removed our intuition shifts and we conclude that it is sometimes permissible to do so; but unreliable testimony is closely bound to the practice of torture, such that removing it means that we no longer seem to be assessing whether *torture* is ever morally permissible. Removing the support factor means that our intuition misses the philosophical target.

1.5. In search of an appropriate supplement

Given the problem of excluding potential support factors in the process of streamlining, a supplement to philosophical thought experiments is required. This section will assess whether fiction might provide a suitable supplement to the very austere thought experiments that are typically used in philosophy. It will be argued that cases from fiction do not offer an appropriate supplement, because they are subject to the same risks as thought experiments, and that, as such, real-world cases should be used to fill this function. I will conclude this section by noting that precedent exists in the sciences for using real-world cases as a supplement to experimental methods.

Elgin argues that cases from fiction ought to be included in philosophical methodology, because they provide more detail than very schematized thought experiments do. Fictional cases are also more manageable than examples taken from the real world, due to the streamlining process that authors subject reality to in the production of fiction. Fictional cases thus seem like the ideal solution: more detailed, but still manageable (Elgin, 2014 , p. 232). The example that she uses to illustrate this point is that of Jane Austen's novels. Her argument is that they provide substantially more detail than a standard philosophical thought experiment, but by focusing on only three or four families in a boring English village (Austen's characters rarely venture into London where there is a lot going on, there are no peasant revolutions in her novels, etc.), Austen is able to remove extraneous factors and focus in on that which is relevant: the relationships. Elgin makes the point as follows:

> Austen devises a tightly controlled thought experiment. Restricting the factors that impinge on her protagonists enables her to elaborate on the effects of those that remain… Real families, however, are affected by too many forces for the social and moral trajectories exhibited by Austen's characters to stand out. Too many other factors impinge on them; too many descriptions are available for characterizing their lives. *Any sociological study would be vulnerable to the worry that unexamined factors played a non-negligible role in the interactions studied, that other forces were significant* (Elgin, 2014 , p. 233) [emphasis added].

As the italicized text makes clear, Elgin argues for the inclusion of fictional cases over real-world cases because she is concerned that important support factors might go unnoticed when studying cases taken from the real world. Her worry is that we will be so distracted by all of the other things going on in the case that we will fail to pick out that which is really important. However, as I have already argued in this chapter, missing out on support factors is more of a problem for those making use of highly schematized thought experiments, where almost all factors have been intentionally excluded from the description, thus substantially increasing the likelihood that support factors have been excluded in the process. Further, given the streamlining process involved in fiction (which is precisely what Elgin thinks is good about these cases), fictional cases cannot provide a suitable supplement to thought experiments, because they also involve the risk of unwittingly streamlining out support factors.

Fiction also involves an additional risk, which is that the authors might not be streamlining the world in ways that are philosophically useful. The author's aim in producing a fictional case is typically to create a compelling narrative, not to create a rigorous thought experiment. Some of the extraneous and distracting factors that are left behind by the author might make for good fiction, but might not be conducive to philosophical rigour. Writers might also include atypical factors in their descriptions, which may act as confounders when they are taken up by philosophers.

The only suitable supplement to thought experiments is cases from the real world. Williams (2002) makes a similar suggestion when he argues that: "real history *fills* in the merely schematic picture". For instance, in the Ticking Bomb case, it was only by checking the thought experiment against the real world that it became clear that salient factors had been removed in the creation of the thought experiment.

A potential criticism that could be made at this point is that we do not need to turn to the real world to check that support factors have not been accidentally excluded, we can do this just by having more thought experiments, with more of

the potentially relevant factors varied across the experimental cases. The way that Frances Kamm uses trolley problems is a good example of this – she assesses a vast number of trolley problems in which very subtle things are changed in each of the permutations, so that potentially relevant and irrelevant factors are screened in and out of the thought experiment, and that principles can be devised at a suitable level of context specificity. For instance, it was by looking at Thompson's (1985) Footbridge case in contrast to the standard Trolley Problem that it became clear that distance might be a support factor that was present in the standard version of the problem and absent from the Footbridge case. No consultation of the real world was required. However, it still seems like the risk of inadvertently excluding potential support factors persists, because what is and is not varied across the permutations of the thought experiments (even if there are many of them) will still be dependent on the imagination of the philosopher, and they might miss out on something crucial. To reiterate Williamson's point from before: "Any humanly compiled list of such interfering factors is likely to be incomplete" (2007, p. 185).

An example of how we would use real-world cases as a 'supplement' to thought experiments (which is my suggestion in this chapter) is offered by Judith Jarvis Thomson's (1971) Violinist Case, and its relationship to recent philosophical thinking about real cases of pregnancy. Thomson's thought experiment asks that you imagine waking up to find that your circulatory system has been plugged into the circulatory system of an unconscious famous violinist. It turns out that the famous violinist is suffering from kidney disease. The Society of Music Lovers has checked all the available medical records and has determined that you are the only person whose blood is a match to his, so they kidnapped you during the night and hooked you up to the violinist so that your kidneys can circulate his blood. If you stay attached to the violinist for nine months he will make a complete recovery and you can both go back to your regular lives. If you detach yourself at any point in that time period the violinist will die (Thomson, 1971, pp. 48-49). Thomson's argument is that nobody would think that you are morally required to remain attached to the violinist for nine months (although that might be a very nice thing for you to do). Similarly, no woman should be morally

required to act as a human life-support system for a foetus for nine months in the case of pregnancy (Thomson, 1971, pp. 49-50).

In the thought experiment Thompson relies on what has become known as the 'foetal container' view of the relationship between the pregnant woman and her foetus (Kingma, 2016, forthcoming, p. 7; Purdy, 1990). That is, a woman is an incubator in which the foetus resides for nine months. This assumption is what makes the thought experiment work – being pregnant, according to the thought experiment, is like having a stranger plugged into your circulatory system. However, recent work on the relationship between mother and foetus argues that given the various ways that the anatomy of the foetus and the mother are integrated – for instance the foetus "resides not in the uterine cavity, but is implanted in the uterine wall, within the maternal deciduous tissue and is, at least in its early stages, completely covered by it" (Kingma, 2016, forthcoming, p. 18) – it makes more sense to think of the mother-foetus relationship as a part-whole relation (where the mother is the whole and the foetus is one part) rather than thinking of the foetus as something that is merely inside the mother "the way 'a bun is in the oven' or 'a tub of yogurt is in the fridge'" (Kingma, 2016, forthcoming, p. 5). I am not going to venture into what the implications are for the abortion debate by questioning the foetal container picture of pregnancy on which Thomson relies. The point here is just that looking at the real world (for instance, the integration of a mother and foetus's anatomy) can be a useful supplement to philosophical thought experiments (for instance, Thomson's Violinist case), and for checking that nothing important has been excluded or missed out on in the creation of the thought experiment.

It is also useful to note that, maintaining the analogy that has been used throughout this chapter between thought experiments in philosophy and laboratory experiments in the sciences, precedent exists for using real-world cases (or natural experiments) as a supplement to experimental methods (John, 2011, p. 496). One area where this is particularly clear is in the history of psychology. Historically, there was a strong tradition of only studying psychological responses in laboratory settings, in order to maintain rigour and to

ensure that irrelevant factors were excluded. However, concern began to grow within the field that the phenomena of interest might not occur in laboratory conditions, particularly aspects of individuals' social lives – for instance, how would one study psychological aspects of friendship in a laboratory setting? It seems as though many psychological phenomena that we might be interested in just cease to exist in the context of the laboratory. Studies of individuals in their everyday environments were thus included to capture the phenomena that had been previously streamlined out of experimental methods (Dechesne & De Roon, 2014, pp. 186-188). That is, real-world cases were included as a supplement to streamlined experimental cases.

Given that schematized thought experiments and cases from fiction both run the risk of unwittingly excluding support factors when confounders are removed, real-world cases should be included as a supplement in philosophical methodology.

1.6. A criticism

One potential criticism against the position presented in this chapter is that streamlining also occurs when describing real-world cases. That is, even though the case is taken from the world, the process by which it is described before philosophical analysis can occur involves streamlining. Not all of the details can be included in that description. The writer needs to pick out that which they take to be salient, excluding much of what actually happened as they do so. In that selection process, important support factors might be excluded. Someone critical of the position that I advocate might thus argue that real-world cases are just as susceptible to the kinds of problems I have argued are applicable to very streamlined hypothetical cases (like the Trolley Problems) and to cases from fiction (like Jane Austen's novels).

The critic makes a good point. A selection process does occur when real-world cases are described, and so there is a risk that philosophically relevant factors might be factored out in that process. But this is not just a problem that occurs

when *describing* real world cases; it is likely that this would even be a problem for the philosopher experiencing a real-world case directly. For instance, we can imagine a philosopher being present for an interrogation in an actual ticking bomb case. Even when the philosopher is present in the room, there will be aspects of the experience that stand out as salient to them, and others that fade into the background of irrelevancy, and some of the factors that fade into the background might very well be philosophically relevant. Anyone who has had the experience of only noticing a coffee shop in their neighbourhood after walking past it hundreds of times before will be familiar with this experience. This just seems to be a limitation to our ability to engage with the world – we are always streamlining out (hopefully) extraneous factors and doing so imperfectly, and so we are always at risk of inadvertently excluding support factors. The question, then, is whether this problem is more or less severe when describing real world cases for philosophical purposes than it is for streamlined hypothetical thought experiments and cases from fiction.

Part of the problem with the streamlining process that occurs in producing austere thought experiments and fictional cases (in contrast to the streamlining that occurs when describing real cases), is that the philosopher is entirely reliant on their imagination when attempting to figure out what is an extraneous confounder and what might be a support factor. While we are still required to factor things out in our descriptions of real-world cases, and how this is done will be a matter of interpretation, at least the real world places constraints on that interpretation, and we can test our interpretations against the real world – for instance, would the torture in a ticking bomb case really produce the effects that the thought experiment requires to be successful? These checks are absent in fictional cases and more austere thought experiments. As such, real-world cases are less likely than the alternatives to accidentally exclude important support factors. Thus, they still provide a valuable supplement to our philosophical methodology.

In the rest of this chapter, I will consider two further reasons why considering real-world cases might be philosophically beneficial. These are that they might

help us to improve the external validity of our philosophical findings (that is, make the results of our philosophical analysis more likely to transfer to new cases), and they are a useful source of hypotheses and conceptual clarification.

1.7. External Validity

Up until this point the discussion in this chapter has focused exclusively on the issue of 'internal validity', where "a thought experiment is internally valid to the extent that it allows its readers to make judgments that are confident and free from bias or other confounding factors about the hypothesis or point of principle that it aims to test." (Wilson, 2016, forthcoming, p. 5). But as Wilson correctly points out, knowing that a thought experiment has high internal validity, does not tell us how well that result will generalize to further cases (Wilson, 2016, forthcoming, p. 1).[16] Or, to borrow Cartwright and Hardie's language, just because we know that something worked *there* (the test scenario), does not tell us that it will work *here* (the new case that we are trying to assess) (2012, p. 91).

In the case of philosophical thought experiments, knowing that we have a particular intuition about the moral permissibility of an action in the context of a thought experiment (even if that thought experiment was conducted superbly, and we accept that it has high internal validity) does not mean that we would necessarily accept the moral permissibility of similar-looking actions in new scenarios. The result might not generalize – that is, external validity might be low. Wenar (2011) argues that Singer's (1972) Child in the Pond case is like this.

In Singer's original thought experiment, the reader is asked to imagine that they come across a child drowning in a pond. Singer argues that we would all agree that, in that situation, you are morally required to wade into the pond and save the child, regardless of the damage to your shoes that would result (Singer, 1972, p. 231). He argues that if you accept the conclusion that you should save the child

[16] Cartwright makes the same point earlier about Randomised Control Trials, arguing that they have high internal validity, but that does not give us information on how well those results generalize. That is, by themselves, they do not give us information about external validity (Cartwright, 2010).

in the pond, then you are also committed to the idea that you are morally required to donate to international aid organisations in order to save the lives of the distant poor who would otherwise die, based on the principle that if you can prevent a very bad thing from happening without sacrificing anything of comparable moral importance then you are required to do so (Singer, 1972, p. 231). He later goes on to describe the relationship between the Child in the Pond case and our responsibilities to the distant poor like this:

> We are all in the situation of the person passing the shallow pond: we can all save lives of people, both children and adults, who would otherwise die, and we can do so at a very small cost to us: the cost of a new CD, a shirt or a night out at a restaurant or concert, can mean the difference between life and death to more than one person somewhere in the world – and overseas aid agencies like Oxfam overcome the problem of acting at a distance (Singer, 1997).

Singer assumes that translating the results from the Child in the Pond case to the situation of international poverty will be easy: phrased slightly differently, he assumes that external validity almost automatically follows internal validity. Wenar challenges the external validity of Singer's thought experiment. He argues that once we have taken into account the complexity of the international aid system, which sometimes means that our best efforts to help the distant poor actually causes more harm than good (for instance, when donated money is channelled off by the corrupt leadership of a country), it becomes less clear that aiding on the international level is anything at all like wading into the pond to rescue the drowning child (where the required action is straightforward, and the result predictably good) (Wenar, 2011, p. 105).

The questionable external validity of Singer's thought experiment is likely to be a much broader problem in philosophy. Given that a large part of the point of doing moral and political theory is to produce principles that are applicable in a wide range of scenarios, it would be a real problem if the results of our considered cases have low external validity. It also seems likely that this is a widespread problem in philosophical theorising. One of the issues to have emerged in the preceding discussion in this chapter is that philosophical intuitions and the resulting principles are actually quite context-sensitive – whether we think the

agent in the Reach case has an obligation to save the child drowning on the other side of the country, for instance, depends on the details of how the case is imagined.

Wilson (2016, forthcoming; 2014) argues that consulting cases from the real world can help with the problem of external validity. To show how this works he introduces what he terms the 'translational model'. Again, this is a piece of methodological thinking borrowed from the medical sciences. The translational model in medicine developed out of the erroneous way that the relationship between abstract and applied research was historically understood. In the 1940s, the dominant idea was that the bulk of available time and funding should be devoted to abstract research (what was termed "basic research"); the thought being that advances in abstract scientific problems would translate directly into improvements in applied medical science. However, in reality, it was difficult to get advances in the basic sciences to translate into practical biomedical improvements. The translational model in the medical sciences was introduced in an attempt to overcome this. The idea was that research should be funded and conducted across the spectrum – from the very abstract to the very applied – with basic science informing practice, and applied research informing basic science. Wilson argues that philosophical methodology should be thought of in the same way: philosophical study should traverse the spectrum from the most abstract to the most applied, with feedback occurring across the spectrum (Wilson, 2016, forthcoming, pp. 16-17; 2014, pp. 8-13).

Wilson produces a model for thinking about how one would go about moving from a very theoretical piece of research to realising its application (Wilson, 2014, p. 12). However, he does not fully articulate how one might move from the more applied end of the spectrum (real cases) to theoretical developments at the more abstract end, although it is clear that he thinks that this is the case. Cartwright and Hardie (2012) provide us with the theoretical tools to better understand how one might move from applied cases to theoretical developments. In particular, their discussion of "horizontal searches" and "vertical searches" is helpful (and it has the additional benefit of linking back to the previous

discussion on support factors). I discuss each of these terms in the following paragraphs.

Cartwright and Hardie (2012) develop an account of horizontal and vertical searches as part of their broader project of using evidence for policy development, where the 'evidence' will often be an RCT with high internal validity. The concern is then how to move from an experiment with high internal validity to a new policy scenario (this is very similar to the question of how one might move from an intuition about an internally valid case to new scenarios in philosophy). They suggest that doing so involves doing two things. First, you should figure out what the relevant support factors were that produced the effect in the successful experiment, because if you want to achieve the same result in a new scenario, you will need to make sure that those support factors (or appropriate substitutes) are present. Seeking out the relevant support factors is termed the 'horizontal search'. The way this would occur in the philosophical context is to figure out what the salient support factors were in the case being assessed. That is, what it was about the case that generated the specific judgment in response to it – what was the philosopher's intuition latching on to? For instance, this would require that you figure out exactly what it was about the Ticking Bomb case that resulted in your judgment that it is permissible to torture the prisoner in that case. The idea being that if you want to apply your judgment about the case to new cases (for instance, saying it is permissible to torture a prisoner in a new scenario) then you would have to make sure that the same support factors were also present in the new case, which requires that you know what those support factors are. Consulting real-world cases is going to be extremely useful at this stage of the process, given that (as has already been argued) at least in real-world cases we know that all the support factors were present.

The second thing that Cartwright and Hardie suggest is to seek out more abstract causal principles from the specific case being studied (that is, to abstract away from the case), which will allow you to apply the more abstract principle to new scenarios. This is the 'vertical search'. For instance, (to borrow their example),

imagine that you used the claw end of a hammer to pull a nail out of a wall. From the specific case, you can abstract away to a more general causal principle about levers. When you later encounter a large stone that you need to move, you can draw on the more general causal principle about levers that you developed from the hammer scenario and use that to figure out that you need a crowbar (another type of lever) to move the stone (Cartwright & Hardie, 2012, pp. 76-88). In the philosophical analogue, for instance, Singer used his Child in the Pond case to develop the more abstract principle that an agent is morally required to prevent great harm whenever it is the case that nothing of comparable moral importance will be sacrificed in the process of doing so.

In the philosophical context, the idea is that you would start with an intuitive judgment about a case; such as an assessment about the rightness or wrongness of an action in the real world. Then you would try to figure out what contributed to that intuition; potentially listing all of those features. That is the horizontal search. Then you would ask why *those* particular principles brought about that judgement; trying to construct a more general principle that explains why those features yield that judgement – i.e. you develop a more abstract principle. That is the vertical search.

It thus seems, as Wilson suggests, that consulting real-world cases can also help with the problem of external validity.

1.8. Hypothesis Generation and Conceptual Modification

So far, in this chapter, it has been argued that consulting the real world is important for achieving internal and external validity, but little has been said about where the philosophical hypotheses that are being tested come from. In this section it will be argued that considering the real world is also important for generating philosophical hypotheses in the first place (or to keep the analogue

with the sciences going – it provides the 'context of discovery'). Relatedly, real-world considerations help to cast light on new aspects of pre-existing concepts.[17]

The real world is an obvious place to look when one is attempting to generate philosophical hypotheses to test. This offers an additional method for coming up with hypotheses, the alternative being that one studies pre-existing philosophical problems in-depth in an effort to find a conceptual gap in the details that requires filling in (Wilson, 2014, p. 5). A particularly successful example of using the real world to generate a philosophical hypothesis is Miranda Fricker's *Epistemic Injustice* (2007). Here, Fricker argues that we often misjudge individuals who are the members of groups that are the victims of prejudice as not being credible bearers of knowledge. The consequences of this are that we not only miss out on important information that we would otherwise have had access to, but we also subject these individuals to an injustice.[18] It is unlikely that Fricker would have stumbled upon the overlap between issues in epistemology and justice if she had just conducted conceptual analysis of these terms.

Similarly, news from the real world can call into question the plausibility of various pre-existing philosophical positions,[19] and thus creates the impetus for generating new hypotheses to test. Philippa Foot (2009) for instance, reports that it was news on the discovery of concentration camps during the Holocaust that made her question the plausibility of moral subjectivism. She notes: "in the face of the news of the concentration camps, I thought, 'It just can't be the way Stevenson, Ayer, and Hare say it is, that morality is just the expression of an

[17] This can be seen in Chapter 5 of this thesis, where the Mbeki case brings to light the problem of suppressed disagreements at the intersection between culpable ignorance and peer disagreement.

[18] There might be another problem here, which is that many of those who are the victims of epistemic injustice may never read Fricker's work, because they are precisely the people who lack access to university philosophy programmes. But given that the focus in this chapter is on what philosophical benefits can be achieved paying attention to real-world cases, this issue will be bracketed. Philosophy has benefitted from Fricker uncovering the relationship between epistemology and justice, because we now understand a conceptual relationship that we did not before.

[19] This can be seen in Chapter 4 of this thesis, where the disproportionate prevalence of AIDS cases in southern Africa, and the crisis of AIDS denialism, places conceptual pressure on the strict distinction between mono-causal and multi-causal accounts of disease, and provides motivation to resolve this distinction.

attitude,' and the subject haunted me." (Foot, 2009, p. 91). In particular, she was concerned that subjectivism lacked the conceptual resources to be able to tell a Nazi that they were 'wrong', even though it was obviously clear that they were. Foot then spent a substantial portion of her career responding to subjectivism.

In the same vein, consulting the real world can also help to extend and modify pre-existing philosophical concepts. Srinivasan (2016), for example, notes that historically 'marital rape' was considered to be conceptually impossible, but the testimony of women over time eventually made it clear that the pre-existing concept of rape needed expansion to accommodate this (pp. 5-6). Many reasons were offered for why marital rape was a 'conceptual impossibility'. For instance, that marriage marked the merger of the spouses into a single unified entity, making rape within a marriage incomprehensible: "Since nobody can rape himself or herself, it is strictly speaking impossible – a legal oxymoron – for a man to rape his wife" (Burgess-Jackson, 1996, p. 113).[20] Once this spousal merger argument was abandoned (perhaps because of how metaphysically dubious it is), it was argued that the harm of non-consensual sex within a marriage is not as severe as it is in other kinds of situations in which non-consensual sex might take place. For instance, two commenters in the early 1950s arguing against the recognition of marital rape in the law claim that: "though the wife may suffer indignity and shock as a result of [her husband's] action, her suffering is comparably less than that of the victim of the typical rape" (Morris & Turner, 1953 quoted in Burgess-Jackson, 1996, p. 115). Testimony gathered by social scientists called into question this assumption:

> [M]any factors cause wife rape to be more traumatic than rape by strangers and non-intimates; for example, the sense of betrayal, the disillusionment, the fact that it contaminates the entire marriage, and the additional fact that wife rape is often repeated, sometimes for years on end (Russell, 1990 quoted in Burgess-Jackson, 1996, p. 116).

Eventually this concept was expanded upon to include the recognition of marital rape, but this was only done as a result of collecting testimony in the real world.

[20] Burgess-Jackson is reporting an argumentative position here, not endorsing it.

Paying attention to the real world is thus important for hypothesis generation (as in Fricker's case), discovering challenges to existing philosophical theory (as in Foot's case), and for expanding and adapting our concepts (as in the case of marital rape).

1.9. Conclusion

In this chapter I have argued that real-world cases should be included in philosophical methodology. Thought experiments are effective because they streamline out extraneous factors and allow one to focus on just that which is philosophically relevant. However, the process of streamlining also runs the risk of overlooking important support factors, especially because support factors might be invisible until the full mechanism driving the intuition is known, or until a support factor is removed and the intuition no longer holds. In most cases we are producing streamlined cases for philosophical analysis without a good sense of which elements are likely to be confounders and which are support factors (we are streamlining blind). Cases from fiction are subject to related problems because they are produced via similar streamlining processes, and they involve the additional risk that the elements that are included and excluded are decided by narrative effectiveness, not by philosophical rigour. A supplement to thought experiments and fictional cases is thus required. Real-world cases fill this gap, because at least we know that all of the required support factors are present in the real-world case. Critics might argue that when real-world cases are described for the purposes of philosophical analysis, streamlining also occurs and important support factors might be inadvertently excluded. However, this concern is less warranted for the real-world case than when it is posed to thought experiments and fictional cases. At least we can check our descriptions of the world against reality, and determining what may or may not be a support factor (or a confounder) is not entirely reliant on the imagination of the philosopher.

Additionally, consulting real-world cases can be beneficial for better translating our philosophical responses from specific cases to novel scenarios (i.e. they help with external validity). This is because we can subject our assessments of real world cases to horizontal and vertical searches. Through the horizontal search we figure out what the support factors were that drove the intuition in that scenario, and through the vertical search we develop more abstract principles, which are then applicable to new cases. We thus have procedure for translating the results from particular philosophical cases to new contexts.

Finally, the real world provides us with the context of philosophical discovery. Consulting the real world can help to generate philosophical hypotheses in the first instance (such as with Fricker's uncovering of philosophical material at the intersection between epistemology and justice). It can challenge existing philosophical principles (such as Foot's realisation that the concentration camps of the Holocaust pose a challenge to moral subjectivism). Further, consulting the world can help to modify and expand pre-existing philosophical concepts (such as the role that social science research played in showing the possibility of marital rape).

Given all of the philosophical benefits that accrue from paying attention to the world, close study of real-world cases should be a legitimate part of philosophical practice.

2. The Case of Thabo Mbeki's AIDS Denialism

2.1. Introduction

Having justified the detailed philosophical study of real-world cases in the previous chapter, in this chapter I describe the case that is the focus of this thesis. In particular, I provide a chronology of Mbeki's AIDS denialism. There is substantial controversy over the events of this period, and what was motivating Mbeki to act as he did (Steinberg, Forthcoming, 2017, p. 4). In this chapter I aim to cut through the controversy and provide the reader with as simple a description as possible in order to make comprehensible the philosophical problems that follow in the course of this thesis. In chapters 4 and 5 I provide more detail on what I take to have motivated Mbeki. In chapter 4, I argue that at least part of what occurred was that Mbeki (correctly) recognised that there was an explanatory gap in the biomedical literature at the time, which made it difficult to account for the disproportionate prevalence of AIDS in southern Africa, and which led him to (incorrectly) consult non-mainstream AIDS scientists. In chapter 5 I argue that his distrust of various sources of disagreement with his view, and suppression of disagreement within his political party, resulted in him being ignorant on certain key issues, and this also helps to explain his motivations.

I divide the discussion into three chronological sections: 1) the period prior to Mbeki's denialism (1980s-1999); 2) the period of Mbeki's active denialism (1999- 2000)[21]; and 3) the post-denial period (2000-2016).

[21] There is some controversy over using 2000 as a cutoff point, because the government only began to roll out an ARV programme in 2003, and Mbeki seems to remain committed to his denialist views even now. Wang (2007) takes 2003 to be the cut-off point, because this is when the government committed itself to providing ARVs. I take October 2000 to be the end of Mbeki's "official" denial period, because this is when he publically stepped back from the AIDs debate, even though he continued to influence policies through other members of the party (Posel 2008, 18).

2.2. Pre-Denial (1980s- 1999)

Both the apartheid government and the ANC believed that AIDS would never be a problem for South Africa, assuming that it would largely be restricted to white homosexual communities, as had been the case in other parts of the world (Vale & Barrett, 2009, p. 451; Gumede, 2005, p. 152; Feinstein, 2007, p. 130). Homophobia and an unwillingness to discuss sexuality, which united the Christian apartheid National Party (NP) and the anti-apartheid liberation groups, meant that nothing was done about AIDS until the early 1990s (Butler, 2005, p. 593). At this point, prevalence of the disease was extremely low, with only 0.76% of the population infected by 1990 (Karim & Karim, 2002, p. 39).

In 1992, the National AIDS Convention of South Africa (NACOSA), made up of the NP's health ministry, and representatives from the ANC and the United Democratic Front (a coalition of anti-apartheid liberation groups other than the ANC), met for the first time to draw up a strategy for tackling AIDS as the country underwent democratic transition. Motivation for this meeting seems to have been partially due to an awareness of rising infection rates, but also because leading members of the anti-apartheid movement had been living outside of South Africa in areas that had much harder hit by the epidemic, and they had borne witness to the effects of the disease. South Africa had been granted partial early protection from the epidemic due to the isolation that came with sanctions, but individuals like Chris Hani (a major figure in the ANC at the time, but who was assassinated just before democracy was realised) were especially concerned that South Africa might suffer a similar fate.

The ANC assumed office in 1994 with an impressive AIDS plan that had been drawn up at NACOSA, and it was made a 'Presidential Lead Project', with the commitment that adequate funding would be allocated to ensure its realisation. However, despite this plan, in the early days of democracy AIDS was not a priority issue and little progress was made on implementing it (Posel, 2008, p. 15; Butler, 2005, p. 593; Chikane, 2013, p. 257; Gumede, 2005, p. 152). In one description of the ineffectiveness of the Mandela presidency on issues of HIV and

AIDS, Marais comments: “Measured minute by minute, during his presidency Mandela probably spent more time with the Spice Girls and Michael Jackson than he did raising the AIDS issue with the South African public” (Marais 2000 quoted in Nattrass 2007, p. 40).

In 1994, an inter-departmental committee was established to discuss AIDS. Mbeki, then the deputy president, took responsibility for government oversight of all AIDS-related projects, across multiple departments. Both Chikane (2013, 258) and Butler (2005, 594) describe Mbeki as being a “champion” of the ANC’s AIDS policy in the 1994 - 1998 period. Chikane, director-general of the presidency at the time, recalls that Mbeki required all senior government officials and members of Cabinet to wear the red ribbons that had become emblematic of the struggle against AIDS at all times; reprimanding them if he caught them without their ribbons, because he wanted to make sure that the message to the public was clear: The government is taking AIDS seriously. Chikane notes that at “this stage there was no inkling that there would ever be any controversy about HIV and AIDS” (Chikane, 2013, p. 258).

Despite Mbeki’s enthusiasm for the ANC’s early AIDS strategy, there were two high profile scandals associated with the government’s AIDS intervention during this period. The first was the *Sarafina II* debacle in 1996. The then-Minister of Health, Nkosazana Dlamini-Zuma, spent a disproportionate amount of the annual AIDS budget (most of which was international donor money) on producing a musical that would travel the country educating people about HIV and AIDS. This has been overwhelmingly rejected as a waste of scarce resources, which had little impact on public awareness about the disease. There were also a number of “financial irregularities” associated with the production, which ultimately led to the project being scrapped prematurely (Posel, 2008, p. 15; Gumede, 2005, p. 153).

However, despite being included in most chronologies of the denialist period (see Posel (2008), Butler (2005), Gumede (2005) and Nattrass (2007)) the *Sarafina II* scandal was restricted to the Health Ministry, with little connection to Mbeki and

nothing to do with the denialism that came later. The second AIDS scandal from this period – the "Virodene Scandal" – was far more telling of things to come; both in terms of the individuals involved and the content of the incident.

In 1996 a group of researchers from the University of Pretoria claimed that they found a cure for AIDS, which they called 'Virodene', but the MCC (Medical Control Council – the body tasked with overseeing medical research) refused to accept the research protocol, effectively shutting down the possibility of a trial (Gumede, 2005, pp. 153-154; Posel, 2008, p. 15). The MCC continued to reject versions of the protocol over the 1997 – 1998 period, and it was also rejected by the University of Pretoria's own internal research ethics committee (Myburgh, 2009, pp. 5-6). Faced with the prospect of not being permitted to conduct a trial, the researchers made personal contact with the Minister of Health and Mbeki in an attempt to bypass the MCC.

Mbeki found the prospect of Virodene appealing, not just because it was substantially cheaper than the alternative antiretroviral therapies available on the international market, but because it was produced at a South African university, which cohered well with his commitment to "African solutions for African problems." In 1997 the Virodene researchers addressed Cabinet directly, on special invitation from Mbeki himself, bringing with them a small number of patients who had received the treatment. These patients offered testimony of the miraculous effects that the drug had had on their lives (Gumede, 2005, p. 154; Myburgh, 2009, p. 4; Posel, 2008, p. 15). But it turned out that Virodene was not the cure that everyone had hoped for. In an independent review, it was concluded that the Virodene research team had contravened trial procedures, having tested it on humans before successfully completing laboratory or animal trials. It was also found that the main ingredient was an industrial solvent, which was already well-known to cause severe liver damage in humans (Gumede, 2005, p. 154). However, Mbeki's personal support of Virodene and its researchers, and his irritation with not being able to overrule the MCC were prescient of his later denial. In Posel's words:

> Then Deputy President Mbeki's backing of the Virodene trials was ardent and his critique of the MCC pointed and angry, giving perhaps the earliest signs of what would grow into his more thoroughgoing distaste for the power of the scientific establishment with respect of AIDS. (Posel, 2008, p. 15)

There were also rumours that members of the MCC were removed due to their refusal to accept the Virodene trial protocol, and that the ANC had financial stakes in the drug's success (Posel, 2008, p. 15; Myburgh, 2009, p. 6). Whether or not these rumours are true, it is largely accepted that the Virodene incident primed Mbeki to consider non-mainstream approaches to AIDS science, and that it was the start of his antagonistic relationship with the scientific establishment (in this case, antagonism with the MCC). But at this point Mbeki's approach to AIDS was still based on mainstream AIDS science (Gumede, 2005, p. 155).

In June 1999 Thabo Mbeki was elected as South Africa's second democratic president, and his position on AIDS quickly departed from the scientific mainstream.

2.3. Denial (1999-2000)

In July 1999 Anita Allen, a science journalist, and Anthony Brink, a lawyer, sent dossiers of AIDS denialist material to Mbeki. When Mbeki described these events to his biographer, Mark Gevisser, he explicitly cites this as the event that "sparked it off" (Gevisser, 2007, p. 729).

By all accounts, Mbeki proceeded to do a considerable amount of independent research on HIV and AIDS over the following months, mostly on the Internet (Cherry, 2009, p. 18; Gumede, 2005, p. 159). In October he announced at the National Council of Provinces meeting that he had uncovered a large literature online that questioned the safety of AZT (an antiretroviral) and encouraged other members of the Council to read up on this issue (Gumede, 2005, pp. 158-159; Chikane, 2013, pp. 261-262). At his end-of-year holiday with his university friends from England, he showed them a number of AIDS denialist websites – indicating that Mbeki had indeed been reading denialist material online in the

intervening months. In January 2000 the sheer volume of Mbeki's reading on HIV and AIDS became clear, when he sent a dossier of over fifteen hundred pages of denialist material to Malegapuru Makgoba, a leading South African immunologist and president of the Medical Research Council (MRC), who Mbeki was trying to win over (Gevisser, 2007, p. 742).

After extensively reading non-mainstream HIV/AIDS material, Mbeki made contact with the relevant scientists directly. In January 2000 he wrote to David Rasnick, a self-proclaimed "dissident" AIDS scientist at Berkeley, California. Even at this point it seems that Mbeki was still mostly committed to the mainstream view of HIV and AIDS. His letter to Rasnick consisted of a series of questions, along with the answers he had received from his Minister of Health, Manto Tshabalala-Msimang. Tshabalala-Msimang's answers were entirely consistent with mainstream AIDS science at this point. The only hint that Mbeki gave that he might be sympathetic with Rasnick's view was putting the word "HIV" in quotation marks in his letter (Gevisser, 2007, p. 742; Cherry, 2009, p. 25; Gumede, 2005, p. 159). Rasnick's responses included reference to poverty as a cause of AIDS, and shortly after Mbeki began including comments about the causal salience of poverty in his speeches (Gevisser, 2007, p. 743). In February he gave the State of the Nation Address, in which he explained that government would not only be intensifying their public health messaging about AIDS (in line with mainstream scientific views), but also stressed the necessity of considering poverty in conjunction with AIDS (still permissible within the mainstream paradigm, but starting to edge toward what was non-mainstream science at the time, and echoing Rasnick's response letter) (Gevisser, 2007, p. 743). Soon after, he made contact with Peter Duesberg, the unofficial leader of the "dissident" AIDS movement (more will be said about Duesberg in the following chapter) (Gumede, 2005, p. 159).

In March 2000 it was announced that an expert panel was being established to "explore" the AIDS issue and to provide advice to the presidency. Amongst the items on their brief, panellists were asked to consider issues related to the disease's aetiology (Cherry, 2009, p. 25). Just under half the scientists on the

panel represented the denialist view. John Moore, a well-known mainstream AIDS scientist commented that: “The panel has pretty well everyone on it who believes that HIV is not the cause of AIDS, and about 0.0001 per cent of those who oppose this view” (Moore quoted in Cherry 2009, 26).

The panel had their first meeting at the start of May 2000. In July they met again, but the panel had been enlarged from 37 members to 52, with the new members representing the mainstream scientific view, placing the denialists firmly in the minority. The hope was that the change in composition of the panel was indicative of an upcoming shift in government policy (a hope that was unfortunately not realised) (Cherry, 2009, pp. 27-30). Unsurprisingly, the panel was unable to reach consensus and their final report consisted of contradictory recommendations. Mbeki took this to indicate that AIDS science was deeply contested, and he used this as justification to delay the provision of antiretroviral therapy via the public health system until 2003 (Cherry, 2009, pp. 34-35; Nattrass, 2007, p. 91).

Mbeki came under severe pressure from local civil society groups, notably the Treatment Action Campaign (TAC), and lost credibility in the international press, who “widely believed that insofar as his views were sincerely held, he was mad” (Steinberg, Forthcoming, 2017, p. 5). This was particularly worrying to Mbeki and the ANC because of the local elections that were scheduled for later that year. The senior party leadership convinced Mbeki to publicly back down from the AIDS issue. Importantly, they did not criticise him directly, but rather expressed concern that there had been a breakdown in communication, and that the message about AIDS had become confused (that they did not criticise him directly will become important in Chapter 5, when suppressed disagreement within the ANC will be discussed).

Mbeki eventually agreed that he would abstain from public comment on AIDS, at least until after the local elections, but on the understanding that he would not recant any of his previous statements (Gumede, 2005, pp. 165-166). On 26 September 2000 Mbeki publically announced his withdrawal from the AIDS

debate (Feinstein, 2007, p. 123; Cherry, 2009, p. 32). Joel Netshitenzhe (the head of communications for the ANC) was put in charge of undoing the public relations damage that the AIDS issue had created, a task that was all the more difficult because of Mbeki's insistence that he would not back down from any of his previous statements. Netshitenzhe embarked on an extremely well-funded publicity campaign. In particular, he focussed on arguing that the media had intentionally misunderstood the government's position on AIDS, and that no member of the ANC had ever claimed that there was no causal connection between HIV and AIDS (Gumede, 2005, p. 166).

2.4. Post-Denial (2000- Present)

Despite Mbeki's exit from the public debate, in practice he continued to influence South African HIV/AIDS policy via other members of the party, most notably through Tshabalala-Msimang (the Health Minister), despite the fact that she had initially adhered to the mainstream scientific view. Posel refers to Tshabalala-Msimang's role in the AIDS controversy as being Mbeki's 'chorus' (i.e. uncritically reflecting his views) (Posel, 2008, p. 18). Steinberg makes similar comments about Tshabalala-Msimang's role, describing her as being under Mbeki's "sway" (Steinberg, Forthcoming, 2017, p. 4). Fred Mouton, a popular satirical cartoonist, published a cartoon at the time in which he drew Tshabalala-Msimang as a puppet and Mbeki as the puppeteer (Nattrass 2007, 74).

That Mbeki had no real intention of exiting the AIDS debate was immediately clear. Two days after the public announcement, he addressed the internal ANC caucus, and made statements about AIDS that were more extreme than he had ever made in public,[22] such as questioning whether the virus exists at all (Feinstein, 2007, pp. 124-126). There are other occasional instances in which it is

[22] The ANC caucus occurs in private. The only reason we know about these events is because Andrew Feinstein, a member of the ANC and a Member of Parliament, leaked his notes from the caucus meeting to the press, and subsequently included his version of events in his book on this period of the ANC's history, *After the Party* (2007).

clear that despite being (mostly[23]) quiet about AIDS in public, he continued to promote his denialist stance within the privacy of the party.

One particularly noteworthy instance of this occurred in March 2002, when an anonymous monograph entitled, *Castro Hlongwane, caravans, cats, geese, foot & mouth and statistics: HIV/AIDS and the struggle for the humanisation of the African*, was circulated to Members of Parliament. In this document some of the most extreme versions of the denialist stance are endorsed – for instance, that HIV might not be causally relevant to AIDS at all, and that pharmaceutical companies were deliberately inflating the estimates of the number of HIV positive people in southern Africa in order to increase sales (Anonymous, 2002). At the time there was widespread speculation that Mbeki had written the *Castro Hlongwane* document, due to the similarity between it and Mbeki's writing style, and because an embedded signature on the document suggested that it had been written on Mbeki's computer (Wang, 2007, p. 16; Barrall, 2002). In a public letter this year (2016) Mbeki has finally claimed co-authorship of the document (Mbeki, 2016). There is also strong evidence that Mbeki continued to quietly work on the document and update it for many years after he exited the public debate. While Gevisser was preparing Mbeki's biography in 2007, Mbeki asked him if he had read the *Castro Hlongwane* document, to which Gevisser responded that he had. Shortly after, Mbeki couriered him an updated version, which was twice as long as the original and included references up to August 2006. While Mbeki did not confess authorship of the document to Gevisser at the time, he did admit that the document reflected his views (Gevisser, 2007, p. 736).

While Mbeki continued to work on the denialist project privately, he was losing the public battle, especially due to the activities of the TAC. The TAC had been established in 1998 to campaign for a comprehensive care package to be

[23] Occasionally he slipped up and did make public comments on AIDS that echoed his previous denialism. In particular, at a public lecture at the University of Fort Hare in October 2001 he commented on the "strange opinions" about disease that South Africans have been forced to adopt, and reprimanded civil society groups for what he took to be the implied racism of their activism against the state in their efforts to secure anti-retroviral therapy via the public health system (Mbeki, 2001). However, it seems as though Mbeki might plausibly have believed he was in keeping with his agreement with the ANC because he did not mention AIDS by name in the speech (Forrest and Streek 2001).

delivered by government. They initially thought that their biggest challenge would be battling the high prices charged by the pharmaceutical companies, not the government itself, and they even partnered with government in legal proceedings against the pharmaceutical industry (Geffen, 2010, pp. 48, 55-56). When it became clear that government would continue to block ARV access, even once the prices had been brought down, the TAC's activities turned toward government.

The TAC organised a number of large public protests, and ultimately took the Department of Health to court in an effort to force them to provide the one-off shot of Nevirapine that was required to prevent mother-to-child transmission at the time of birth (Nattrass, 2007 , p. 95). In December 2001 the courts found in favour of the TAC, compelling government to provide Nevirapine at birth to those who required it (Nattrass, 2007 , p. 97). Despite delays created by the decision being appealed at the Constitutional Court, the court decision marked the beginning of the end of the government policy of withholding ARVs (Gevisser, 2007, p. 754). The Treatment Action Campaign was once again successful at the Constitutional Court. Under compulsion of the court ruling, in 2002 government declared that Nevirapine would be made available to all pregnant women. Shortly after, it was announced that ARVs would be made freely available to rape victims, and shortly after that the government declared that ARVS would be provided for all South Africans living with AIDS (Gevisser, 2007, p. 757; Steinberg, Forthcoming, 2017, p. 5). The Department of Health began to roll out the national ARV programme in 2004 (Gumede 2005, 172).

The era of South African AIDS denialist policy ended in 2004 when ARVs became available via the public health system. Mbeki, however, has remained committed to his view (Gumede 2005). He commented to Gevisser that he believed it was "very unfortunate" that he was made to withdraw from the public debate in 2000. More recently, in a public letter this year (2016) he reiterated that he does not believe that a virus can cause a syndrome (while also claiming that he never rejected the position that HIV causes AIDS – it is unclear how these two positions can be compatible). In the same letter he also restated that poverty needed to be

given as much consideration as the virus, and he quoted (at length) an interview from *House of Numbers: Anatomy of an Epidemic* (2010), a very well-known denialist documentary. Mbeki's denialism persists.

2.5. Conclusion

This chapter has provided the reader with a chronology of the development of Mbeki's denialist beliefs in the late 1990s and early 2000s.The next chapter will take a chronological step back to examine the state of AIDS science at the point when Mbeki entered the debate. In particular, I will make use of Lakatos's theory of scientific research programmes to assess whether it was rational for the scientific community to have accepted HIV as the definitive cause of AIDS in 1984.

3. Was it rational to accept the viral account of AIDS?

3.1. Introduction

The previous chapter described the case of South African AIDS denialism in the early 2000s, but very little was said normatively about Mbeki's role in these events. Was it a good idea for Mbeki to include non-mainstream scientists on his advisory panel? Can he be held morally responsible for the consequences of having done so (the subject of chapter 5)? In order to assess the reasonableness of Mbeki's decision about who was included on his advisory panel (a mixture of both mainstream and denialist scientists), and, ultimately, to determine whether he was morally responsible for the consequences of the resulting policy, we must first assess the state of debate about AIDS within the scientific community. This chapter provides a picture of that debate; beginning with the scientific uncertainty at the start of the 1980s when the first cases of the disease were identified, through to the isolation of the HIV virus in 1983, and into the period of mainstream science versus denialism.

The overall aim of this chapter is to assess whether it was rational for the scientific community to accept HIV as *the* cause of AIDS in the 1980s, because if it was rational to do so, then the denialists were reasonably excluded at that point and ought not to have been included on Mbeki's advisory panel. The assumption here is that policy should be based on the 'best' science, leaving the scientific debates to the scientific community itself (Jones, 2002). If the state of the AIDS debate within the scientific community was genuinely contested, or there was evidence that debate within the scientific community had been suppressed or was biased in some way, then including non-mainstream scientists on the Presidential Advisory Committee might have been understandable. If there was no such contestation or evidence of suppression, and the viral account of the disease became dominant because it was just the better theory, then the decision to include non-mainstream scientists in an advisory role is more difficult to justify.

Contemporary AIDS denialists claim that the HIV hypothesis was adopted irrationally and prematurely, thus focussing scientific attention and resources on the viral research programme, and neglecting competing (possibly better) theories. For example, Treichler, in *AIDS, HIV, and the Cultural Construction of Reality* (1992), argues that the scientists working on viral explanations of AIDS were able to:

> ... stake out a fairly ambitious territory: By repeatedly citing each other's work a small group of scientists quickly established a dense citation network, thus gaining early (if ultimately only partial) control over nomenclature, publication, invitation to conferences, and history (p. 76).

She goes on to argue that the concentration of resources in the viral research programme stunted the growth of competing theories. Contemporary AIDS denialists claim that they were victims of this process, and that as a result they were never able to properly pursue their research. Peter Duesberg, the unofficial head of the denialist movement (more will be said about him later in this chapter), claims the same when he states that: "This fatal assumption [that HIV causes AIDS] mostly was the result of a rush to judgment in 1984" (Duesberg, 1994). He goes on to argue that questioning the causal connection between HIV and AIDS cut him out of the scientific debate:

> I had all the students I wanted. I had all the lab space I needed. I got all the grants awarded. I was elected to the National Academy. I became Californian Scientist of the Year. All my papers were published. I could do no wrong. Almost, professionally that is, until I started questioning the claim that HIV is the cause of AIDS. Then everything changed (Duesberg, 1999 quoted in Kalichman 2009,p.32).

If it was the case that the viral theory of AIDS was adopted irrationally and prematurely in 1984, resulting in scientists working on non-mainstream views being unfairly excluded, then perhaps they are victims of an epistemic injustice.[24] If so, they might have a moral claim to being recognised as credible bearers of knowledge, and Mbeki's inclusion of non-mainstream AIDS scientists would have been the just thing to do, making Mbeki a moral hero. It is also worth noting that "Duesberg's chief gripe… is that scientists have simply not taken him seriously" (Kalichman, 2009, p.

[24] I am borrowing Miranda Fricker's term here, specifically her description of a situation in which "the hearer makes an unduly deflated judgement of the speaker's credibility, perhaps missing out on knowledge as a result; and the hearer does something ethically bad—the speaker is wrongfully undermined in her capacity as a knower" (Fricker, 2007, p. 17)

54). Further, Nattrass (2010) argues that it is precisely this claim of unfair exclusion that has motivated the picture of 'Duesberg-as-oppressed-hero-scientist' and has contributed to the lasting appeal of the denialist movement; bestowing Duesberg's supporters "with the thrilling identity of being in receipt of 'the truth', and as brave whistle-blowers standing up for 'real' scientific progress" (Nattrass, 2010, p. 248; Nattrass, 2012, p. 5). Duesberg's supporters strongly endorse the picture of him as a 'latter-day Galileo' (to borrow Nattrass's (2012, p.5) turn of phrase), and the denialist documentary *Positively False* (2012) starts with a montage that includes both Duesberg and Galileo's images.

Given that claims of unfair exclusion have motivated the persistence of AIDS denialism, it important to examine this complaint closely. Mbeki himself is obviously drawn to the image of AIDS denialists as oppressed "dissidents" (their preferred descriptor), and in his capacity as an anti-apartheid struggle hero in South Africa, he sympathises with their apparent experience of oppression:

> It is suggested…that there are some scientists who are 'dangerous and discredited' with whom nobody, including ourselves, should communicate or interact. In an earlier period in human history, these would be heretics that would be burnt at the stake! Not long ago, in our own country, people were killed, tortured, imprisoned and prohibited from being quoted in private and in public because the established authority believed that their views were dangerous and discredited. We are now being asked to do precisely the same thing that the racist apartheid tyranny we opposed did, because, it is said, there exists a scientific view that is supported by the majority, against which dissent is prohibited (Mbeki, 2000).

It is also clear that the denialist scientists who were on Mbeki's advisory panel felt that an epistemic injustice had been rectified by their inclusion. David Rasnick, an associate of Peter Duesberg's, and a member of Mbeki's advisory panel reflected on the experience that: "it was a victory from the moment we stepped off the plane" (Rasnick quoted in Conlan, 2000).

In this chapter I take seriously the denialist complaint that the viral account of AIDS was accepted prematurely, to the exclusion of other promising theories. In order to do this, I provide a description of early AIDS science in the 1980s, with emphasis on the two most prominent theories of the time: the immune overload

theory – the theory that AIDS is the result of immune system collapse due to lifestyle factors such as drug use and repeated STD exposure (this theory is the precursor to the most plausible AIDS denialist theories) – and the viral theory of AIDS – the theory that AIDS is caused by a virus. I will then provide a description of how the denialist position emerged after HIV was isolated, and I will use Lakatos's theory of scientific research programmes (with some additions from Thagard (1978)) to examine the reasonableness of the shift from multiple hypotheses about the cause(s) of AIDS to the acceptance of HIV as the 'most likely cause'. I conclude that it was reasonable to accept that HIV causes AIDS, especially by the time effective antiretroviral therapy became available in the mid-1990s. However, I will conclude that the application of Lakatos's approach to the AIDS case leaves us with two lingering concerns: first, that Lakatos's approach allows for denialist scientists to continue rationally working on the immune overload thesis; and second, that there was some explanatory loss that occurred in the transition from multiple theories of the disease's aetiology (one of which was the multi-causal theory) to the viral theory alone. In particular, it is difficult to explain the disproportionate prevalence of the disease in southern Africa without appealing to factors beyond the virus. Accounting for this explanatory loss will be the task of Chapter 4.

Before proceeding, it should be noted that much of the denialists' work is focussed on discrediting the claims of mainstream AIDS science, but not making many positive claims of their own (Kalichman, 2009, p. 11; Steinberg, 2009, p. 34). While the aim of this chapter is to give denialist scientists a fair hearing by taking their position seriously, there are often silences from the denialist quarter on issues for which it would be good to have some response in order to fully assess the position – however, they might argue that they have been deprived of the resources to do the research that would be required to resolve these questions. This is just an unavoidable problem with the way this debate actually unfolded, but it poses some limitations on reconstructing the debate in order to assess it.

Another challenge for reconstructing this debate is that so little of it has occurred within the formal academic sphere: both because mainstream scientists have typically ignored the denialists, and because the denialists have found alternative non-academic platforms to make their points. Kalichman suggests that the lack of engagement from mainstream scientists with the denialist view is because much of the scientific establishment is unaware of the persistence of denialism, due to the fact that those who have been persuaded by the denialist stance are precisely those who no longer engage with the scientific establishment, granting them a kind of invisibility (Kalichman, 2009, p. 1). As a result, mainstream AIDS scientists often do not publish responses to denialists because it is assumed that denialists no longer exist. Further, mainstream AIDS scientists view it as a waste of time and resources to continue debating these topics. In 2003 the editor of the *South African Medical Journal* (SAMJ) announced that they would no longer be publishing material on this debate, stating that continuing to do so would distract from the important work of actually tackling the epidemic (SAMJ, 2003).

With AIDS denialists engaging so little with the mainstream scientific community and its publications, much of the denialist stance is articulated in places that are not traditionally academic: such as periodicals, self-produced documentaries, and most importantly, the Internet (Kalichman, 2009, pp. 59, 93-96; Steinberg, 2009, p. 34). John Moore, an immunologist and leading AIDS scientist, has commented that: "Denialism has been relegated to the fringes of the internet" (Moore quoted in Steinberg, 2009). Given that much of the denialist debate takes place on non-academic platforms, when mainstream scientists *do* respond, they tend to do so via the same non-academic platforms as the denialists (Nattrass, 2012). This poses a challenge for dealing with this debate in an academic way, because it means that reconstructing various positions requires drawing on traditionally non-academic sources.

3.2. Setting out the Philosophy: Lakatos and the Demarcation Problem

This section describes Lakatos's theory of progressive and degenerating research programmes, which will be used later in this chapter to assess whether it was

rational to accept the HIV theory of AIDS over the immune overload theory in the 1980s. Lakatos's approach to the demarcation problem – the problem of distinguishing between science and non-science (Larvor, 1998, p. 48; Lakatos, 1978, p. 1) – is favoured for the purposes of this chapter (with some additions from Thagard (1978)) over the alternatives provided by Popper and Kuhn. I begin by justifying this choice.

The accepted interpretation of Popper's description of 'good science' is that it involves developing theories that are susceptible to falsification – that is, the structure of the theory should be such that it has the potential to be disconfirmed by counterexamples (Popper, 1959/2002). However, this approach would have left all competing approaches to AIDS science in the 1980s 'dead in the water', because each theory had problems which would have (on Popper's account) resulted in their falsification and dismissal, because "for a Popperian, a single counterexample kills a theory" (Larvor, 1998, pp. 57, 59). For instance, the early inability to detect HIV in the cells of AIDS patients would have acted as a counter-example to the HIV thesis, while the extension of the AIDS pandemic to non-traditional risk groups (such women and children) would have resulted in the dismissal of the immune-overload thesis. Lakatos articulates the same concern about Popper's approach, when he states that all theories "are born refuted and die refuted" (Lakatos, 1978, p. 5). An approach that would have us dismiss all early AIDS science as irrational is unhelpful, because it cannot allow for any of the theories to have been rationally accepted, and this makes it difficult to account for the progress that has been made in AIDS research in the last thirty years. Popper's approach does allow for some initial error, in that he accepts that something might look like a counterexample when it is actually not, due to problems with the reliability of evidence collection (Popper, 1959/2002, p. 28). However, it does not allow for theories being a bit wrong at the start.

Kuhn (1962/1970) is also not particularly useful for the project at hand, because his concerns are on a different scale to what would be appropriate for assessing the AIDS debate. Kuhn is concerned with 'scientific revolutions' – that is, when major paradigm shifts occur. Phrased slightly differently, Kuhn is concerned with

those scientific periods when the very "rules and standards for scientific practice" come under scrutiny (Kuhn, 1962/1970, p. 11). Kuhn argues that the rest of the time, 'normal science' persists. Normal science occurs in those periods when the rules and standards are agreed upon, and scientists are focussed on 'puzzle-solving'. During these periods, scientists do not "aim to produce major novelties" and focus instead on figuring out manageable problems within the existing paradigm (Kuhn, 1962/1970, p. 35). All AIDS science in the 1980s falls well within the ambit of Kuhn's normal science – none of the competing aetiological accounts of AIDS in the 1980s were concerned with changing the rules and standards of biomedical science. Accounts of disease that appeal to infectious microbial agents (such as the aetiological descriptions of the flu, Tuberculosis, Ebola, etc.) and accounts of disease that make use of lifestyle related factors in their explanations (such as the explanations of diabetes, heart disease, etc.) are both acceptable modes of explanation within the biomedical sciences (more will be said about this in Chapter 4). The puzzle in the case of early AIDS science was to figure out what kind of disease the scientists were looking at (infectious or lifestyle-related). The problem fits well within the scope of Kuhn's normal science.

Lakatos is favoured in this chapter, because his account allows for us to specify under what conditions the shifts in early AIDS research may have been rational, despite all of the available theories at the time having their own problems, and without having to appeal to any major paradigm shifts.

The unit of analysis for Lakatos is the 'scientific research programme', which refers to a set of theories, some of which make up the 'hard core' – the central claim to be defended – and others which constitute the 'protective belt' – ancillary theories that can be modified or dismissed in defence of the core (Lakatos, 1978, p. 179). Lakatos describes research programmes as being either 'progressive' or 'degenerating'. Progressive research programmes need to show progress in two areas: they need to be theoretically and empirically progressive. Theoretical progress requires that "each new modification leads to new unexpected predictions" and empirical progress requires that "at least some of

these novel predictions are corroborated" (Lakatos, 1978, pp. 178-179). A research programme is thus progressive if and only if:

i) it makes predictions
ii) evidence is discovered that supports those predictions.

Further, the predictions of progressive research programmes should derive from the theory itself – they should 'fall out' of the theory. This means that the prediction should have the following form: if my theory is correct, then we would expect to find *x* (a phenomenon should occur, an entity should be found, etc.), and *x* would be unlikely to occur if my theory were incorrect. It is possible that some 'predictions' may not temporally succeed the theory, so long as the same piece of evidence is not used both in theory construction and in support of the theory (Worrall, 1978, pp. 48-49). Those instances in which the 'predictions' do not temporally succeed the discovery of the corroborating evidence will not be strictly 'predictive'; rather the claim is that this is something that the theory *would* have predicted had the theory preceded the corroborating evidence. What Worrall is really precluding here is scenarios in which a new piece of evidence is uncovered, that piece of evidence is used in the development of the theory, and then that same piece of evidence is used to "corroborate" the theory.

A research programme becomes degenerative when modifications are made to the theory in order to explain some unexpected phenomenon and the modified theory fails to predict anything novel. (Lakatos, 1978, pp. 179, 185). A theory might also degenerate if no evidence is found to corroborate predictions that are made, or if evidence is found that contradicts a prediction made by the theory. Importantly for the project of this chapter, one research programme supersedes another when it "has excess truth content over its rival, in the sense that it predicts progressively all that its rival truly predicts and some more besides" (Lakatos, 1978, p. 179). The question is therefore whether the HIV research programme was able to explain a substantial portion of that which is explained by its competitors and whether it has made more true predictions.

Lakatos's approach comes with a catch, which is that it permits scientists working on a degenerating research programme to continue to do so, in the hope of turning it around. More than Lakatos's account would be needed to claim that scientists should abandon their commitment to a degenerating research programme (this becomes an issue later in this chapter). Thagard (1978) provides us with the additional conceptual resources to identify research programmes that are so far beyond the pale that no rational person should continue their commitment to them. He claims that these research programmes are not just 'degenerative', but that they are also 'pseudoscientific', where being declared a 'pseudoscience' is a more damning criticism. On his account:

> A theory or discipline which purports to be scientific is pseudoscientific if and only if:
>
> 1) it has been less progressive than alternative theories over a long period of time, and faces many unsolved problems; but
> 2) the community of practitioners makes little attempt to develop the theory towards solutions to the problems, shows no concern for attempts to evaluate the theory in relation to others, and is selective in considering confirmations and disconfirmations (Thagard, 1978, pp. 227-228)

What this amounts to is that a theory is not only degenerative, but is also pseudoscientific, if it persists in the face of an alternative progressive theory, and those who are committed to the degenerative research programme do not make much of an effort to resolve its problems. The idea of there being an alternative progressive research programme is included in Thagard's account to make sense of those scenarios in which a research programme is clearly not doing very well, and has not been progressive for some time, but there are no better alternative theories to work on. It would be unreasonable for us to expect scientists to abandon their degenerating research programmes if there is nothing more promising to focus on instead (Thagard, 1978, p. 229). Thagard uses his criteria to argue that astrology is a pseudoscience, and this is meant as a severe criticism (Thagard, 1978, p. 228). Having declared astrology to be a pseudoscience, we would think that anyone who continued a commitment to that research programme would be irrational in doing so.

Now that we have the conceptual tools to assess the competing theories in early AIDS science, the next section will provide an introduction to that history.

3.3. Early AIDS Science in the 1980s

In June 1981 five young homosexual men in Los Angeles were admitted to hospital with *Pneumocystis carinii* pneumonia (PCP) (The Lancet, 1981). This was unusual, because although the microbe that causes PCP is relatively common, it is usually dealt with fairly easily by the immune system and is therefore extremely rare. PCP is normally only seen in patients who have severely suppressed immune systems, typically as the result of some other medical treatment, such as chemotherapy or as a result of receiving an organ transplant (Zuniga & Ghaziani, 2008, p. 7). These men were previously healthy. At the same time, 26 young men were admitted with Kaposi's sarcoma (KS) – a rare form of skin cancer that typically presents in elderly male patients, or (like PCP) those suffering from immune-suppression, and it is hardly ever fatal. Five of the initial 26 KS patients also had other opportunistic infections (Lancet, 1981; Epstein, 1996, p. 46; Oppenheimer, 1988, p. 270). At the end of that year the *Lancet* issued the following statement, which represented the state of available knowledge at the time about the unusual combination of diseases:

> Currently the C.D.C. (Centers for Disease Control) in Atlanta is aware of nearly 180 cases, and the numbers are increasing at 7 – 10 a week. The epidemic seems to be largely confined to urban areas of New York State and California. Most of those patients are young white males (95% aged less than 50) and 94% of them are homosexual or bisexual. The case fatality rate, due to the effects of tumour or overwhelming infection, has been alarming 40% (Lancet, 1981)

It was clear by the end of 1981 that these cases were related. The most important unifying factor was that all of the patients exhibited immune suppression resulting from a shortage of 'helper T-cells'– the white blood cells that help to fight off infections (Epstein, 1996, p. 47). Already, at this very early stage of the epidemic, there was some demarcation between 'lifestyle' theories of the disease and aetiological accounts that posited an unidentified infectious microbial agent as the cause of this cluster of diseases:

> The CDC report zeroed in on the question of sexuality – "the fact that these patients were all homosexual" – to put forward two tentative hypotheses: that the PCP outbreak was associated with "some aspect of homosexual lifestyle" or with "disease acquired through sexual contact [i.e. there was an unknown infectious microbial agent causing the disease]." However, "the patients did not know each other and had no known common contacts or knowledge of sexual partners who had similar illnesses"[thus making the infectious microbe theory seem initially less likely] (Epstein, 1996, p. 46). (Own commentary in square brackets.)

In May 1982 the term 'AIDS' was used for the first time. The name was selected to denote the following aspects of commonality between cases:

> Acquired Immunodeficiency Syndrome, or AIDS: "acquired" to distinguish it from congenital defects of the immune system; "immunodeficiency" to describe the underlying problem, the deterioration of immune system functioning (and specifically, a decline in the number of the helper T cells, causing the body to lose most of its capacity to ward off infection); and "syndrome" to indicate that it was not a disease itself, but rather was marked by the presence of some other, relatively uncommon disease or infection (like PCP or Kaposi's sarcoma), "occurring in a person with no known cause for diminished resistance to that disease." (Epstein, 1996, p. 55)

The hunt was on to find the "cause for diminished resistance to that disease". In the early 1980s there were multiple research programmes that attempted to explain the aetiology of AIDS. One early hypothesis was that AIDS is caused by the use of nitrate inhalants, which were already suspected of being immune-suppressants and which had become associated with drug use in the urban American 'gay scene' as 'poppers' (Spekowitz, 2001). However, nitrate inhalants were already an established treatment for cardiac patients, and had never been associated with causing PCP or Kaposi's sarcoma in those cases, so more would need to be said about the difference between the effects of nitrates in cardiac patients and the recreational use of nitrates to explain why the latter caused damage to helper T cells while the former did not (Epstein, 1996, p. 47).

Another early hypothesis was that homosexual men's repeated exposure to other men's semen could result in a condition similar to 'graft versus host disease.' Graft versus host disease occurs when the white blood cells of transplanted tissue identify the host's body as a foreign threat and attack it. One

of the sites of attack can be the host's immune system. Considering that graft versus host disease does sometimes occur as a result of blood transfusions, and is therefore possible in instances of fluid transfer, the 'semen exposure' hypothesis at least had a plausible mechanism (Spekowitz, 2001; Mavligit, Talpaz, Hsia, *et al*, 1984). But this theory faced obvious problems; such as why would this only be a problem for homosexual men exposed to semen, and not to heterosexual women who are similarly exposed? This theory seems to be a product of the disease initially being framed as a 'gay disease' (more will be said about the initial framing of the disease later).

The initial state of AIDS science was typified by multiple disparate hypotheses. However, in the 1982- 1983 period, the scientific community began to cluster around two main research programmes (which had already been gestured towards in the first CDC report in 1981): the immune overload research programme and the viral research programme.

At the start of 1982 the immune overload research programme was the favoured approach.[25] As the name suggests, the immune overload hypothesis proposed that AIDS patients were susceptible to opportunistic infections because their immune systems had been 'worn out' from having to cope with multiple infections, mostly in the form of Sexually Transmitted Diseases (STDs). However, this alone was insufficient to explain the sudden onset of AIDS. Homosexuality was not new – Durack (1981) asks: "were the homosexual contemporaries of Plato, Michelangelo, and Oscar Wilde subject to the risk of dying from opportunistic infections?" (p. 1466) – and neither was the existence of STDs. Two additional hypotheses were thus included to make sense of what was so different now to give rise to AIDS. The first was the use of recreational drugs, and the second was the 'sexual revolution' of the 1970s and 1980s. A little more will be said on each of these.

[25] However, Epstein (1996) argues that it was never the favoured approach within the mainstream scientific community, where the hope had always been to isolate an infectious agent. Rather, he suggests, the immune overload theory was favoured by policy makers, the media, and members of the public, because it cohered with pre-existing homophobia within those groups. That the "gay lifestyle" was making people sick tied in with people's pre-exising beliefs about the American gay community(p.52).

The literature that focuses on drugs as the relevant difference-maker most commonly refers to the recreational use of nitrate inhalants, but references to other drugs, such as heroin and cocaine, are also included (see for instance Duesberg, 1992, p. 237). As noted above, nitrates were already suspected to be immune-suppressants, so the additional strain that they placed on immune systems that had already been stressed by multiple STDs might explain why AIDS patients were so much more susceptible to opportunistic infections, and could explain what had changed to bring about AIDS over and above pre-existing STDs (Epstein, 1996, pp. 45-48; Oppenheimer, 1988, p. 273; Duesberg, Koehnlen, & Rasnick, 2003, p. 386). Versions of this hypothesis place varying degrees of emphasis on the contributory force of drugs. One example of a hypothesis that puts a great deal of emphasis on drug use in its explanation is Artie Felson's research:

> [H]e had interviewed between 300 and 400 gay men with 'AIDS', and had interrogated each of them with regard to sex and drug use. Though none of his respondents were virgins, some of them had not been especially 'promiscuous.' However, they were all drug users... And without a single exception, they had all used poppers (Shenton, 1998, p. 23).

Felson's research gave drugs a substantial role in explaining AIDS. Similarly, Duesberg claims that "Not even one male homosexual at behavioural risk for AIDS or with AIDS was found to be drug-free by the CDC" (2003, p. 387).

Other versions of the immune overload research programme de-emphasise the role of drugs, and focus more on sex and STDs in the development of AIDS. They claim that what had changed was not necessarily the use of drugs, but that poppers were part of the broader 'sexual revolution' of the 1970s and 1980s (Shenton, 1998, p. 20). On this account, it is really the unprecedented number of sexual partners and associated STDs that are the problem. Regardless of the extent to which drugs and STDs are causally weighted, the hard core of this research programme is that AIDS is caused by the immune system becoming overloaded, which renders it incapable of fighting off opportunistic infections.

Even though the very first CDC report on AIDS suggested that an infectious microbial agent might be responsible, the possibility of AIDS being virally based only become a real alternative to the immune overload research programme in July 1982 when the first cases of AIDS in haemophiliacs were reported. In each case the patient had received 'Factor VIII', a blood product made from the plasma of multiple donors, and so it seemed that there was something in the blood causing AIDS. Some kind of transmission 'agent' was present. The plasma was screened before being converted into Factor VIII, so any bacteria and typical contaminants would have been noticed. However, it was possible that something as small as a virus could have been missed in the screening process (Epstein, 1996, p. 56). However, without having actually isolated a virus, this suggestion was initially treated with scepticism.

In May 1983 the first announcement that such a virus had been isolated was made by Luc Montagnier, who was leading a virology research team at the Pasteur Institute in Paris. They named the isolated retrovirus 'Lymphadenopathy-Associated Virus' (LAV) and shared their samples with the American CDC. In April 1984 it was announced that a research team at the American National Cancer Institute, led by Robert Gallo, had isolated the viral causal agent of AIDS, which they named 'Human T-Lymphotropic Virus Type III' or HTLV-III. It was later found that LAV and HTLV-III are the same virus, and that Gallo had probably used the samples that Montagnier shared with the American CDC.[26] In 1986 the name 'Human Immunodeficiency Virus' (HIV) was settled on to denote the viral causal agent of AIDS, and this ended the dispute over whether Montagnier or Gallo's name would stick (Zuniga & Ghaziani, 2008, p. 11).

From Gallo's announcement in 1984 onward, the virological research programme became dominant and shortly after the immune overload thesis was

[26] As an interesting aside, Gallo appeared before the Department of Health and Human Services' Office of Research Integrity on charges that he intentionally suppressed the fact that he had used Montagnier's samples in order to take credit for the research findings. He was initially found guilty and the disciplinary committee stated that "his false statement had 'impeded potential AIDS research progress' by diverting scientists from potentially fruitful work with the French researchers." However, on appeal the charges were dropped because it could not be proven that Gallo had 'intentionally' suppressed his French sample sources (Shenton, 1998, pp. 29-30).

abandoned by mainstream science, only to be taken over by the AIDS denialists in 1987.

3.4. Enter the AIDS denialists

In 1984, shortly after the official announcement of the isolation of the virus, the first 'AIDS denialist' paper was published. This was 'The Group-Fantasy Origins of AIDS' by Casper Schmidt. In it he argues that AIDS as a disease only exists in so far as is a form of "epidemic hysteria", whereby a vicious group fantasy within the population of the United States is so compelling that it gives rise to what *seems* like a physical disease. He argues that it is particularly telling that groups who are the most at risk of AIDS are homosexuals and drug-users – the most ostracised members of American society in the 1980s. The view did not gain much traction, and it is completely unlike the denialist views that follow, but it does mark the first instance of a "denialist" paper being published after the isolation the HIV virus. Schmidt died of AIDS in 1994 (Kalichman, 2009, p. 26).

In 1987, Peter Duesberg enters the denialist debate, by publishing a paper entitled 'Retroviruses as Carcinogens and Pathogens: Expectations and Reality' in *Cancer Research*. This is taken to be the starting point of the AIDS denialist movement (Steinberg, 2009, p. 33; Shenton, 2011; Lenzer, 2008). In it, he argues that retroviruses cannot cause diseases. His direct target at the time was cancer-causing retroviruses. He had isolated the first cancer-causing gene in a retrovirus in 1970 with Peter Vogt, which prompted an enormously fruitful research programme isolating further "oncogenes". In Lenzer's (2008) profile piece on Duesberg in *Discovery Magazine* she describes this as "a celebrated breakthrough that truly put the young German [Duesberg] on the map". It is unclear why Duesberg did an about-turn on his views on cancer, but given that he came to hold the belief that retroviruses cannot cause diseases, this applied to AIDS as well, because HIV is also a retrovirus.

It should be noted that Duesberg has never done research specifically focussed on AIDS, only on cancer (Steinberg, 2009, p. 33). He did submit a grant

application to the National Institute of Drug Abuse, a subsidiary body of the NIH (National Institute of Health) in 1993, to test his theory that recreational drugs could cause AIDS, but it was unsuccessful on the initial application and again on resubmission.

Duesberg and his associates claim that his NIH applications were rejected because he dared to question mainstream AIDS science (Lange, 1995). On the other hand, the review committee claims that a large motivating factor for refusing to fund his application was that in the initial application, and again in his later resubmission, Duesberg failed to produce evidence that his laboratory had conducted any preliminary studies. The NIH takes preliminary studies very seriously, because they indicate whether the grant-seeking laboratory has the capacity to carry out the research that they propose. In the case of Duesberg, this was taken especially seriously, because he "is not a toxicologist or immunologist, the specializations that would be most well-equipped to perform studies of drugs and the immune system" (Kalichman, 2009, p. 47) and so he would need to be able to provide strong evidence that he could perform research so far outside of his area of expertise. Further, the committee takes a hard line on those who fail to respond to criticism from reviewers in the first round of applications:

> Revised grant applications that are not responsive to the first-round comments are extremely unlikely to move forward. The reviewers were very clear that Duesberg would need to do preliminary studies to demonstrate the promise of his ideas, a requirement of all grants like the type he applied for (Kalichman, 2009, p. 47).

He showed no evidence that he had paid attention to the first round of criticism from reviewers, or that he had attempted to conduct the required preliminary studies, and so his application was dismissed, and he was asked not to re-apply for the same topic (Kalichman, 2009, p. 48).

In Duesberg's book *Inventing the AIDS Virus* (1996) he provides a list of things that do cause AIDS, given that he believes that HIV cannot. These include: recreational drugs (especially poppers – i.e. nitrate inhalants) (p.260-284); blood transfusions in the case of haemophiliacs (p.259); and AZT (the earliest

antiretroviral therapy) (p. 299-339). When it comes to Africa, he argues that the African AIDS epidemic does not exist, and that there are not appropriate numbers of graves in Africa for all of the people who have apparently died of the disease (which is eerily reminiscent of the Holocaust denialists (Kalichman, 2009, p. 9)), and that there are financial incentives for medical authorities in African countries to fake an epidemic (Duesberg, 1996, pp. 290-291). He also suggests that in African countries there are other pre-existing diseases that simply get misdiagnosed: where the African AIDS epidemic is not faked, it is merely a product of misidentification (Duesberg, 1996, pp. 293-297). He later amends his view on Africa, arguing not that the epidemic is fake, or merely a matter of incorrectly diagnosing pre-existing diseases, but that malnutrition diminishes immune function and that is what causes AIDS (Duesberg, Koehnlen, & Rasnick, 2003).

Overall, Duesberg's position is that where AIDS exists at all, it is the product of aspects of poor lifestyle that place strain on the immune system, the one exception being haemophiliacs who receive blood transfusions. He accepts that HIV exists, but argues that it is just another opportunistic infection that AIDS patients are susceptible to as a result of already diminished immune function. On Duesberg's account, HIV is a harmless passenger virus (Duesberg, interviewed in Shenton, 2011).

Duesberg is the most credible scientist associated with the denialist community, given his outstanding track record in cancer research prior to his about-turn on the role of retroviruses in causing disease (Larvor, 1998; Steinberg, 2009). But the denialist community is large and diverse. For instance, the Perth Group, led by Eleni Papadopulos-Eleopulos (whose highest qualification is an undergraduate degree in nuclear physics), is a group in Australia who argue that the virus does not exist at all (Steinberg, 2009, p. 36; Shenton, 2011). The Perth Group is so committed to this position that they have offered a $20, 000 cash prize to anyone who can prove the existence of HIV. In a strange twist of irony, Duesberg came forward claiming that he has evidence that the virus exists (which is consistent with his position that HIV is an actual virus, but that it is an

effect of AIDS rather than its cause). He was not awarded the prize money (Kalichman, 2009, p. 61). The Perth's Group position will not be considered further in this chapter, because their views are so fringe and their leadership so under-qualified that is not worth taking their position seriously.

Also in the diverse camp of AIDS denialists is Henry Bauer – who initially looks like a promising candidate as a serious scientist committed to the denialist movement, due to his position as Professor Emeritus of Chemistry and Science Studies and Dean Emeritus of Arts and Sciences at the Virginia Polytechnic Institute and State University. But on closer inspection, he begins to looks more suspicious. He is also a leading authority on the Loch Ness Monster, and editor of the *Journal of Scientific Exploration* – "the major outlet for studies on UFOlogy, paranormal activity, extrasensory powers, alien abductions, etc." (Kalichman, 2009, p. 71). Even if we bracket our concerns about his non-AIDS related research activities, his position in the AIDS denialist debate is not promising. His stance is that HIV does not cause AIDS, and he attempts to use statistical methods to disprove the causal connection. There are two immediate problems with Bauer's approach. The first is that he bases his analysis on testing data from US military recruits from the 1980s – the idea being that this is a representative sample of young Americans at the time. But US military recruits could not be less representative of the sub-groups who were most affected by AIDS in the 1980s, especially because gay men and injection drug users (the two groups who were most at risk of the disease at that point) were explicitly excluded from the military. The other immediate problem is that he runs his analysis of HIV in comparison to gonorrhoea and syphilis. The rationale is that if HIV were contagious its spread pattern would look like that of gonorrhoea and/or syphilis, both of which are known to be infectious and sexually transmitted. HIV's epidemic pattern looks nothing like these diseases and so it cannot be a sexually transmitted infection (on his account). However, both gonorrhoea and syphilis are bacterial, not viral, and so it is expected that they will have entirely different spread profiles to HIV (Kalichman, 2009, pp. 72-73). His position will also not be given any further consideration in this chapter, because his methods are so obviously incorrect.

Having excluded the Perth Group's position that there is no virus, and Bauer's particular version of the claim that "HIV does not cause AIDS", on the grounds that neither is credible, the rest of this chapter focuses on Duesberg's approach to AIDS denialism. Largely because of Duesberg's track-record as an extremely well-respected scientist:

> [He] was part of the team that first mapped the genetic structure of retroviruses, [he was] codiscoverer of the first viral cancer gene in 1970… In 1986, at age 49, he was elected to the National Academy of Sciences. That same year he was given a National Institutes of Health Outstanding Investigator Award, one of the most prestigious and coveted grants. Robert Gallo, codiscoverer of HIV and a former friend of Duesberg's, praised him in 1985 (sic) as a "man of extraordinary energy, unusual honesty, enormous sense of humor, and a rare critical sense." He added, "This critical sense often makes us look twice, then a third time, at a conclusion many of us believed to be foregone." (Lenzer, 2008)

Lenzer gets the year wrong when quoting Gallo, who made these comments about Duesberg in 1984 at the National Cancer Institute in Maryland, not in 1985 (this is due to a mix up between the year in which the speech was given, and the year when the speech was published). The year is relevant because Gallo's lengthy speech in praise of Duesberg was made shortly before the discovery of HIV was announced in the US, and thus just before Gallo and Duesberg's collegial relationship began to break down. Lenzer also cuts Gallo's quote off short, leaving out much that is prescient, given the role that Duesberg would shortly play in the AIDS debate. Gallo's speech goes on as follows:

> However, his [Duesberg's] critiques are sometimes a major problem for the casual observer. When is he truly debating? When is he being the devil himself? The casual observer is also often at a loss to determine which of the many weapons he possesses he is using. Peter, it is hard for us to know when you are using your machine gun or your slingshot, or simply exercising your vocal cord (Gallo, 1985, p. 8).

The first page of Gallo's introduction for Duesberg is nothing but glowing praise for his outstanding research career. If we should pay attention to anyone in the denialist camp, it should be Duesberg, even if based on his credentials alone. Additionally, his view should be taken more seriously than the others because it

is continuous with immune overload thesis, which was a leading contender to explain the aetiology of AIDS in the early 1980s, and so it has not always been a fringe view – it is not inconceivable that there is something we missed when shifting to the strictly viral account. Although Duesberg was not one of the scientists working on the immune overload theory of AIDS in the 1980s, it is clear that he intends for his position to be an extension of that approach. In the abstract of his paper, 'The chemical bases of the various AIDS epidemics' (Duesberg, Koehnlen, & Rasnick, 2003), he explicitly states that he intends for his own approach to be a continuation of the immune overload theory, which he believes was abandoned too early in the wake of premature enthusiasm for the viral account. The rest of this chapter will therefore compare the immune overload theory with the viral account.

3.5. Assessing the Immune Overload and Viral Research Programmes

This section assesses the immune overload theory of AIDS in comparison to the viral account, using Lakatos's theory of scientific research programmes, as outlined earlier in this chapter. The aim is to determine whether the viral research programme was accepted rationally over the immune overload account in 1984, when the formal announcement was made in the United States that HIV had been isolated, and when the immune overload theory was abandoned by mainstream science. I begin by assessing the viral account, because elements of that account are relevant to addressing questions that are later posed by Duesberg.

3.5.1 Appraising the Viral Account of AIDS

Lakatos's theory requires that a research programme generate novel predictions, and that at least some of those predictions be corroborated in order for it to be considered progressive. This section will look at two such predictions generated by the viral account of AIDS: 1) that the virus would be found in the T-cells of all AIDS patients; and 2) that treatment targeted at various parts of the virus

(antiretroviral therapy) would be effective. It will be found that both predictions were eventually corroborated, but that process took some time.

3.5.1.1. If HIV causes AIDS then the virus will be in the T-cells of AIDS patients

The viral theory of AIDS operates on the premise that the virus kills off helper T-cells, which are necessary for the immune system to function. The immune system is thus weakened and the patient becomes susceptible to opportunistic infections. It would therefore be reasonable for the viral account to predict that the HIV virus will be found in the T-cells of AIDS patients.

It initially seemed as though this prediction would not be corroborated, and this was one of the core concerns raised about the theory at the Proceedings of the National Academy of Sciences in 1985. When T-cell samples taken from AIDS patients were tested, very few were found to contain the virus, "sometimes as few as one in one hundred thousand cells" (Epstein, 1996, p. 90). This was a real worry to the scientific community at the time and opponents to the viral research programme used this as their 'killer argument'. However, on Lakatos's account it is permissible for there to be problems at the start of a research programme (Lakatos, 1978, p. 5). Eventually this concern was resolved as the tests become more sophisticated, and the virus was found to be present in a significant number of the T-cells in AIDS patients. The change came with the development of polymerase chain reactions (PCR), a procedure that allowed for DNA to be manipulated, thus making it easier to spot the virus (Epstein, 1996, p. 120). In 1993 an even more refined version of PCR was developed and using this method it was shown that 10% -30% of the T-cells had the virus, and within those cells the viral concentration was high (Epstein, 1996, p. 163). So the mystery of T-cells was eventually resolved. However, this corroboration only came in 1993 – almost a decade after the viral research programme was actually accepted by the scientific community. This might add credibility to the denialists' claim that the viral account was accepted *prematurely*.

Related to the prediction that the HIV virus would be found in the T-cells of AIDS patients, is the prediction that if more of the virus is found in any particular patient then we would expect the disease to be more severe in that patient (this is typically referred to as the 'dose-response relationship'). Again, this prediction was corroborated (albeit somewhat later than we might like). In 1996 it was shown that higher concentrations of the virus in the patient's plasma (i.e. higher 'viral loads') are correlated with more rapid onset of the disease and more rapid progression to death (Mellors, Rinaldo, Gupta, White, & Todd, 1996). This supported the prediction that increased levels of the virus in patients would result in the illness being more severe. However, again, this confirmation only became available a significant period of time after the viral account was accepted in 1984 (roughly 12 years later).

3.5.1.2 If HIV causes AIDS then ARV treatment will be effective

The second main prediction to come out of the viral account of AIDS is that if HIV causes AIDS, then antiretrovirals (which target specific viral processes) would be effective, and we would see AIDS patients developing fewer opportunistic infections, and living longer and healthier lives.

Prior to the isolation of the virus, treatment had focussed on targeting the opportunistic infections that AIDS patients developed as a result of their diminished immune function. For instance, if the patient developed a fungal infection as a result of AIDS, they would receive a standard anti-fungal treatment. This approach allowed for specific opportunistic infections to be dealt with, but it did not prevent subsequent infections from taking hold, and physicians were only able to extend the lives of AIDS patients by short periods of time this way (Zuniga & Ghaziani, 2008, p. 18). Ultimately, the patient's immune system would collapse, and the opportunistic infections would become too many and too severe to treat. It was recognised that little could be done to help patients in a significant way until the underlying immune dysfunction was targeted directly, but at that point the cause of the underlying immune dysfunction was unknown. When the HIV virus was isolated, the possibility was opened up that manipulations could

be made on the virus directly and that the underlying immune suppression could thus be remedied (Epstein, 1996, pp. 181-182).

Initially, very little was known about the isolated virus, and so it was difficult to target the treatment at specific viral processes. But there was some evidence that HIV was a retrovirus, as opposed to a standard virus. Standard viruses are composed of DNA; and via the process of infection they enter the nucleus of a new cell and make that cell replicate the virus as per the instructions coded in its DNA. A retrovirus is composed of RNA, rather than DNA, and so before it can integrate itself into the newly infected cell, it needs to convert its RNA into DNA by rewriting its genetic code backwards, which it does through a process called 'reverse transcription'. In order for reverse transcription to take place, an enzyme called 'reverse transcriptase' needs to be present. The first thought that scientists had about how to treat the disease was to find a way of inhibiting reverse transcriptase, thus preventing the reverse transcription process from occurring, and saving further cells from being infected (Epstein, 1996, p. 183; Pomerantz & Horn, 2003, p. 867).

The first generation of ARV treatments were thus transcriptase inhibitors – also known as NRTI's (dideoxynucleoside reverse transcriptase inhibitors) (Vella, Schwartlander, Sow, Eholie, & Murphy, 2012). In 1985, trials were underway for a transcriptase inhibitor called 'suramin', and while it did seem to inhibit the reverse transcription process, its side effects were severe. Some of the patients suffered adrenal failure and died more quickly than they would have been expected to without treatment (Epstein, 1996, p. 191).

Those who were suspicious of the early acceptance of the viral theory interpreted the initial problems with ARVs as evidence that HIV does not cause AIDS. Sonnabend, for instance, commented:

> If we have agents that effectively inhibit the replication of this virus but [those agents] make no impact on the course of this disease, I think it will make apparent, for some people, the actual role of HLTV-III [HIV prior to the name was officially settled] in causing this disease. (Sonnabend, 1985 quoted in Epstein, 1996, p. 191)

That is, according to Sonnabend, HLTV-III has no role in causing the disease. Criticisms, such as Sonnabend's were not entirely fair, given that transcriptase inhibitors were actually dampening the reverse transcription process, and were thus inhibiting the replication of the virus (as predicted). Their side effects were just so overwhelmingly bad that they could not yet be used as an effective treatment.

Things began to look more promising in July 1985, when azidothymidine (AZT), a new generation of reverse transcription inhibitors, was tested in a small Phase I trial. AZT is a 'nucleoside mimic', where nucleosides are the building blocks of DNA:

> AZT "fooled" the reverse transcriptase enzyme into using it, in place of the nucleoside it imitated, when transcribing the virus's RNA into DNA. Then, once AZT was added to the growing DNA chain, AZT's structure prevented any additional nucleosides from being added on: reverse transcription simply came to a halt at that point, and the virus stopped replicating (Epstein, 1996, p. 193).

Nineteen patients were included in the initial Phase I trial. After six weeks it was found that AZT had prevented the virus from replicating in fifteen of the nineteen patients, and that as such they had improved immune function and they experienced fewer opportunistic infections (Epstein, 1996, p. 193).

In 1986 a Phase II trial for AZT began. It involved 282 AIDS patients. 145 patients were given AZT and the remaining 137 received a placebo. Six months into the trial the results were that: "only one patient who had received AZT had died, compared with 19 patients in the control group. Twenty-four patients in the AZT experimental arm had developed OIs [Opportunistic Infections], compared to 45 placebo recipients" (Zuniga & Ghaziani, 2008, p. 19; Fischi, et al., 1987). The trial was shut down prematurely because regulators had ethical concerns about continuing to administer placebos to the control group, given the early success of the treatment (Zuniga & Ghaziani, 2008, p. 19; Gazzard, 2005; Epstein, 1996, p. 198). There has been speculation from the denialist community over the 'real reasons' for the study's closure. Celia Faber, a journalist and AIDS denialist, published an article in 2006 in which she claimed that the trial ended because

the double-blinding process had been compromised due to drug sharing between study participants. However, this conclusion is not supported by the FDA (Food and Drug Administration) report on the issue (Geffen, 2010, pp. 33-34). In 1987 the FDA approved AZT as safe for use (Pomerantz & Horn, 2003, p. 867).

Shortly after the FDA accepted AZT, three further transcriptase inhibitors were approved. However, the initial enthusiasm over transcriptase inhibitors was short-lived, because "each had its particular toxicities", thus making the side effects of these drugs severe (Vella, Schwartlander, Sow, Eholie, & Murphy, 2012, p. 1231). The side effects of these early drugs were so bad that many argued that they were making patients more ill than if they had not received any treatment (Epstein, 1996, p. 149).

Part of what was going wrong with the early use of ARVs was that it was unclear how they should be administered: What should the dosage be? How often should it be administered? At what point in the disease's progression should the treatment begin? These were all questions that had yet to be resolved, and very often practitioners opted for precautionary over-dosing, leading to even more severe side effects. A doctor working at the 'frontlines' of AIDS treatment in New York in the mid-1980s describes the situation at the time as follows:

> [It was] really scary. We had no clue what we were doing, and in retrospect we made a lot of mistakes… We used medication in the wrong doses… We didn't anticipate things… You just didn't know what the heck was going on (Wafaa El-Sadar quoted in Zuniga & Ghaziani, 2008, p. 23).

So some of what was going wrong with early ARV treatment was down to dosage problems – particularly overdosing leading to more severe side effects. It also quickly became clear that 'mono-therapies' – that is, administering only one drug – are only effective for short periods of time before the virus adapts and the treatment no longer works (Pomerantz & Horn, 2003, p. 867). Attempts were made to combine the various approved reverse transcriptase inhibiters, and to administer them sequentially, but this only bought patients marginally more time (Vella, Schwartlander, Sow, Eholie, & Murphy, 2012, p. 1232). Despite these problems, it was always clear that even early transcriptase inhibiters were

blocking the reverse transcription process, thus corroborating the viral account of AIDS. Their positive effects were just being swamped by the extremely bad negative side effects. Some AIDS denialists claim that ARVs cause AIDS and when they do, they always make reference to how 'toxic' AZT in particular is (Shenton, 2011; Duesberg, 1996, pp. 300-301; Duesberg, Koehnlen, & Rasnick, 2003, p. 393). This may be a reference to the severe side effects that were associated with early treatment, but note that AZT is no longer in use today.

In 1996 everything changed for ARV therapy. The FDA approved two additional ARV classes. The first of these were nucleoside reverse transcriptase inhibitors (NNRTIs), which were also focussed on inhibiting reverse transcription, but did so at a different point in the process to what had been the target of previous treatments. The second were protease inhibitors, which suppressed the enzyme that was required for the virus to reach maturation (Pomerantz & Horn, 2003, p. 867; Vella, Schwartlander, Sow, Eholie, & Murphy, 2012, p. 1232). These treatments were offered to patients as a combination therapy (that is, multiple treatments were mixed together into a 'cocktail' – mono-therapies, which were known to lead to drug resistance, were no longer in use). This became known as 'HAART' – Highly Active Antiretroviral Therapy – and it was extremely effective. Steinberg describes HAART as having a 'Lazarus effect', whereby "AIDS patients who had been mortally ill were rising from their beds, putting on their jackets and ties, and reporting for work" (Steinberg, 2009, p. 34). In the US and Western Europe there was a "60-80% reduction in new AIDS-defining conditions [opportunistic infections], hospitalizations and deaths" (Zuniga & Ghaziani, 2008, p. 24). From 1996 onward, ARV treatment was an unequivocal success.

The effectiveness of ARV therapy marks the viral research programme as progressive in Lakatos's terms. The theory predicted the effectiveness of the treatment, based on the virus being a retrovirus that operates in very particular ways (making use of a reverse transcription process, and the enzyme protease being required for the virus to reach full maturity). ARV therapy targeted these very specific processes by disrupting the function of particular mechanisms and enzymes. Had those mechanisms and enzymes not been present, the treatment

would not have been effective. However, again, this only took place in 1996 when HAART became available, and so it took some time before this aspect of the theory was corroborated.

The response from AIDS denialists to the resounding success of ARVs shows that their engagement with the debate may be somewhat disingenuous. Duesberg's explanation for the seeming success of ARV therapy is that he claims that ARVs have an antibiotic effect that treats opportunistic infections and so makes it appear as though the patient is getting better, even though the long-term damage to the immune system persists. However, as just noted, the mechanisms by which ARVs operate are very specifically targeted to processes of the HIV virus, and so it would only be effective if the cause of disease were HIV.

The viral research programme is thus progressive. It predicted that the virus would be found in the T-cells of AIDS patients, that greater levels of the virus would be associated with more severe cases of the disease, and that ARV therapy (which targeted very particular aspects of the HIV mechanism) would be effective. All of these predictions were corroborated. However, they were only corroborated in the early to mid-1990s. The AIDS denialists might thus be right that the viral theory of AIDS superseded other competing theories prematurely in the 1980s, but it was definitely a progressive research programme by the time Mbeki was elected President of South Africa in 1999, which is what really matters for the overall project of this thesis. The following section will assess the immune overload theory.

3.5.2. Appraising the Immune Overload Account

The immune overload theory was introduced with the first reported AIDS cases of the early 1980s. All of the initially reported cases were of gay men in the US, and it was assumed that there was something about their "lifestyle" that was making them ill. None of the early patients knew each other, and so it seemed unlikely that they were all experiencing a single underlying infection, because it was assumed that they would need to have come into contact with each other to

pass on an infectious microbial agent. Epstein explains that the immune overload theory: "represented the initial frame for understanding the epidemic: the syndrome was essentially linked to gay men, specifically to the 'excesses' of the 'homosexual lifestyle.'" (Epstein, 1996, p. 48). Fee and Krieger agree with Epstein as follows:

> Epidemiologists, the first scientists to lay claim to understanding the mysterious new ailment, were struck by its seemingly exotic preference for young, homosexual men; they therefore searched for causes in the behaviors or "life-styles" common to gay men. In the process, they looked for risk factors prevalent in this "risk group" and indicted life in the fast lane, including "promiscuity," "poppers" (amyl nitrate), and anal sex (Fee & Krieger, 1993, p. 1478)

Framing the disease in this way also cohered with the social stigma attached to the group at the time. So strong was the early commitment that the disease must be linked to 'homosexual lifestyle', that initial counter-examples were dismissed as either lies on the part of patients, or as instances of an entirely different disease. Male patients who presented with AIDS but claimed not to be gay were accused of lying about their sexuality – "He says he is not homosexual, but he must be" (Epstein, 1996, p. 50). Further, when a New York paediatrician reported that he was seeing cases of AIDS amongst children in 1981, he was mocked for thinking that the disease observed in children could possibly be the same as that being reported by gay men (Epstein, 1996, p. 50). The initial version of the immune overload theory was deeply committed to the idea that it was a combination of extreme exposure to STDs and drugs that caused the immune system to break down and increased individuals' sensitivity to opportunistic infections (Shenton, 1998).

As early as August 1981, AIDS cases were reported in heterosexual patients (including women) and in some children (Oppenheimer, 1988, p. 279; Epstein, 1996, pp. 49-50). In 1982, cases of AIDS were discovered in haemophiliacs and Haitians living in the United States, and both were added to the list of 'risk groups' (Epstein, 1996, p. 59). The way that the immune overload theory coped with the disease appearing in additional groups was to expand the list of 'lifestyle' factors that could cause the disease to include blood transfusions, poverty, and even taking antiretrovirals (especially AZT). Cases of children

developing AIDS, once it was accepted that this was a possibility, were explained by poor lifestyle choices on the part of their mothers while pregnant. For every new group of people who presented with the disease, a new cause of AIDS was added to the ever-increasing list. The ad hoc way in which this list of AIDS causing factors was expanded is illustrated well by the following quote by Kalichman:

> [I]llicit drugs cause the immune dysfunction labelled AIDS. What about the people with AIDS who do not use illicit drugs? His [Duesberg's] reply is that AZT and other chemicals used to treat HIV actually cause AIDS. And what of those people who had AIDS before AZT was approved and had never used drugs? He says there are no such people. What about haemophiliacs and blood transfusion recipients who have never used drugs or been treated with AZT? Duesberg claims that haemophiliacs develop AIDS from contaminated blood clotting factors. What about Africa, where AIDS kills millions and there is no widespread drug use or AZT? Duesberg claims that AIDS in Africa has existed long before there was an HIV test, and that many old diseases result from malnutrition, contaminated drinking water, and poor sanitation; in a word, poverty… What cause of AIDS we are talking about depends on where you live as well as your lifestyle? For Gay men, drug use causes AIDS. For Gay men who do not use drugs, HIV medications cause AIDS. In Africa, malnutrition causes AIDS. If you are a wealthy African, AZT causes AIDS. If you are a Haemophiliac, treatments for haemophilia cause AIDS. No research has ever suggested the Duesbergian view of AIDS is true (Kalichman, 2009, pp. 38-39).

This shows that the immune overload theory adapted to account for new anomalous evidence (the expansion of the disease into new groups). This is permissible on Lakatos's account, so long as the adjustments yield novel predictions, and the predictions are corroborated. But we do not see such predictions and corroborations developing out of the adjustments to the immune overload thesis. One prediction that we might expect to come out of the immune overload thesis is that as ARVs become increasingly available, so we should witness more AIDS deaths, given that the account claims that one of the causes of AIDS are ARVs. However, "AIDS declines everywhere in the world when antiretroviral medications become available" (Kalichman, 2009, p. 43).

The immune overload research programme also failed to 'predict' the cases of gay men with AIDS who do not conform to the 'risk profile' suggested by the

research programme. For instance, not particularly 'promiscuous' (and so less likely to have significant medical histories of repeated STD infections), non-drug using homosexual men were also found to have AIDS and these cases were reported by doctors right from the outset of the disease (Epstein, 1996, p. 49). These are anomalous cases that the theory cannot account for, except perhaps to tell us that these individuals are lying when reporting on their lifestyles – which is one of those ad hoc moves that Lakatos warns against (adapting the theory to account for anomalous evidence, and that adaptation not leading to any novel predictions).

As mentioned in the introduction, much of the project of those who remain committed to the immune overload theory of AIDS is about raising questions of mainstream AIDS science, as opposed to putting forward a substantial positive theory of the disease. Having covered the positive account that is on offer, I now move on to assessing two of the main questions that they pose to those who endorse the viral account of AIDS. First, if HIV causes AIDS, why is there still no vaccine? Second, if HIV causes AIDS, why is there such a long latency period between contracting the virus and the onset of illness?

In various places, Duesberg claims that if the viral account of AIDS were correct, then a vaccine would have been found by now, given how much time and money has been devoted to studying HIV (Duesberg, 1996, p. 438; Duesberg, Koehnlen, & Rasnick, 2003, p. 389; Duesberg interviewed in Shenton, 2011). However, while it is true that there is no vaccine, we know why this is the case, and it is consistent with the viral account being true. The idea behind vaccination is that if you expose the immune system to the infectious agent, under controlled conditions, it will learn to recognise that agent and develop defences against it. The major challenge is that HIV mutates rapidly (which the reader will recall was a major problem for the effectiveness of early ARV mono-therapies):

> The most salient characteristic of HIV, is its remarkable capacity for generating 10^9 - 10^{10} virions every day coupled with a high mutation rate of approximately 3×10^{-5} per nucleotide base per cycle… In fact, after 10 to 20 years of infection, the HIV in a patient's body will have mutated at a

> rate equal to the average mutation rate of human beings over the course of the past few million years (Zuniga & Ghaziani, 2008, p. 18).

This means that even if an individual were 'vaccinated' against the disease, the version of HIV that the patient would come into contact with later would be so different to the virus that they were initially inoculated with, that their immune system would not recognise it as the same virus, and would thus not identify it as a threat. Developing an HIV vaccine is therefore an extremely difficult task, perhaps impossible. That there is no HIV vaccine does not tell us anything about the plausibility of the viral account of AIDS (Zuniga & Ghaziani, 2008, p. 18).

The second criticism levelled against the viral AIDS account by denialists is that the delay between infection and disease onset is so long that HIV cannot possibly be the cause of AIDS. In Duesberg's own words:

> I believe AIDS is not, or cannot be an infectious disease. An infectious disease, believe it or not, has a certain criteria [sic] to it. How it happens, when it happens. For example, if you get infected by a bug or by a virus, within weeks or months after contact or after that infection you will have symptoms of a disease. In HIV and AIDS, however, we are told you get sick ten years later, ten years after infection. That is not how viruses or bacteria work. They work fast or never. They are a very simple mechanism like a little clock that can do only one thing – go around the dial once and that takes 24 to 48 hours with a virus. There is no way a virus could possibly slow down or wait a week or wait ten years. That is totally absurd (Duesberg, 1990 quoted in Shenton, 1998, p. 88).

However, this contradicts the position accepted by mainstream virology. Note the following extract taken from an introductory epidemiology text:

> Epidemiologic studies of kuru, the first spongioencephalopathy to be identified in humans, were initiated in the late 1950s…In 1959, the veterinarian William Haddow had spelled out the clinical and pathological similarities of kuru to scrapie, an infectious disease of sheep with a very long incubation period. Awareness of this zoonosis moved Gajdusek's team to consider the possibility of an infection rather than a genetic condition. With monumental patience in the laboratory, after many attempts Gajdusek and his colleagues eventually succeeded in transmitting the disease to primates; upon exposure, the disease became manifest only several years thereafter. Gajdusek's unique discovery led to the recognition of what was then called a slow virus… In 1976, Gajdusek's persistence was rewarded by the Nobel Prize (Susser & Stein, 2009, p. 206).

From this extract it is clear that the scientific community had accepted the existence of viral diseases with long incubation periods (also known as 'slow viruses') since at least the 1970s. The long incubation period in the case of AIDS therefore might be unusual, but it is not unheard of in medical science. Duesberg responds to being presented with this mainstream scientific view about scrapie and the existence of slow viruses by arguing that the scrapie virus does not exist. He argues that all cases of the disease can be explained by exposure to toxins, and that no slow viruses exist at all (Duesberg, 1996, pp. 76-80). It is ignoring this piece of mainstream science that makes Kalichman speculate that perhaps Duesberg is intentionally misrepresenting the mainstream scientific facts (2009, p. 50).

The immune overload theory quite clearly looks degenerative. Adaptions were made to the theory to accommodate each new group who presented with AIDS, and novel predictions did not flow from those adjustments. Further, those predictions that we might have expected from the theory were not corroborated – such as the case of ARVs. It was noted that a large portion of the AIDS denialists' project involves questioning aspects of the mainstream account of AIDS, and two of the main questions that they pose were addressed: if AIDS is viral then why is there no vaccine, and how can HIV cause AIDS after such a long latency period? It was argued that those who endorse the viral account of AIDS have the resources to respond to both of these criticisms: there is no vaccine because the virus mutates too quickly; and mainstream virology and epidemiology have accepted the existence of viruses with long latency periods since the 1970s. Duesberg resists the latter claim by arguing that this is just another piece of mainstream science that is wrong (1996, pp. 76-80). Thagard (1978), as previously discussed, argues that one of the factors that nudges a research programme toward pseudoscience is an unwillingness to engage with the available alternative theories. Deusberg's dismissal of large swathes of mainstream science seems to be indicative of the research programme not just being degenerative, but also being pseudo-scientific (more will be said about this below).

3.6. What motivates the denialists?

Given the overwhelming evidence that has been presented in favour of the viral account of AIDS, and the obvious problems with the immune overload view, why would anyone still stick with the latter?

Some portions of the denialist camp are obviously motivated by financial gain, and they have used their commitment to the immune overload theory to sell various 'immune enhancing remedies' (Kalichman, 2009, p. 21; Thom, 2009; Geffen, 2010). However, Duesberg is clearly not financially driven in this regard, and has arguably lost money as a result of his commitment to the denialist programme (Lenzer 2008).

Kalichman makes various speculations about what might be motivating Duesberg. One speculation is that he remains committed to the denialist stance because he something of a rock star within the denialist community (Kalichman quoted in Steinberg, 2009, p. 35). Another might be his peculiar rivalry with Robert Gallo (the co-discoverer of AIDS) – they were on friendly collegial terms until Duesberg began to backtrack on his views about cancer, and then Duesberg found himself increasingly excluded from the research community, even before he entered the AIDS debate. That Duesberg's animosity is linked to Gallo in particular seems likely – in the original 1987 AIDS denialist paper, Duesberg thanks Gallo in the acknowledgements for having prompted the paper (Kalichman, 2009, pp. 35-37; Lenzer, 2008). In another speculation, Kalichman wonders whether Duesberg might just be contrarian in character:

> Contrarian may be an understatement. You almost get the feeling that if suddenly it were discovered that AIDS is caused by toxic drugs Duesberg would refute the evidence and pose an alternative theory (p.48)

Kalichman goes on with his speculation that Duesberg might just be contrarian when he states that:

> Is it possible that Peter Duesberg himself did not seriously question whether HIV causes AIDS? In a position paper that he wrote for a South African meeting, he referred to the classic novel *The Plague* by Albert

> Camus as the most readable modern depiction of an epidemic. Addressing a country where hundreds of people die of AIDS each day, Duesberg's reference to a work of fiction seems quite telling. Was he signalling that we should not take him too seriously? Perhaps he pushes on his academic rivals only because he enjoys the argument – as a kind of sport, debate for the sake of debate. (p.49).

It is unlikely that we will get a definitive answer to the question of what is motivating the AIDS denialists to hold on to their views. When Duesberg is asked directly why he maintains his stance, he says "I don't want to be a 'good German" (Duesberg quoted in Lenzer, 2008), referring to his childhood in Nazi Germany and the disdain he feels for those Germans who stood back and did nothing while evil was allowed to continue around them.

3.7. Where does this leave us?

It seems clear that the HIV research programme should be accepted over the immune overload thesis, by at least the mid-1990s. The success of ARVs makes the HIV research programme both theoretically progressive (it predicted that the ARVs would work) and empirically progressive (that prediction was corroborated). It is also extremely unlikely that ARVs would have been effective, had the viral account not been true. In contrast, the immune overload account adjusted the theory to accommodate new evidence (individuals from new risk groups getting the disease), without novel predictions being made and corroborated. The immune overload theory was degenerating. Further, much of the AIDS denialist project involves asking questions of the viral account that can easily be addressed. But, we know that individuals have stuck with the immune overload theory despite all of this, what can we say about that?

The catch of making use of Lakatos's methodology is that it allows for those who support the immune overload thesis to continue being so committed, because Lakatos argues that it is permissible "to stick to a degenerating research programme and try to turn it into a progressive one" (Lakatos, 1978, p. 6). This echoes Popper's idea that "the scientific future is essentially unknowable" (Urbach, 1978, p. 100) and so it is plausible that a currently degenerating

research programme may become progressive in the future, and it is not possible to know in advance whether or not this will be the case. After all, one of the appealing characteristics of Lakatos's approach is that it allows for substantial problems to exist within a scientific research programme without requiring that it be dismissed outright, and it permits scientists to work through those issues. We cannot insist that scientists involved in degenerating research programmes should give up on them, while also maintaining the idea that the scientific community should be given time to resolve the problems with their theories. But this is an unsettling conclusion, given the obvious problems highlighted with the immune overload approach above, and the harms that have been associated with the persistence of this particular position (most notably, South African AIDS denialism).

Worrall highlights an aspect of Lakatos's theory that may provide a more nuanced way of addressing this problem. He describes Lakatos's position as follows:

> Of course the methodology does not predict that, whenever some new programme comes along which it appraises as more progressive than the old one, all scientists will switch to work on the progressive programme. Nor does the methodology pronounce 'irrational' those scientists who, in such circumstances, stick to the old programme. Such a scientist may, in perfect conformity with this methodology, agree that the new programme is, at the moment, superior, but nevertheless declare his intention to work on the old programme in an attempt to improve it so that it becomes even better than the new programme (Worrall, 1978, p. 61).

Initially, this just seems to be a restatement of the part of Lakatos's theory that creates the problem. However, in the note for this section, Worrall continues:

> A scientist *would* be pronounced 'irrational' (or rather mistaken) by the methodology if he stuck to the old programme denying that his own programme needed improvement in order to catch up with the new one. It is in such circumstances that we shall begin to suspect the operation of extra-rational motives (Worrall, 1978, p. 70).

And it is the latter behaviour that we see in AIDS denialism. It is not just that scientists continue to be committed to the immune overload research programme, despite recognising that their research programme may currently be in a rocky patch. Rather, they deny the truth of certain crucial pieces of evidence

that support the HIV thesis and ignore evidence that places their own position in an unfavourable light. For instance, the denial of evidence that suggests that ARVs are effective, or the denial that viruses can have long incubation periods. It is this kind of denial that leads to the suspicion that 'extra-rational motives' might be at play.

Thagard's (1978) contribution, as explicated earlier in this chapter, is also relevant here. His position is that a theory can be declared not just degenerative (which although undesirable, is not pejorative) but also 'pseudoscientific', if: "the community of practitioners makes little attempt to develop the theory towards solutions to the problems, shows no concern for attempts to evaluate the theory in relation to others, and is selective in considering confirmations and disconfirmations" (Thagard, 1978, pp. 227-228)

All of these factors are present in the case of the AIDS denialism. There is an alternative progressive research programme available in the form of the viral account of AIDS, which has not been seriously engaged with by those who remain committed to the immune overload theory. There has not been much of an effort to resolve the problems of the immune overload account – the one exception being Duesberg's application for a NIH grant in the early 1990s to study the effects of nitrates on immune suppression, but he failed to do the required preliminary studies and did not respond to the reviewers' questions from his first attempt at the application, so it is unclear how seriously he really took this application. Mostly, what we see from the denialist camp are restatements of very early AIDS science (Steinberg, 2009), which is indicative of the selective way in which they consider confirmations and disconfirmations (they only consider the early AIDS science supports their view, neglecting the scientific developments that followed).

We can thus conclude that the continued commitment to the immune overload theory not just degenerative, but also pseudoscientific. As such, it would be irrational for anyone to continue research in this domain.

3.8. A final worry: loss of explanatory content

A final concern one might have is that some explanatory content has been lost in HIV's acceptance as the mainstream view. Explanatory losses can occur when one research programme supersedes another, because it has excessive truth content over its rivals, but there are aspects of the phenomenon that the alternative research programme explains and the accepted view cannot (Worrall, 1978, pp. 61-62).

In the AIDS case, having access only to the viral account of AIDS, makes it more difficult to explain the disproportionate prevalence of AIDS in southern Africa, without resorting to problematic assumptions about Africans and their sexuality (which would require evidence to make convincing – more will be said about this assumption in the following chapter). When the immune overload theory, and its explanatory tools, were still available, one could appeal to factors such as poverty to explain the AIDS epidemic in southern Africa. The next chapter will try and account for this explanatory gap that was created by the shift to the strictly viral view.

3.9. Conclusion

The aim of this chapter was to gain a better understanding of the state of the scientific debate surrounding AIDS, and to take seriously the claim of contemporary AIDS denialists that they were unfairly and prematurely cut out of the debate in 1984 when HIV was declared to be the cause of AIDS.

This was done using Lakatos's theory of scientific research programmes, which requires that a theory make novel predictions, and that at least some of those predictions be corroborated in order for it to be progressive. When a theory starts to make ad hoc adjustments to accommodate unexpected evidence, and no novel predictions and corroborations follow from those adjustments, then the theory starts to degenerate. Further, Thagard tells us that a theory is not just degenerative, but also pseudoscientific, when its devotees fail to engage with

promising progressive alternative research programmes; they make little effort to solve the theory's problems; and they are selective in the way they deal with confirmations and disconfirmations.

The viral account of AIDS comes out as progressive on Lakatos's terms. It predicted that HIV would be found in the T-Cells of AIDS patients, that higher viral loads would be correlated with more severe cases of the disease and more rapid progression toward death, and that ARVs would be effective. All of these predictions had been corroborated by the mid-1990s (although this corroboration came somewhat later than we might have liked). By contrast, the immune overload theory made ad hoc adjustments to accommodate the extension of the disease into new groups. These adjustments did not make novel predictions, and those elements of the theory that we might expect to be predictive did not yield corroborations – if, for instance, ARVs causes AIDS (as those committed to the immune overload theory claim), then we would expect there to be an increase in the number of AIDS cases as ARVs become more readily available, and this has not been the case (quite the opposite has been true). The immune overload theory is thus degenerative on Lakatos's account. Further, using Thagard's criteria, we can say that the immune overload theory has become pseudo-scientific. It fails to recognise the existence of an alternative progressive AIDS research programme in the form of the viral theory; it is unclear that an effort has been made by those who are committed to the immune overload theory to resolve its problems; and the AIDS denialists continual restatement of very out-of-date science is indicative of the selective way that they deal with confirmations and disconfirmations. The viral account was thus rationally accepted over the immune overload theory, at least by 1996.

However, it was also noted shifting to the strictly viral account of AIDS created an explanatory gap, in that it became more difficult to explain aspects of the disease (specifically the disproportionate prevalence of AIDS in southern Africa) without reference to aspects of the immune overload theory. The following chapter will argue that this explanatory gap might help to explain Mbeki's denialism, and I will attempt to remedy this problem by thinking more carefully about how

causation is understood in explanations of disease in epidemiology and biomedicine.

4. Thinking mono-causally and multi-causally about disease in the case of AIDS

4.1. Introduction

In 1984 the scientific community moved from multiple theories of the possible aetiology of AIDS to accepting HIV as *the* cause of AIDS. While this move was rational (the subject of the previous chapter), it resulted in the loss of some explanatory power – there were aspects of the disease that could not easily be explained in strictly viral terms, but which seemed to require reference to elements of the now discarded 'immune overload' theory; the theory that AIDS is due to diminished immune function resulting from various 'lifestyle related' factors, such as multiple STD infections, drug use, and malnutrition. For instance, on one plausible theory, at least part of the disproportionate prevalence of AIDS in southern Africa could be partially explained by drawing on the idea that malnutrition and already having been exposed to multiple other infections places strain on the immune system, making individuals more susceptible to opportunistic infections. Mosley makes the point as follows:

> It has been estimated that in America, the odds of contracting AIDS from an infected heterosexual partner is 1 in 500. But in many parts of Africa, the odds are 1 in 10. This difference in susceptibility to infection is explained, not by reference to HIV, nor to commensurably higher rates of unprotected sex that can be corrected by a focus on sex education. People suffering from malnutrition, parasites, and other forms of illness have compromised immune systems that make them more susceptible to infection by HIV than comparable healthy, well-nourished individuals in industrialised countries (Mosley, 2004, p. 409).[27]

While Mosley's point is controversial (more will be said about this later), it does seem straightforward that one should be able to hold the commitment that a disease is viral and still be able to draw on various socio-economic factors to explain aspects of that disease.

[27] Gray *et al* (2001), Sanders & Sambo (1991) and UNAIDS (1999) make similar points.

In this chapter I argue that being committed to a disease being both microbially-caused and subject to socio-economic factors became difficult because of an ambiguity about the relationship between mono- and multi-causal accounts of disease in biomedicine and epidemiology – where mono-causal accounts of disease require that there be only one, often microbial, source of illness for the purposes of explanation, and multi-causal accounts allow for a range of behavioural and socio-economic causal factors to be included in the explanation of disease. These two approaches are often taken to explain different types of disease, where mono-causal accounts are used to explain infectious diseases, while multi-causal accounts are used to explain chronic non-communicable disease (diabetes, hypertension, heart disease, certain forms of cancer, etc.). This becomes a problem in cases like that of AIDS, because the available conceptual framework for understanding disease requires that one focus on either viral or socio-economic factors, but not on both.

Once I have spelled out the distinction between mono-causal and multi-causal approaches to disease, I will argue that this division might help to explain Thabo Mbeki's AIDS denialism in the early 2000s. On this description, Mbeki wanted an explanation for why issues, such as poverty, seemed to be so intertwined with AIDS in southern Africa. The mainstream viral research programme at the time (a strictly mono-causal account) failed to provide an adequate explanation, which led Mbeki to consult non-mainstream AIDS scientists (who were very multi-causal in their approach). The non-mainstream scientists were able to provide a plausible-sounding account of the link between poverty and AIDS, but also incorrectly rejected the causal role of the virus in the process.

If I am correct about the role that these conceptual frameworks played in Mbeki's AIDS denialism, then this should put pressure on the mono-causal/multi-causal divide, motivating us to find an integrated approach that allows for multiple causal factors while still maintaining the causal salience of the virus for the purposes of explaining AIDS. Susser's (1973) layered multi-causal model seems to solve this problem. However, Susser's approach raises a number of philosophical issues in the process. In particular, it creates problems related to

the ontological status of the levels in his account, and the relationship between those levels (Russo, 2011, p. 77). It also fails to account for causal factors that are not the most proximal to the disease, but which are the most salient for explaining that particular disease. I will argue that instead of Susser's account, the best way to think about this problem is as one of 'causal selection' – that is, how one should distinguish between 'causes' and so-called 'enabling conditions'. Further, I will suggest that Woodward (2011) offers a more objective approach to solving the causal selection problem than the available philosophical alternatives, and that this provides a promising way to think about the relationship between mono-causal and multi-causal accounts of disease.

A caveat should be noted before proceeding. No attempt will be made in this chapter to assess the metaphysical status of causation in the medical and health sciences. There is already a vast literature on this topic, and the focus in this chapter is on a different set of issues. Specifically, the focus here is on the models and metaphors that underpin thinking about disease, and their implications for policy and health practice. For the purposes of this chapter, I assume that causation is a matter of probabilistic association, underpinned by actual mechanisms (Russo & Williamson, 2007).

It should also be noted that in some ways this is not a debate about HIV/AIDS at all – this is an issue for all infectious diseases that are subject to 'social drivers'. Stillwaggon's comments on the outbreak of the plague in Europe are particularly instructive in this regard:

> Throughout history it has been clear that the epidemic spread of disease requires favorable conditions. Rats (or soldiers) aboard a ship from an eastern port carried plague-infected fleas into Italy in 1348 and sparked the epidemic spread of plague in Europe, wiping out one-third of the population in most of the continent. This introduction was a random event, but it was certainly not Western Europe's only exposure to rats or plague. In 1348, plague entered a continent weakened by 30 years of failing per capita food consumption and increasing immiseration of the peasantry due to increased feudal demands. The population of Europe had already been falling in the decades leading up to 1348, and a series of disastrous harvests exacerbated the effects of war… Even though many nobles and townspeople perished in the Black Death, the ecologic context

> for the epidemic was the worsening economic situation of the peasantry (Stillwaggon, 2006, p. 8).

As Stillwaggon's example makes clear, there are some ways in which this is all old news. Practitioners of epidemiology and the biomedical sciences have been aware for a long while that both infectious agents and socio-economic factors need to be considered simultaneously in order to make sense of disease. However, the AIDS case makes it clear that this historical lesson has not been taken to heart and attention needs to be paid to this theoretical concern. We need a more sophisticated conceptual framework for discussing multiple causes of an effect – here, disease. That is the focus of this chapter.

4.2. Mono-causal and multi-causal approaches to disease

The mono-causal account of disease came into existence in the nineteenth century. Prior to that, the miasma theory (the theory that diseases, such as cholera, arose due to poor air quality) and the humour theory (that disease came about when the four fluids of the body – black bile, yellow bile, blood and phlegm – became unbalanced) dominated (Lee, 2012, p.117; Codell Carter, 2003, p.10). Granted, both the humour theory and the miasma theory also appear to be 'mono-causal', in the sense that one can appeal to a single cause to explain the onset of disease ("it was the bad air"), but the standard nomenclature takes mono-causalism about disease to be almost synonymous with the 'germ theory of disease' (the idea that microbes are causally responsible for diseases). For the purposes of this chapter, mono-causalism will be taken to refer to the germ theory of disease, unless otherwise specified.

The germ theory of disease came into being in the 1860s, when Louis Pasteur identified the microbes involved in fermentation. The realisation that something microbial could cause fermentation, and the ability to identify microbial agents for the first time, opened up the conceptual space in which it became conceivable that microbes could be necessary for all kinds of biological processes (this is often referred to as the "age of bacterial discovery") (Evans, 1993, p.8-10). This heralded in the mono-causal account of disease, as it is typically understood

(Thagard, 1999, p. 24- 25). In its broadest construal, the mono-causal account is just the insistence "that every disease has one cause that is necessary and, in limited circumstances, sufficient for disease" (Broadbent, 2009, p. 302).[28] But, of course, nobody thinks that any disease is *literally* mono-causal (Lee, 2012, p. 134). Broadbent makes this point as follows:

> … [B]ut no disease - and no other event we ever encounter – is mono-causal in the literal sense of being the operation of just one cause. The ingestion of *Vibrio cholerae* bacteria is a cause of cholera, but so too – in different ways – are the presence of human waste in the drinking supply, the lack of fuel to boil the water, the continued presence of oxygen in the atmosphere while the bacteria multiply inside the breathing host, the operation of the host, and so on. This observation suggests that the way to understand "monocausal" is as asserting a restriction, not on the number of causes, but on the number of causes that meet a certain restriction. (Broadbent, 2013: 149).

Broadbent goes on to claim that this 'certain restriction' is provided by Koch's Postulates (Broadbent, 2013: 149). Although there have been many versions and interpretations of the postulates, they can be outlined as requiring that an identified microbial agent, which is suspected of causing a particular disease, should satisfy the following criteria:

> 1). The microbe must be found in every instance of the disease.
>
> 2). It must be possible to grow the microbe in pure culture.[29]
>
> 3). It must be the case that when the pure culture of the microbe is injected (or otherwise introduced)[30] into a healthy animal it develops the disease.
>
> 4). It must be possible to find the microbe in the animal that was experimentally infected (Lee, 2012, p. 121).

[28] It should be noted that Lee (2012) and Broadbent (2013) both point out that the germ theory of disease and the mono-causal account are conceptually distinct. One can be committed to the mono-causal account without being committed to the single cause of any particular disease being microbial. Diseases of deficiency provide good examples of this. For instance, scurvy can be explained mono-causally by reference to a Vitamin C deficiency. However, historically, those endorsing the mono-causal account of disease have largely focussed on microbial agents.

[29] It is accepted that viruses are never able to fulfil this requirement, because they require cells to 'survive' and so it is not possible to grow a virus in pure culture.

[30] In an interesting side example, Barry Marshall drank a petri dish of the cultured bacterium *Helicobaterpyloris,* in an attempt to satisfy this postulate and establish the causal link between the bacterium and the development of ulcers (Lee, 2012, p. 123). Hence the need to specify that there are other ways of introducing the purported causal agent into the animal model.

Criteria 1) and 3) are the most important for establishing causal claims; with 1) establishing the necessity of the microbe for the disease and 3) providing evidence for the microbe being 'sufficient' for the onset of disease under a range of normal background conditions (underpinned by an interventionist picture of causation). 2) and 4) are more practical laboratory guides. 2), the requirement that the microbe be grown in pure culture, is included in order to isolate the particular microbe under investigation from potential sources of contamination. Incidentally, this is why it is not much of a worry that viruses fail to satisfy this criterion, because so long as precautions have been taken to isolate the virus from potential sources of contamination, then the function of the criterion 2) has been fulfilled. Again, 4) is an attempt to isolate the microbe under investigation and to provide evidence that it is *really* responsible for producing the disease (i.e. to rule out 'confounders', to use the common talk). We can imagine a scenario in which the microbe is injected into an animal, but its immune system successfully fights off the infection, and yet the animal still becomes sick, but not as a direct result of the microbe that was injected into it (perhaps it is allergic to the material that the syringe is made of) – the fourth postulate is meant to rule out that possibility.

From the start it was recognised that no microbe was likely to satisfy all four postulates. However, the postulates provide an ideal to be emulated when a researcher identifies a microbe as a plausible cause of a disease, even though these criteria might never be perfectly realised in practice. The aim of Koch's Postulates is to provide a guide for what one (ideally) ought to do when attempting to establish causation between an identified microbe and the disease it is suspected of causing. It sets high standards for the causes to look for – that is, causes that are necessary and sufficient under the range of normal background conditions.

In practice, the mono-causal approach proved to be a very fruitful research programme. On discovering a new disease, the practical task became one of identifying a plausible microbial causal agent, and then using Koch's postulates to

establish that the microbe is necessary and sufficient for the effect under the range of normal background conditions (as opposed to the microbe being coincidentally present or the secondary by-product of an underlying common cause). This approach allowed for the identification of the microbes responsible for a wide range of diseases (Koch himself, in the nineteenth century, identified the microbial agents responsible for Cholera, Tuberculosis and Anthrax), and the discovery of pasteurisation and immunisation (Lee, 2012, p. 117; Evans, 1993, p. 10). There was nothing about 'mono-causalism' per se that made this research programme so successful – that is, it was not the commitment that there should be only *one* cause for every disease that drove innovation during this time period. Rather, what allowed progress to occur was the thought that for every disease there was an underlying responsible microbial agent, because it meant that whenever a new disease was encountered the task became one of identifying the relevant microbe; thus providing a practical research strategy to scientists encountering new or unexplained diseases. This proved to be a very successful strategy for a large number of diseases that had previously been aetiologically mysterious. In those cases in which it was unsuccessful, it was just assumed that more time was needed to discover the relevant microbe, and that eventually the infectious agent would be found.

Despite the early successes of the mono-causal approach, the multi-causal account arose in the mid-twentieth century as a reaction against it. This was because the growing incidence of chronic non-communicable diseases (CNCDs) (diabetes, heart disease, certain forms of cancer, hypertension – diseases typically associated with lifestyle) could not be explained by reference to a single salient causal factor (Susser, 1985, p.150; Broadbent, 2009, p. 305). The multi-causal model came to be identified with the central metaphor of the 'web of causation', which Krieger describes as follows:

> Conceptually, the metaphor evoked the powerful image of a spider's web, an elegantly linked network of delicate strands, the multiple intersections representing specific risk factors or outcomes, and the strands symbolizing diverse causal pathways. It encouraged epidemiologists to look for multiple causes and multiple effects, and to identify the many – as opposed to singular – routes by which disease could be prevented. (Krieger, 1994, p.891).

Again, this is a piece of practical methodological advice for epidemiologists ("look for multiple causal pathways") and not a shift in thinking about ontology of disease causation, especially given that nobody had ever actually thought that diseases were literally mono-causal to begin with.

At this point there are two plausible theories of disease explanation on the table: the mono-causal account, which was developed to explain infectious diseases, and the multi-causal 'web' which came about in the middle of the twentieth century to explain CNCDs. It is not immediately clear what the relationship between these accounts of disease explanation is, or should be. The normative question of what the relationship between these two accounts *should* be will be addressed later in this chapter. At this point, it should just be noted that in the case of HIV/AIDS, practicing scientists working on the disease appear to have been strongly methodologically committed to mono-causalism. For instance, Robert Gallo, one of the scientists credited with the discovery of the HIV virus (as discussed in the previous chapter), explicitly states that he is committed to the position that "… multifactorial is multi-ignorance. Most of the factors go away when we learn the real cause of a disease" (Gallo, 1991, p. 148). This quote is taken from his autobiography, in a chapter entitled "A Single Disease with a Single Cause" – a title that further emphasises Gallo's commitment to mono-causalism. Further, Oppenheimer (1992) argues that the isolation of the HIV virus resulted in a substantial decrease in the number of epidemiological AIDS studies that were undertaken in the 1990s (p. 49-50). Given that epidemiology is the multi-causal health science par excellence, this is another indicator that the mono-causal account had taken centre-stage in AIDS research as the viral research programme grew in the late 1980s and 1990s, to the exclusion of multi-causal factors in the explanation of the disease. Oppenheimer makes this point as follows:

> With the isolation of the HIV, the relative importance of epidemiology in the definition of the disease lessened… Increasingly, the "bench" scientists – virologists, immunologists, cancer researchers – determined the definition of HIV infection. In effect, they redefined AIDS as a set of biomedical problems open to a chemical resolution in the form of drugs

> and vaccines… The change in the types of professionals studying infection and in their defined fields of observation and analysis affected a subtle shift in the characterisation of the disorder. The disease was increasingly conceptualized in terms of the infectious agent, the virus. Interest in cofactors or a multifactorial model diminished. (Oppenheimer, 1992, p. 63).

Fee and Krieger (1993) agree with Oppenheimer:

> Once the virus was identified, scientists tended to lose interest in the social factors accompanying transmission. They instead turned to laboratory studies of the virus and its action within the body… (Fee & Krieger, 1993, p. 1478)

This strong commitment to mono-causalism in the case of AIDS (as seen in Gallo, Oppenheimer, and Fee and Krieger) meant that it was not permissible to draw on socio-economic explanations of disease (because the mono-causal account allows only for a single, typically microbial, source of disease to be included in the explanation). Without being able to draw on factors, such as poverty, in the explanation of disease, this rendered the disproportionate prevalence of HIV/AIDS in southern Africa somewhat mysterious.

4.3. Was there really an explanatory gap in the HIV/AIDS case?

This chapter started with the concern that the transition from multiple theories of AIDS's aetiology (some of which accommodated socio-economic factors) to the strictly viral account may have resulted in a loss of some explanatory power. In particular, the concern was that some of the disproportionate prevalence of AIDS in southern Africa could not be explained in strictly viral terms, but that it might be explicable by appealing to socio-economic factors. However, if one is committed to strict mono-causalism about disease, as the AIDS research community in the 1990s appears to have been, then it is not permissible to draw on non-viral factors to explain any aspect of any particular infectious disease.

In this section, I will look more closely at whether it was reasonable to be concerned about an explanatory gap created by strict adherence to mono-causalism in the case of AIDS. In particular, I assess whether we can explain the

disproportionate prevalence of AIDS in southern African in strictly viral terms – either in terms of differences in the viral strains themselves or differences related to individuals' exposure to the virus[31] – or whether we need to make use multi-causalism to draw socio-economic factors into the explanation. I will look at three mono-causal explanations of the disproportionate prevalence of HIV/AIDS in southern Africa (in increasing order of plausibility): 1) that there is a more virulent (or infectious) strain of HIV in southern Africa; 2) that southern Africans participate in riskier sexual practices (broadly construed); and 3) that southern Africans are more likely to have concurrent sexual relationships, and that this explains the variation (a narrower version or the explanation presented in 2).[32]

One way of explaining the disproportionate prevalence of HIV in southern Africa might be to appeal to how virulent or infectious different strains of the virus are – if strains of HIV that are typically found in southern Africa are worse than strains found in western Europe and the United States, then this would explain the difference without reference to any socio-economic factors. It is well known that HIV-1 is a more virulent strain of the virus than HIV-2 (Flint et al, 2009, p. 167). If the regional variation of prevalence tracks the variation in the strains of the virus, then a strictly viral explanation of the disease's distribution could be produced. However, this is just empirically not the case. HIV-2 is most prevalent in West Africa and does not really extend beyond that region (Flint et al, 2009, p. 167). This means that all of the areas we are trying to account for (western Europe, the United States, and southern Africa) are affected by HIV-1. Limited research is available on how virulent various sub-strains of HIV-1 are, but the research that does exist suggests that there is no difference (Gray *et al*, 2001). No mainstream accounts suggest that differences in strains of the virus have

[31] Whether this would count as a strictly mono-causal explanation is debateable, because it makes reference to population behaviour in the explanation and hence something other than the virus is being appealed to. However, population level behaviour could be compressed into something approximating a mono-causal explanation by describing this as merely amounting to increased instances of exposure to the virus. This is how those who are committed to mono-causal explanations conceive of the issue.

[32] Again, it is somewhat controversial to describe 2) and 3) as strictly mono-causal accounts, given that they both appeal to behavior in their explanations. But again, one might claim that this just amounts to differences in exposure to the virus, and so all that really matters to the explanation is the virus itself.

anything to do with the increased prevalence in southern Africa (Epstein, 2007, p.52).

If there is nothing about the strains of the virus itself that makes them more or less virulent, then perhaps the differences in prevalence between southern Africa and the rest of the world could be explained in terms of variations of sexual practices, which would leave the viral theory intact – the only difference in the case of southern Africa being a greater number of instances of possible transmission (Iliffe, 2007 , p. 63).[33]

It seems like common sense that sexual practices might make a difference to the HIV/AIDS epidemic, given that most instances AIDS in southern Africa are sexually transmitted (as opposed to transmitted via injection drug use, or blood transfusions) (Epstein, 2007, p. 50). Additionally, the migrant labour system that supports the South African mining industry is put forward as an obvious example of how sexual practices increased instances of exposure to the virus and thus drove the epidemic. Quoting Steinberg (2011):

> [T]he mining system was structured in such a way that it could potentially disperse the virus to all four corners of the region. South Africa's quarter-of-a-million migrant mineworkers lived two lives, many with at least two long-term lovers, one at work and the other at home in the countryside.

However, the evidence on sexual practices being the only difference-maker to HIV prevalence is patchy. A report issued by UNAIDS in 1999 highlights this. In this report, four African cities (two high prevalence and two low prevalence) were studied in an effort to isolate the causal factors that lead to higher prevalence – these cities were Kisumu (Kenya), Ndola (Zambia), Cotonou (Benin)

[33] Note that Iliffe is presenting this as an explanatory stance that one might adopt toward the southern African AIDS epidemic. He does not endorse this position himself. His own explanation for the disproportionate prevalence of HIV/AIDS in southern Africa is that the disease has just been in Africa for longer than in other places, giving it more time to spread unidentified than had been the case in other parts of the world – again, this would be another mono-causal account because it just has to do with instances of transmission over time (2007, p. 58). However, while Iliffe's description might be true in some parts of Africa, it cannot explain the South African HIV/AIDS epidemic, where prevalence was at less that 0.1% in 1986 (Steinberg, 2011) at least partially because international sanctions during apartheid meant that South Africa was very isolated and was only exposed to the virus quite late.

and Yaoundé (Cameroon).[34] Various aspects of sexual practice (such as number of sexual partners, and contact with sex workers) and potential co-factors (incidence of additional sexually transmitted diseases, and male circumcision) were included as potential drivers of prevalence. The study concluded that there was insufficient evidence to put differences in prevalence solely down to differences in sexual behaviour. They also exclude the possibility that different sub-types of the virus were responsible for differences in prevalence:

> From these data, *it would be difficult to argue that the divergence in the rate of HIV spread between the East African and West African sites can be explained solely by differences in sexual behaviour*. Nor was there evidence that differences in circulating strains of HIV-1 are a major factor in the rate of spread of HIV. Subtype A was the most prevalent subtype in both of the low-prevalence sites as well as in one high-prevalence site…Finally, how can one explain the dramatically high prevalence of HIV infection (15%-23%) in girls under 20 in Kisumu and Ndola? Unmarried girls in these cities reported a median of 1 to 1.5 lifetime sex partners, an estimated 10-12% of whom were HIV infected… it is hard to explain the high HIV prevalence in female teenagers. For the girls to have become infected so soon after their sexual debut as a result of relatively few exposures to an infected partner, *HIV transmission co-factors must be part of the explanation*… In conclusion, *differences in the rate of HIV spread between the East African and West African cities studied cannot be explained away by differences in sexual behaviour alone*. In fact, *behavioural differences seem to be outweighed by differences in HIV transmission probability*. (UNAIDS, 1999)[own emphasis added].

Stillwaggon comments on this UNAIDS report that it "provided a clear empirical challenge to the behavioural assumption. Unfortunately, UNAIDS has not followed up on unanswered questions…" (Stillwaggon, 2006 , p. 19). So while this study provides some empirical reason to doubt the view that differences in prevalence are solely due to differences in numbers of sexual partners, this study has not been appropriately followed up.

The broad theory that differences in "sexual practices" (which is presumably encompasses things like increased numbers of sexual partners, more visits to

[34] It is recognised that not all of these locations are in southern Africa, but this study offers general evidence that 'differences in sexual practices' (broadly construed) are not the sole or primary difference maker in these areas, and so it is unlikely that they will be the sole difference maker in southern Africa.

sex-workers, and the like, but is very non-specific) drives the AIDS epidemic in southern Africa, does not seem well-supported by the evidence. For instance, it seems that prevalence could be higher in a location without individuals having a higher number sexual partners. That is not to say that those individuals were having no sex, or that they were strictly monogamous, just that they had no more partners than the contrast case, but that the contrast case maintained a lower prevalence rate. This suggests that the number of partners alone cannot explain divergent prevalence rates.

The alternative theory that was offered to the very broad 'sexual practices' account is the 'concurrency theory' (Sawers & Stillwaggon, 2010). The idea is that there is more of a norm of southern Africans having a small number of concurrent sexual relationships, which can last for long periods of time, as opposed to the "serial monogamy that is more common in Western cultures" and that this explains the differences in rates of prevalence (Epstein, 2007, p. 55). There is a plausible mechanism on offer for this account. Individuals have a high concentration of the virus in their blood in the period just after they have been infected, which makes them more infectious during this period than later on:

> This means that sexual networks in which people sleep with two or three partners at intervals of days or weeks are probably [very dangerous] …The existence of a "viremic window" early in infection, when transmission is especially likely, also sheds light on why HIV spreads so slowly in populations practicing serial monogamy. By the time the serial monogamist has moved on to a new partner, his viral load will have fallen, so he is unlikely to infect her (Epstein, 2007, p. 61)

In addition to the plausible mechanism, mathematical models have been produced that show increased HIV prevalence when concurrency is modelled as opposed to serial monogamy (Kretzschmar & Morrus, 1996 ; Morris & Kretzschmar, 1997 ; Morris & Kretzschmar, 2000). A combination of the plausible mechanism and mathematical modelling made the concurrency theory the mainstream view for explaining the disproportionate prevalence of HIV in southern Africa (Allais & Venter, 2012; Sawers & Stillwaggon, 2010). However, on closer examination, the theory is less convincing than hoped.

A major problem with the evidence for the concurrency theory is that much of its plausibility rests on the mathematical models, but the models themselves are based on implausible assumptions. Most implausibly, the models "assume that everyone in a partnership has sexual contact with every one of their partners every day" (Sawers & Stillwaggon, 2010, p. 3). Furthermore, it assumes a very high rate of infection and that concurrency is gender symmetric (that both men and women have multiple partners),[35] neither of which is supported by the evidence. Once these assumptions are corrected to more closely align with the data from the areas being studied, the model fails to generate any difference in HIV prevalence between concurrency and serial monogamy (Sawers & Stillwaggon, 2010).

None of the strictly mono-causal accounts of the disproportionate prevalence of HIV/AIDS in southern Africa have been particularly convincing. Lurie and Rosenthal (2010) argue that we should not be very surprised that no single mono-causal account of HIV/AIDS completely explains the geographical variation of the disease, arguing instead that:

> In reality, HIV epidemic dynamics are complex and unlikely to be explained by a single variable. Instead, a combination of factors likely drives the epidemic – with some factors playing a more important role in some geographic areas than others (Lurie & Rosenthal, 2010, p. 18).

A non-viral factor that may play an explanatory role in the geographical variation is poverty. As noted before, the likelihood of contracting the virus in any particular sexual encounter is substantially higher in southern Africa than in the United States or Western Europe (Mosley, 2004; Gray *et al*, 2001; Sanders & Sambo, 1991). The mechanism for this is not mysterious – note that we are happy to accept that the poor are more susceptible to all sorts of other diseases because of poverty-induced diminished immune function (remember the example from Stillwagon about the introduction of the Plague into Europe in the

[35] Gender symmetry is an important assumption for the model to work, because it allows for a whole network of infectiousness to be created. This is clear in the case of the South African mine worker who has a long-term partner at home and a long-term partner at the mine. If neither of the miner's long-terms partners have additional partners of their own, then the virus remains trapped within the two concurrent relationships, and the virus is not distributed any further.

1300s). In the case of HIV/AIDS, the increased susceptibility of the poor to the disease operates as follows. HIV positive individuals are more infectious when they have a high 'viral load'.[36] Individuals who already have diminished immune function – due to malnutrition, concurrent exposure to other infections, etc. – spend more of the time at higher viral loads, thus meaning that they are more infectious for more of the time (Nattrass, 2007 , pp. 18-22). It thus seems as though our explanation for the disproportionate prevalence of HIV in southern Africa should include some reference to the role of poverty and malnutrition in the spread of the epidemic.[37]

A question that remains at this point is why the mainstream HIV/AIDS research community at the time was so wedded to a mono-causal account of AIDS, given how little credibility the various mono-causal theories of the epidemic have when they are scrutinised.[38] My own suspicion for why this occurred is that non-viral causal factors had lost credibility when HIV was accepted as *the* cause of AIDS, and had become too closely associated with the AIDS denialists – who continued to argue that diminished immune function is the result of malnutrition, STDs and drug use, and that it is *not* viral. To draw on non-viral causal factors in an explanation of AIDS was to place oneself too closely in the territory of the AIDS denialists, and was thus a risk to one's credibility. However, this is more of a question for social scientists than for philosophers. Krieger (1994), for instance, would argue that the overall dominance of the viral research programme, relative to other forms of research in the medical and health sciences, would explain why the research community lost interest in non-viral issues once HIV had been identified, because viral accounts of disease are just taken more seriously than others. Alternatively, Sawers and Stillwaggon (2010) argue that

[36] This is a good explanation for why 'treatment as prevention' is successful. Being on antiretrovirals lowers the viral load, making one less infectious. Thus having an effective national antiretroviral programme lowers overall prevalence (Attia, Egger, Zwahlen, & Low, 2009).
[37] None of this is to say that sex plays no role in AIDS transmission rates, just that sexual practices alone do not always explain differences in prevalence at the population level, and so conceptual space needs to be opened up to include additional drivers of disease.
[38] In the early days of mono-causalism, this kind of commitment made sense, because it was thought that every disease literally had a single microbial cause. But the advent of multi-causalism in the wake of chronic non-communicable diseases (hypertension, heart disease, diabetes, etc.) makes such a strong commitment to a mono-causal explanation less sensible, and the continued commitment to a strictly mono-causal account of disease in the face of counter-evidence less sensible.

resistance to a more multi-causal account was the result of a commitment to the idea that the explanation *must* be sexual (Mbeki would argue that the commitment to the idea that the sole difference maker must be sexual is due to racist attitudes towards African sexuality); so much so that when the evidence showed that sexual behaviour could not be the only difference maker (such as the evidence offered by the UNAIDS report cited above), the concurrency theory was offered as a last-resort attempt to explain the disproportionate prevalence of HIV/AIDS in sexual terms (p. 13). Regardless of why the mono-causal account was so dominant in scientific HIV/AIDS research in the 1990s, this was the state of the debate when Mbeki became president of South Africa in June 1999.

4.4. Explaining the South African case in terms of mono- and multi-causal accounts of disease

In this section I argue that Mbeki's AIDS denialism can plausibly be explained as a result of the scientific community's strict commitment to mono-causalism about AIDS in the 1990s, to the exclusion of socio-economic factors (i.e. to the exclusion of factors that would typically be associated with a multi-causal account). If I am correct about this, then the tragedy of the case should motivate us to think more carefully about how to better understand the relationship between mono-causal and multi-causal accounts of disease.

One plausible (but charitable) reading of the Mbeki case is that he noted that there were aspects of the disease that the viral account did not seem able to explain on its own. In particular, that the viral account of AIDS seemed incapable of explaining the disproportionate prevalence of the disease in southern Africa, when compared to the United States or western Europe and this led Mbeki to consult scientists who strayed from the viral orthodoxy.

On this description of the Mbeki case, he began consulting non-mainstream AIDS scientists in order to find answers to the unresolved questions that he had identified, especially related to the regional variation of the disease, and the apparent connection between poverty and AIDS (even within southern Africa,

HIV/AIDS prevalence is highest amongst the poor). Phrased slightly differently, the strictly mono-causal account of AIDS did not resolve issues that a more multi-causal account might have been able to, because a more multi-causal approach would have been able to draw on socio-economic factors in its explanation.

Mbeki's speeches from this period add plausibility to the idea that this was motivating his action. His opening speech at the 13th International AIDS conference in Durban, on 9 July 2000, is particularly instructive in this regard. Much of this speech focuses on the impact of poverty on health in Africa. Concluding the discussion on poverty he states:

> One of the consequences of this crisis [poverty] is the deeply disturbing phenomenon of the collapse of the immune system among millions of our people, such that their bodies have no natural defence against attack by many viruses and bacteria (Mbeki, 2000).

This seems like an obvious reference to the immune overload theory of AIDS: poverty causes the immune system to collapse, which makes individuals vulnerable to opportunistic infections. Further, the idea that Mbeki turned to non-mainstream AIDS scientists *because* there were gaps in the mainstream viral research programme is alluded to again in the same speech when he states:

> Some in our common world consider the questions I and the rest of our government have raised around the HIV-AIDS issue, the subject of the conference that you are attending, as akin to grave criminal and genocidal misconduct. What I hear being said repeatedly, stridently, angrily, is – do not ask questions!... As I listened to the whole story being told about our country, it seemed to me that *we could not blame everything on a single virus* (Mbeki, 2000) [italics added].

The above quote, particularly the portion in italics, suggests that Mbeki was consulting non-mainstream scientists because he believed that there were aspects of the disease that required reference to concepts beyond the virus. Further support is added to this explanation of Mbeki's action by comments in his welcome address to the members of the Presidential Advisory Panel. This speech begins with Mbeki quoting AIDS prevalence statistics from a then recent WHO (World Health Organisation) report, particularly that sub-Saharan Africans make up "85% of the global total [of people diagnosed with AIDS], even though

only one-tenth of the world population lives in sub-Saharan Africa" (Mbeki, 2000a). This is the first suggestion that Mbeki was suspicious about the distribution of the disease. He then goes on to note that both the prevalence and the distribution of AIDS have changed in Sub-Saharan Africa, but not in the US or Europe. He takes this to be a strange outcome, and one that requires further investigation, which had not at that point been undertaken within the viral research programme. In Mbeki's words:

> The situation has not changed in the United States up to today, nor in Western Europe with regard to homosexual transmission. But here [in southern Africa] it changed radically in a short period of time and increased radically in a short period of time. Why? This is obviously not an idle question for us because it bears very directly on this question: How should we respond? (Mbeki, 2000a).

In this speech he also explicitly states that the reason the Presidential Panel had been assembled (and presumably the reason why non-mainstream AIDS scientists had been included on the panel) was to resolve these concerns. In this regard, Mbeki states that:

> It is truly our hope that this process will help us to get to some of the answers, so that as public representatives we are able to elaborate and help implement policies that are properly focused, and that actually have an effect. I'm quite certain that *given the people who are participating in this panel*, we will get to these answers (Mbeki, 2000a) [italics added].

These excerpts from Mbeki's speeches indicate that it is plausible that Mbeki's behaviour might be at least partially explained by reference to the divide between mono-causal and multi-causal accounts of disease, and the resultant emphasis that this placed on viral aspects of AIDS research, to the exclusion of socio-economic factors.

Mark Gevisser (2007), Mbeki's biographer, suggests that Mbeki himself viewed the problem in this way. Gevisser describes Mbeki's thinking on this as follows:

> And then that word again: 'paradigm'. Scientists and doctors are 'committed to a particular paradigm,' he [Mbeki] told me [Gevisser], which is that a retrovirus is the sole cause of AIDS, and therefore that AIDS is best combated by antiretrovirals (ARVS), the exorbitant drugs

> marketed by pharmaceutical companies rather than more cost-effective nutritional solutions. It is not, as he has repeatedly said, that he denies the existence of AIDS. But the question he insisted on asking – the question asked by the small group of AIDS-dissident scientists – was this: 'What causes the collapse of these immune systems? Now the answer [commonly given] is a virus, but… if you look at the literature you can see that it's actually wrong, to attribute collapse of immune systems which might be happening to millions in South Africa and elsewhere on the continent, to attribute it just to a virus.' Scientists were 'jumpy' about the questions he was asking, Mbeki told me, because 'what happens if this entire paradigm gets changed?' (Gevisser, 2007 , pp. 727-728)

Other writers on Mbeki's denialism have made similar suggestions. Butler (2005) and Mosley (2004) both argue that the mono-causal/multi-causal divide might have underpinned Mbeki's denialism (although they each use slightly different terminology to make this point). Butler (2005) remains neutral about the relative merits of these perspectives, and focuses instead on describing the context that allowed for the ascendancy of the multi-causal approach within the South African government at the time (amongst Mbeki and his supporters – more will be said on this in following chapter). Mosley (2004) argues that the mono-causal and multi-causal approaches are both legitimate perspectives to adopt toward disease, depending on what one is interested in – if you are interested in individual cases of the disease then you should focus on mono-causal viral descriptions of disease, but if you are interested in population level distributions of disease then you should go with a multi-causal epidemiological approach. As such, Mosley concludes that Mbeki's decision to go with a more multi-causal approach might have been justified, because he made use of a legitimate alternative approach to thinking about disease.

Van Rijn (2006) makes a similar point to Mosley, suggesting that 'virological' (mono-causal) and 'epidemiological' (multi-causal) perspectives both provide legitimate ways of viewing disease, and that Mbeki just happened to favour a more epidemiological approach. Further, Van Rijn seems to approve of Mbeki's decision, because he argues that Mbeki moved the discussion about AIDS beyond purely viral concerns, which is taken to be a good thing. Van Rijn is unclear about what aspect of Mbeki's behaviour he supports – presumably he means the inclusion of non-mainstream scientists on his advisory panel (given that this

would have shifted the debate) and not the policy decision of not making ARVs available (given that this was due to a rejection of the causal role the virus played, and didn't just make the debate more inclusive of non-viral accounts of disease). Note that, even if Mbeki did not deny the causal role of the virus, he may nonetheless have made the mistake of focussing on the wrong cause when it came to intervening to prevent transmission in South Africa. Even if 'immune overload' is the relevant difference maker when it comes to explaining regional trends, it does not necessarily mean that this is the relevant cause to focus on when it comes to developing a policy to stop HIV transmission in South Africa, as the quote above from Mbeki suggests.

Fassin (2007) similarly suggests that Mbeki's behaviour might be explained by reference to the division between 'viral' (mono-causal) and 'sociological' (multi-causal)[39] accounts of the disease (p.15). However, he does not adopt Bulter's value-neutral perspective on the division. Nor does he agree with Mosley that either approach is legitimate. He also does not follow Van Rijn's suggestion that Mbeki productively advanced the discussion about AIDS beyond its purely viral elements. Instead, Fassin argues that the Mbeki case should put pressure on these underlying accounts of disease, such that an attempt should be made to unify the mono- and multi-causal approaches, stating that "one seeks a kind of third way, a means of making biological and social theories compatible" (p.15). I am sympathetic to Fassin's position here, and much of the rest of this chapter will focus on trying to figure out what unifying these accounts would amount to.

4.5. An attempt to resolve the mono-causal/multi-causal divide

At this point we have seen that there are two prominent accounts of disease explanation on offer. The mono-causal account, which emphasises infectious

[39] I am aware that "multi-causal" and "sociological" are not synonymous. However, given the emphasis on microbial sources of disease in the mono-causal account, and the idea that any account of disease that is not mono-causal (in the standard sense of being concerned with microbes) is multi-causal, it seems that Fassin is pointing to a multi-causal approach to thinking about HIV/AIDS when he uses the term "sociological". His overall point here is just that we need to think about HIV/AIDS in a way that takes account of viral and non-viral drivers of disease at the same time.

microbially-based diseases, and the multi-causal account, as typified by the 'web of causation'. We have also seen that it is unclear what the relationship between the mono-causal and multi-causal account of disease is or should be, and that this is a source of theoretical as well as practical concern (as illustrated by the Mbeki case). Although some cursory points have already been made about this issue, in this section I will look more closely at this relationship.

There are at least three plausible descriptions of this relationship on offer: 1) the mono-causal and multi-causal accounts are just different approaches to explain different sorts of diseases; 2) the multi-causal approach subsumed the mono-causal approach, and the microbial sources of disease were placed on the causal web as one type of cause amongst many; 3) the multi-causal account of disease subsumed the mono-causal account, but the causal salience of microbial causes was preserved. Each of these will be discussed in turn.

4.5.1) Different accounts for different diseases

One way of dealing with the fact that there two different available accounts of disease is to argue that they are just different approaches for explaining different sorts of diseases – the mono-causal approach is used to explain infectious, typically microbial diseases and is restricted to the realm of the biomedical bench sciences (virology, immunology, microbiology, etc.), and the multi-causal account, which explains CNCDs and is most closely linked to epidemiology. Russo suggests this as a possible interpretation when she states that: "[D]isease causation may be properly described by the mono-causal or the multi-causal model, but that depends on the disease at hand …" (Russo, 2011, p. 75-76). I suspect that this way of thinking about the relationship between these accounts of disease is the most descriptively accurate, or at least the one that best captures the way scientists were thinking in the HIV/AIDS case in the late 1980s and the 1990s (see Gallo (1991), Oppenheimer (1992), and Fee and Krieger (1993) above). There are, however, a number of problems that result from adopting this approach, making it an undesirable way to think about this relationship.

Given the preceding discussion, the most obvious problem is that it is unclear how one should deal with diseases that are clearly infectious, but for which socio-economic factors are causally relevant. This is an issue for the AIDS case, but it would also be a concern for dealing with any disease that involves both microbial and social components, such as cholera or TB. Indeed, for any disease, some details about behaviour and context are going to be causally relevant. Given the historical dominance of the mono-causal approach (Krieger, 1994), the microbial aspects of disease are likely to take priority, and maybe rightly so. However, if one is committed to the mono-causal and multi-causal accounts of disease being strictly distinct, then this results in the exclusion of potentially relevant socio-economic factors.

Russo (2011) alludes to a related issue. A large part of the problem of ignoring socio-economic factors in order to be fully mono-causal about disease is that this results in non-microbial causes being ignored for the purposes of health policy development (Russo, 2011,p.90-91).[40] It is, however, obviously not the case that being mono-causal about disease will result in socio-economic factors being ignored completely in policy. But then socio-economic concerns would need to be pursued in policy for reasons external to the disease in question.[41] If one is *really* committed to mono-causalism about disease, then one cannot appeal to the role that socio-economic factors play in any particular disease in order to pursue policies to improve socio-economic conditions.

One might be concerned that adopting a strictly mono-causal approach toward infectious diseases, especially in cases where it seems that socio-economic factors are actually at play, would leave one with an incomplete picture of the disease at hand. And this is conceptually unsatisfying.

[40] See Fee and Krieger's "Thinking and Rethinking AIDS: Implications for Health Policy" (1993) for a fascinating account of how thinking about AIDS through different paradigms in early AIDS science resulted in radically different policy approaches.
[41] It is likely that poor socio-economic conditions will be bad for people in lots of ways that have nothing to do with disease, and so there will be lots of non-disease related reasons that policy makers can appeal to in order to pursue policies that will improve these conditions.

4.5.2) The multi-causal account of disease subsumes the mono-causal account

Broadbent (2009; 2013) suggests that the multi-causal account of disease subsumed the mono-causal account. This provides a second way of thinking about the relationship between mono-causal and multi-causal accounts of disease.

On this description, the multi-causal account replaced the mono-causal account, such that the previously mono-causal sources of disease were placed on the web of causation as one cause amongst many. That is, the restrictions on the mono-causal account were merely removed, leaving us with just the multi-causal model. Broadbent describes this as follows:

> A multifactorial approach could simply reject the strictures of the monocausal model, asserting that diseases have many causes... This is the *bare* multifactorial model: it consists of no positive assertions about disease causation, and places no restrictions on what causal structures are specific to disease (Broadbent, 2009, p. 306).

This resolves the concern that was present when the mono-causal and multi-causal accounts were strictly divided. That was, because (on that account) the mono-causal and multi-casual accounts were different descriptions of different types of disease, if one classed a disease as infectious and thus mono-causal, then one was unable to appeal to aspects of the multi-causal approach (such as socio-economic drivers of disease). By removing the 'one cause' restriction from the mono-causal account and placing infectious agents on the 'web of causation', this problem no longer exists, because one is able to appeal to microbial and socio-economic sources of disease within the same model.

However, this is an obviously unsatisfying view of disease, because it fails to pick out the causal salience of microbial agents. Broadbent describes this problem as: "There is no discrimination and no hierarchy among causes, no 'primarily caused by' – just a 'constellation' of causes which may come together in one or more than one way to give rise to a case of a disease." (Broadbent, 2013, pp. 154-155).

In the case of AIDS, this would mean that HIV is placed on the causal web along with all the other factors associated with the disease, such as poverty, malnutrition and migrant labour paths. However, this does not pick out the fact that HIV is more causally important to AIDS than the socio-economic drivers of the disease. We do not think that poverty is connected to AIDS in quite the same way as HIV (eliminating HIV would end AIDS, but eliminating poverty would not). Our model of disease should make this distinction clear.

4.5.3) Joining mono-causal and multi-causal accounts. Preserving causal salience.

Given the preceding discussion, what we want is an account of disease that unifies the mono-causal and multi-causal accounts (so that we can draw on explanatory tools from both), but which preserves the causal salience of certain factors (in the case of AIDS, HIV). There are two ways to attempt this resolution. The first is the epidemiological textbook approach, offered by Susser's (1973) multi-level multi-causal model, which initially seems to be successful on both of these counts. However, it introduces philosophical problems related to levels, and it fails to account for how one ought to deal with causes that are the most salient to a particular disease but not the most proximal. The second way of attempting to join the two accounts is by treating this as a causal selection problem – i.e. how do we pick out the 'causes' from the 'supporting factors'? Once it has been agreed that we are dealing with a causal selection problem, there are two approaches for selecting the salient causes. First, there is the standard philosophical approach, whereby salient causes are identified on the basis of the interests and perspective of the individual doing the selection, but this process lacks rigour. Second, there is Woodward's (2011) approach, whereby causes are distinguished from so-called enabling conditions based on the 'stability' and 'specificity' of the relevant factors, and on the basis of contrast cases. I will argue that Woodward's approach is the most promising.

4.5.3.1 Susser's multi-level multi-causal model

I start by discussing Susser's textbook epidemiological approach to the problem. Susser maintains that our account of disease should be multi-causal, but suggests that the central metaphor of the 'web of causation' be replaced with one of nested levels. The levels are determined by "systems", where a system is:

> … [A] set or assembly of factors connected with each other in some form of coherent relationship. A system is an abstraction. It allows a set of related factors to be described in terms of coherent structure or coherent function (Susser, 1973, p. 48).

Examples of systems include the cardiovascular system, made up of the heart and blood vessels; the body made up of many subsidiary biological systems; or a society made of many bodies. Part of what defines a system, on this account, is that coherent analysis can occur within its structural limits. For instance, a cardiologist might (hypothetically) be able to study the cardiovascular system by restricting her analysis to just its component parts, without recourse to other systems (such as the external political system) (Susser, 1973, p. 48).

Despite the fact that each system is a conceptually self-contained unit, they are layered in a nested fashion and different systems interact. For example, smoke in the environmental system might damage the blood vessels in the cardiovascular system. Studies can therefore be conducted on either one of two axes – they can be horizontal, when analysis occurs entirely within a single system (the cardiologist conducts a horizontal study when she studies just the components of the cardiovascular system), or they can focus on the vertical axis, in which case the analysis cuts across multiple systems (an epidemiologist conducting a study on AIDS might pursue a vertical study looking at multiple systems at once– the immune system, the social system, etc.). Importantly, whether a study is horizontal or vertical, and which systems are assessed, depends on the subject of study (Susser, 1973, pp. 49-50).

The distinction between horizontal and vertical studies becomes relevant for Susser's view of causation. On his account, causation can either be 'direct' or

'indirect'. Direct causation occurs when cause and effect operate within the same system level (Susser, 1973, p. 51). For instance, the calcified artery caused the heart to fail – the artery and the heart failure both occur within the cardiovascular system. Indirect causation occurs when cause and effect operate at different system levels. For instance, prolonged smoking causes the artery to calcify, which in turn causes the heart to fail. Smoking is an indirect cause of the heart failure, because the smoking and the heart failure occur within different system levels (smoking might occur at the behavioural level, while heart failure occurs at the level of the cardiovascular system).[42] This latches on to the familiar idea in philosophy of chains of causation, where a cause is more indirect the further it is mediated by other causes in the chain. Reiss (2015) usefully borrows a quote from the television show *Frasier* to illustrate this point:

> Frasier: I cut myself because I was shaving without water. And why was there no water? Because I had to move your chair, which gouged the floor, which made me call Joe, who found bad pipes, who called for Cecil, who ate the cat that killed the rat that lived in the house that Frasier built! (Reiss, 2015, pp. 16-17)

Frasier wants to claim that what caused him to cut himself was that he had to move the chair. But this would be a very indirect cause, given the number of intervening causes that are required to connect the chair to the cut (Reiss, 2015, p. 17).

Susser's layered account allows one to appeal to both microbial sources of disease and socio-economic factors within a single account, because both are included in the same model. Further, in the case of HIV, it picks out the causal salience of the virus, because both the virus and the T-Cells that are targeted by the virus exist at the same system level, and so the virus is the *direct cause* of the disease, thus making it the most salient cause in the explanation of the disease. This produces the correct result, because if we were able to remove the virus, we would also remove the disease and so the virus must actually be the most salient cause. However, socio-economic factors are still included in this aetiological

[42] This is just an example to illustrate the point. It is not mean to be an accurate description of cardiology or the causes of heart attacks.

picture of the disease, as indirect causes. Perhaps malnutrition diminishes immune function, such that the individual is more susceptible to infection when they are exposed to the virus: malnutrition is thus an indirect cause of AIDS, and gets included in the causal explanation on Susser's view.

Susser's account therefore seems to resolve both of the problems that had previously been identified. It unifies the mono-causal and multi-causal accounts of disease, and it allows for certain causes to be picked out as salient. As Russo (2011) points out, Susser's account is a definite improvement on the 'bare-multifactoralism' of the web of causation, but it also creates some philosophical concerns by introducing the notion of levels (Russo, 2011, p.68).

Drawing on "levels" to resolve the divide between mono-causal and multi-causal accounts of disease introduces its own problems, because, as Craver (2007) points out, "the term 'level' is multiply ambiguous" (p. 163). In the special sciences, one can mean a number of different things when talking about levels:

> [T]here are levels of abstraction, analysis, behavior, complexity, description, explanation, function, generality, organization, science, and theory. Consequently, scientific and philosophical disputes about levels cannot be addressed, let alone resolved, without first sorting out which of the various senses of "level" is under discussion (Craver, 2007, p. 164).

Given that Susser specifies that the content of his levels are "systems", where a system is defined by a number of component parts operating together to fulfil a specific function (1973, p.48), it might seem clear that he takes levels to be functional. However, he goes on to state that "a system is an abstraction" (1973, p.48), which makes it seem as though he considers the levels in his account to be levels of abstraction. He later goes on to argue that different systems/levels are the subject of different academic disciplines, which makes it seem as though his levels are organised around (what Craver terms) the 'products of science' (Craver, 2007, p. 171). There is, therefore, some ambiguity surrounding the ontological status of Susser's levels. It is also unclear what the nature of the relationship between the levels on this account is. Susser tells us that different levels can interact with each other, but not much else. As Russo (2011) suggests,

Susser offers an improvement on "bare multi-factoralism", but this account comes with its own problems.

Further, even if we were able to sort out the ontological problems with Susser's account (which seems conceivable if some more effort were directed at precision), his account suffers from a more difficult problem, which is that it cannot account for causes that are indirect, yet salient. For instance, look at the role of smoking in the causal picture of lung cancer. Smoking damages the cells that line the lungs, and the damaged cells cause lung cancer. On Susser's account, smoking would only be an indirect (and thus less salient) cause of lung cancer, because it operates on a different system level to the effect. The direct cause of lung cancer would be the damaged cells in the lung lining. But smoking is a more salient cause of cancer than the damaged cells, but Susser's account cannot pick out causes that are indirect yet salient.

4.5.3.2 Causes and enabling conditions

An alternative way of trying to unify the mono-causal and multi-causal accounts of disease, while still preserving the causal salience of microbial sources of illness might be to distinguish between *causes* and *enabling conditions/background conditions*. Cheng and Novick explain this distinction by way of example as follows:

> Consider a particular plane crash for which the malfunctioning of a component and gravity were necessary factors. These two factors hold the same logical relationship to the effect in terms of necessity and sufficiency: the crash would not have occurred either if the component had not malfunctioned or if there had been no gravity; moreover, the malfunctioning of the component and gravity, along with other necessary factors such as the failure of a backup system, were jointly sufficient to have produced the crash (Cheng & Novick, 1991, p. 84).

The idea is that if an aeronautical engineer were asked what *caused* the crash, they would say the malfunctioning part did, not that gravity caused the crash. The malfunctioning part is the 'cause' of the crash, while gravity is just an enabling condition.

However, it is unclear how one ought to pick out 'causes' from the enabling conditions. The standard philosophical position has been that we pick out the 'causes' pragmatically, based on our viewpoint and our aims. Mosley (2004) in his discussion of the Mbeki case represents the standard philosophical approach as follows:

> What is considered the cause may also be relative to the point of view of the inquirer. From the point of view of the peasant farmer, the cause of his crop failure might be identified with God causing the river to flood his farmland because of the farmer's sinfulness; from the point of view of the agricultural planner, the cause might be the government's refusal to construct the dam he had recommended. The peasant may view the government's incompetence as normal, a "mere condition" of his social existence. While for the planner, the recurrent floods are normal, and the government's failure to act accounts for the farmer's misfortune (Mosley, 2004, p. 403).

Additionally, Woodward (2011) describes the mainstream philosophical approach to the problem thus:

> [M]any philosophers have claimed that there is no "objective" basis for the distinction between causes and conditions, and that distinction instead has to do with what the speakers or their audiences find the most salient, interesting, important, or satisfying (Woodward, 2011, p. 247).

However, the view that what counts as the causes relative to the enabling conditions is entirely dependent on the perspective of the individual making the decision is also problematic. Lurie and Rosenthal (2010) in their discussion of the HIV/AIDS case note that "understanding the main drivers of the epidemic is important in shaping future outreach programs, prioritizing interventions, and determining appropriate resource allocation" (p. 18). Deciding which factors are the causes and which are merely enabling conditions has a real impact on which policies are adopted and how resources are allocated. It would be preferable if there were a more rigorous way of making this distinction than merely leaving it up to the sole discretion of the person who is deciding. Indeed, Mosley (2004) uses the standard philosophical approach to this problem to argue that Mbeki was adopting just one permissible viewpoint amongst many.

Woodward (2011) offers a suggestion for how we might approach the problem of causal selection more objectively. He suggests that when trying to differentiate between 'the cause' and its enabling conditions, we might consider issues of *stability* and *specificity*; where stable causal relations are those that hold across changes to various background conditions, and a specific cause is one where there is a systematic dependence between it and its effect. The notion of specificity might be less clear than that of stability, and Woodward provides a useful illustration to clarify. The example is of a copier and the specific print outs that it produces; the idea is that although the machine must be on in order for any copies to be made (a background condition), the specific copies that will be produced are systematically dependent on the text that is input – the input text is the *specific* cause of the resulting copies (Woodward, 2011, pp. 251-252). Woodward is careful to point out that he does not intend for these to be necessary and sufficient conditions for selecting causes from enabling conditions, but rather he takes these to be considerations that one might refer to when making this assessment. Additionally, Woodward notes that what counts as a cause rather than enabling condition is often due to the relative contrast case that is chosen: "*C*'s being in state c_1 rather than state c_2 causes *E*'s being in state e_1 rather than state e_2" (Woodward, 2011, p. 254). However, unlike the standard position on selecting contrastive cases, which again argues that this is entirely a subjective and pragmatic process (Reiss, 2015, p. 140), Woodward argues that the contrast case itself is selected by keeping in mind the considerations of stability and specificity (Woodward, 2011, p. 255).

Does this help us to better understand the relative importance of various causally relevant factors in the southern African AIDS epidemic? Importantly, does it allow us to pick out the causal salience of microbial sources of illness? Woodward's considerations do seem to help in this regard. Importantly, all of the factors that epidemiologists typically argue for as the 'drivers' of the epidemic in southern Africa (poverty, migrant labour routes, etc.) can be included in the causal picture as 'causally relevant factors' – where causal relevance is understood thus:

> *C* [the purported causally relevant factor] is causally relevant to *E* [the effect] if, were the value of *C* to be different in some appropriate way, then the value of *E* would be different"(Woodward, 2011, p. 248).

Further, it seems as though we are able to use the considerations of stability and specificity to pick out the relative causal salience of HIV in the aetiological account of AIDS. HIV is a stable cause of AIDS, in that we could alter many of the other causally relevant factors to the epidemic (alleviate poverty, cut off the migrant labour paths, etc.) and still have HIV producing AIDS. Additionally, HIV is a very specific cause of AIDS – while poverty might produce general ill-health, HIV produces AIDS in particular. Woodward's account of causal selection therefore offers a promising way to understand the relationship between mono-causal and multi-causal accounts of disease, in that it allows us to include socio-economic causal factors in our explanation, while still maintaining the relative causal salience of microbes to explaining infectious diseases.

Additionally, Woodward's comments on contrast cases are also relevant. He emphasises that what we are interested in is the differences in the variables values: why is *E* (the effect) in state e_1 rather than state e_2? When trying to explain the disproportionate prevalence of AIDS in southern Africa we might be interested in at least two relevant contrast cases: the case of AIDS in contemporary South Africa versus the case of AIDS in the contemporary United States and western Europe; and the case of AIDS in contemporary South Africa versus the case of AIDS in South Africa in 1986 (when prevalence was at less than 0.1% (Steinberg, 2011)). Mbeki seems to be gesturing toward this point himself in a quote cited earlier in this chapter when he states that:

> The situation has not changed in the United States up to today, nor in Western Europe with regard to homosexual transmission. But here [in southern Africa] it changed radically in a short period of time and increased radically in a short period of time. Why? This is obviously not an idle question for us because it bears very directly on this question: How should we respond? (Mbeki, 2000a).

Note that he highlights both the contrast between South Africa and the United States and western Europe on the one side, and the contrast between

contemporary South Africa and South Africa of the recent past (by emphasising the short time period over which the change occurred) on the other. He is also clearly aware that identifying the correct contrast is important for pursuing the correct policy ("This is obviously not an idle question for us because it bears very directly on this question: How should we respond?" (Mbeki, 2000a))

When we compare the case of AIDS in contemporary South Africa to its relevant contrast cases, particular causes stand out as more important than others. What differs between the case of AIDS in South Africa in 2000 when compared to AIDS in South Africa in 1986? South Africa in 1986 was subject to sanctions and so relatively isolated from HIV, when the sanctions were lifted the country became more exposed to the virus. Poverty in South Africa was a constant factor from 1986 through to 2000, while the virus was new. HIV is the salient difference-maker. What differs between the case of AIDS in South Africa in 2000 versus the case of AIDS in the United States or Western Europe in 2000? Based on the discussion above, one plausible difference maker between those cases is likely to be poverty. HIV is going to remain the most salient cause due to specificity and stability (which do not apply to poverty in the case of AIDS), but the contrast between South Africa and the United States and western Europe highlights that paying attention to poverty would be relevant to dealing with AIDS in South Africa in terms of policy interventions.

Woodward's account therefore helps us both to unify the mono-causal and multi-causal accounts, while still allowing us to pick out some causal factors as more salient than others. It also allows us to do so in a more rigorous way than the philosophical alternatives, in which causal selection is completely up to the interests and aims of the individual doing the choosing. Woodward, despite not providing us with a completely objective set of criteria for making the selection, does at least offer us guiding principles that should be taken into account – stability, specificity, and the difference makers when a particular case is compared to its contrasts – thus making this process more rigorous.

4.6. Conclusion

In this chapter I have argued that there is some ambiguity surrounding the relationship between mono-causal and multi-causal accounts of disease (in both the descriptive and the normative sense). Further, that this confusion helps us to better understand Thabo Mbeki's HIV/AIDS denialism in South Africa the early 2000s. On this description, Mbeki noted that there were aspects of AIDS that could not be explained on the strictly mono-causal account of the disease, and so he turned to non-mainstream scientists, who were more multi-causal in their approach, but who ultimately rejected the causal role of the virus in the development of AIDS. If this is the correct description of the Mbeki case, then this should put pressure on the mono-causal and multi-causal accounts of disease, such that an effort ought to be made to integrate the two approaches. Various ways of doing this were discussed, and it was concluded that Susser's multi-level multi-causal account is promising; because it allows us to unify the mono-causal and multi-causal accounts, while still preserving the causal salience of microbial sources of disease. However, Susser's model creates some philosophical problems, because it is unclear what the ontological status of the levels is or what the relationship between the levels might be, and it does not allow for causes to be both salient and not proximal. Rather, it was suggested that this is better thought of as a causal selection problem – that is, the problem is one of how to select the 'causes' from the 'enabling conditions'. The standard philosophical response to the problem has been that the choice is a pragmatic one, based entirely on the viewpoint and interests of whoever happens to be doing the choosing. This is not particularly rigorous. Woodward provides a more promising account of how we should select the salient causes, which is that we should keep in mind the criteria of stability and specificity, and that we should consider the relevant contrast cases. This is not perfectly rigorous, but at least he provides us with some elements to consider, rather than leaving it entirely to the discretion of the individual doing the selection process. This also helps us with the problem of trying to develop a multi-causal account of disease explanation that still maintains the salience of microbial causes. On Woodward's picture, the social drivers of disease would all be included as 'enabling factors', and microbial

causes would be picked out as salient because of their stable and specific relationship to disease.

5. Moral Responsibility, Culpable Ignorance and Suppressed Disagreement

"Thabo Mbeki is a prophet-in-the-wilderness. This is what gets him up in the morning. This is what gets him through the day. He was the one who said, when nobody else believed it, that the ANC had to embrace the market and the West if it was to survive. He was deeply unpopular for it, but he was proven right. He was the one who said, at the height of the conflict, 'Lay down your guns and talk to the enemy.' He was called a traitor, an impimpi, a black Englishman in tweeds. But he was right, again. Now, in the era of the dream deferred, in the difficult transition, he found himself once more in a tiny minority of free thinking dissidents. Once more, he might be overwhelmed by conventional thinking. But once more, in the long run – he believes, with absolute conviction – that he will be proven correct." (Gevisser, 2007 , p. 735)

5.1. Introduction

In the previous two chapters I have focussed on taking Mbeki's position seriously, and assessing whether it was reasonable for him to have adopted it. In this chapter I shift emphasis and focus on blameworthiness. In particular, I explore the problem of suppressed disagreement in the context of culpable ignorance.

The puzzle is as follows. Expressed disagreements place an obligation on the agent to pay attention to the dissenting interlocutor, or risk culpable ignorance for which they might later be found blameworthy. Silence, on the other hand, is typically taken as assent. However, in cases of suppressed disagreement, the silenced interlocutor has information that could save the agent from ignorance in scenarios where that ignorance might lead to harmful action, and silence does not actually indicate assent. The problem is further complicated because the agent might not be aware of the fact that a silenced interlocutor has information that could prevent ignorance, and consequent harmful action. This chapter will provide an account of the obligations on agents in cases of suppressed disagreement. This is particularly relevant to the Mbeki case, because members of his political party had information that could have prevented his ignorance, but their disagreement was suppressed.

This chapter is structured as follows. I begin by explaining the relationship between moral responsibility, culpable ignorance and disagreement, and I introduce the problem of suppressed disagreements. I then look at the standard philosophical response to examples of suppressed disagreement, which has been to distinguish between cases of actual disagreement (disagreement that someone has actually expressed) and merely possible disagreement (disagreement that has not been expressed by anybody, but is possible in the modal sense – there is a possible world in which an interlocutor disagrees). However, as has been correctly pointed out in the literature, giving epistemic weight to merely possible disagreement leads to strange consequences and should be avoided. Instead, I propose distinguishing between expressed disagreements, and suppressed actual disagreements (instances in which someone does actually disagree with the agent, but has been unable to express that disagreement due to suppression). I will argue that in certain circumstances there will be markers that suppression has occurred (such as silence when one would expect there to be debate). Depending on the context, and the role of the agent in that context, the agent might have an obligation to be on the lookout for the markers of suppression. I also argue that the agent will be additionally blameworthy if they created the conditions under which suppression occurred. I conclude by applying the lessons from the rest of the chapter to the Mbeki case; assessing whether he actually suppressed disagreement, and the implications for his blameworthiness.

5.2. Moral Responsibility, Culpable Ignorance and Disagreement

When you undertake an action with harmful consequences you might be found blameworthy for it (Smith, 1983, p. 543). Being blameworthy, on at least one popular account, means that we can legitimately adopt certain negative reactive attitudes toward you – such as anger, resentment, and disappointment (Strawson, 1962/2008) – and that it would be appropriate for you to adopt negative reactive attitudes toward yourself – such as regret, guilt, and remorse (Williams, 1981 ; Fischer & Ravizza, 1998, pp. 5-7; Levy, 2005, p. 2).[43]

[43] However, it should be noted that Williams thought it would be appropriate to adopt negative reactive attitudes toward oneself even if one is not blameworthy for the action. On Williams's

Blameworthiness is distinct from merely being accountable for your actions. Accountability tracks whether the harmful action can be causally attributed to the agent, while blameworthiness tracks the appropriateness of adopting negative reactive attitudes towards them. If you put arsenic in my tea and it kills me, then you are accountable. However, ignorance frequently excuses in moral cases (Rosen, 2004, p. 298; Smith, 1983; Zimmerman, 1997; Strawson, 2008, p. 3; Alvarez & Littlejohn, Forthcoming; Aristotle, NE1113b24-16). For instance, if the sugar in your sugar bowl underwent a freak chemical transformation that created the arsenic, and you did not know that this had occurred when you put it into my tea, then you would not be blameworthy for my death. However, ignorance only excuses if it is itself blameless. If you had been storing arsenic in your sugar bowl and had just forgotten about it when preparing my tea, then your ignorance *is* blameworthy (because it is the result of negligence) and so it does not provide an excuse for your action. It does not get you off the hook morally.[44]

One can avoid culpable ignorance[45] by satisfying one's 'procedural epistemic obligations'. This is the requirement that one should take due care when forming beliefs that inform actions with potentially harmful consequences (Rosen, 2004, p. 301). In Rosen's words:

> Now among the required precautions against negligent harm are certain *epistemic* precautions. As you move through the world you are required to take certain steps to inform yourself about matters that might bear on the permissibility of your conduct. You are obliged to keep your eyes on the road while driving, to seek advice before launching a war and to think seriously about the advice you're given; to see to it that dangerous substances are clearly labelled, and so on. These obligations are your procedural epistemic obligations… As I understand them, these procedural obligations are always obligations to *do* (or to refrain from doing) certain things: to ask certain questions, to take careful notes, to stop and think, to focus one's attention in certain directions, etc. The procedural obligation is not itself an obligation to know or believe this or

account, so long as the agent is causally responsible for the outcome, even if they aren't blameworthy for it, they should feel 'agent regret'.

[44] The arsenic case is taken from Rosen (2004, p. 399-400).

[45] Note that the term "culpable ignorance" will be used throughout. This has the same meaning as "blameworthy ignorance", but better tracks the standard use in the literature.

> that. It is an obligation to take steps to ensure that when the time comes to act, one will know what one ought to know (2004, p. 301).

One way to take care when forming beliefs is to pay attention when an epistemic peer – someone with similar levels of reasoning ability and access to evidence as you (Kelly, 2005, pp. 173-174) – disagrees. For instance, imagine that you are a high-ranking member of the army trying to decide whether to bomb a site and you have a thermo-imaging map to help you make the decision. You notice that there is a warm spot to the side of the map and conclude that the heat pattern is consistent with there being a weapons manufacturing plant in that location. Your colleague, who has the same training as you and is looking at the same map, disagrees and suggests that the heat pattern is also consistent with the site being a hospital. You should pay attention to what your disagreeing colleague has to say.[46] Otherwise, you risk being culpably ignorant and morally blameworthy if things go very badly wrong (in this case, if it turns out that it was really a hospital and not a weapons factory).

Now consider a completely different case. A tyrannical dictator kills off any would-be dissenters before they have a chance to disagree with him. On some particular issue there would have been considerable disagreement, but because everyone who would have disagreed is now dead, there is nobody to give expression to that disagreement (Kelly, 2005, pp. 181-182). We can imagine that the dictator undertakes some harmful action (completely independently of his initial killing of the would-be dissidents) on the basis of a belief that went uncontested. However, as noted above, ignorance can provide an excuse for seemingly blameworthy action – perhaps he did not know that the action would have the harmful consequences that it did. In which case, we need to assess whether the ignorance itself was blameworthy, and the dictator might argue in response that he satisfied his procedural epistemic obligations by looking out for disagreeing interlocutors, but that no disagreement was forthcoming (because

[46] For a defense of the position that one should pay attention when an epistemic peer disagrees, see Christensen (2007). There is considerable disagreement about what one ought to do in the face of peer disagreement, but almost everyone agrees that one should at least pay attention to one's diagreeing interlocutors (assuming they are your epistemic peers or superiors) (Worsnip, 2014, p.1).

everyone who would have disagreed is dead). Would we be convinced by his excuse from ignorance? It seems unlikely that we would (or that we should).

Cases of suppressed disagreement, like this, are puzzling. Expressed disagreement from one's epistemic peers creates an obligation to pay attention to that disagreement (Christensen, 2007 ; Feldman, 2006 ; Elga, 2007). Silence, on the other hand, is often taken as agreement with what has been asserted (Goldberg, Forthcoming; Goldberg, 2010; Goldberg, 2010a). But taking silence as agreement in cases of suppressed disagreement would be a troubling outcome. The implication would be that if we were able to silence potentially disagreeing interlocutors before they have a chance to actually express disagreement that would make our ignorance less culpable and our resulting actions less blameworthy. Or as Lammenranta phrases the problem:"[t]his would mean that we could gain knowledge by killing our opponents" (Lammenranta, 2011, p. 211) – presumably, he means so long as we manage to kill off our opponents *before* they have a chance to express their disagreement. This is clearly the wrong outcome, so we need to think more carefully about how to formulate the requirement to take account of epistemic peers.

The problem of suppressed disagreement is an issue in the Mbeki case. Part of what seems to have gone wrong in this case is that he failed to be appropriately receptive to disagreement with his view (as indicated by the opening quote of this chapter). If correct, this would make his ignorance culpable and his resulting actions blameworthy. The most obvious form of disagreement that he should have paid attention to came from the scientific community in the form of the Durban Declaration – a petition signed by over 5,000 scientists affirming the mainstream view on HIV and AIDS (Durban Declaration, 2000). But, on a charitable reading of the Mbeki case, he was confused about who the experts were, thinking that the denialists had a genuine claim to scientific expertise, and that they had been unfairly marginalised by the scientific community. The petition would have done nothing more than restate one position in the AIDS debate as far as Mbeki was concerned. Thus, when he looked to his epistemic superiors for guidance on what to do (which he did when he assembled the

Presidential Advisory Panel) he got what looked like contradictory advice. The debate thus shifts a level, and instead of being directly about the issues of scientific fact, it becomes a debate about which scientists to trust. As Coady (2006) argues, being able to assess who the experts are is a task which itself requires expertise (p. 71). In assessing which scientists to trust, Mbeki should have turned to his epistemic peers for guidance.

But why pay attention to epistemic peers in what is essentially an expert debate? Suppose that I am a philosopher and I don't know anything about cars. My epistemic peers are all philosophers too and they also don't know anything about cars. I think I need a new car battery. My epistemic peers disagree, but my mechanic agrees with me. Disagreement from my epistemic peers in this case is irrelevant. The reason why disagreement from my epistemic peers is irrelevant is because there is a clear epistemic superior (someone with more/better evidence, or better reasoning skills than me) to whom I ought to defer.

Imagine a slightly different case. I am a philosopher who knows nothing about cars. This is the first time that something has gone wrong with my car. I think it needs a new battery, but because my car has always run well, I don't have the details of a reliable mechanic that I can contact for advice. In this case it seems like a completely appropriate epistemic strategy for me to ask around the philosophy department tearoom to see if anyone has the number for a good mechanic. When there is no obvious epistemic superior to defer to, then consulting one's epistemic peers on which experts to trust is a good strategy.

In the Mbeki case, his epistemic peers on the question of which scientists to trust would plausibly have been other members of his political party (the ANC) – it is reasonable to assume that they had similar levels of reasoning ability to Mbeki and access to similar evidence, thus making them his epistemic peers on this topic. However, when we look at state of disagreement from within the party, there is eerie silence from within the ANC caucus. Commentators on the case argue that the silence was the result of members of the party fearing Mbeki, even though they did actually disagree with him (Steinberg, 2017, forthcoming, p.5;

Feinstein, 2009).[47] But why single out members of his political party as the interlocutors to pay attention to, when there was substantial and actually expressed disagreement from other groups?

Let us briefly have a look at the state of disagreement with Mbeki's view, in order focus in on why the state of debate within the ANC was so important, and thus why the issue of suppressed disagreement is relevant for understanding this case. As already noted, there was disagreement from the scientific community in the form of the Durban Declaration, but from Mbeki's perspective this was just a restatement of one side of a contested debate, with the Declaration merely endorsing the more popular view. There was substantial disagreement from the international press with his view (see Herbst (2005) and Johnson (2007) for examples of this), but Mbeki might plausibly have dismissed the press as being his epistemic inferiors in this case – they did not have access to the same evidence he had in the form of the expert testimony he received from the Presidential Panel (the Panel's meetings were closed to the press, except for the opening and closing events (Cherry, 2009)). There was also disagreement from civil society groups, most notably the Treatment Action Campaign (TAC) – the activist group who won the court case against the Department of Health that eventually made ARVs publically available (see Chapter 2) (Geffen, 2010; Nattrass, 2007), but again he might have dismissed them as his epistemic inferiors, because they did not have access to the evidence that he had.[48]

Mbeki also believed that the international community and local civil society groups were biased, and that their views on AIDS were motivated by racism (Fassin & Schneider, 2003, p. 496). The concern about racism being attached to beliefs about AIDS was not without historical precedent. The Truth and Reconciliation Commission in the early days of South African democracy

[47] There is some controversy over whether this is the correct way to understand what happened in the case. This will be pursued later in this chapter.
[48] I am sceptical of characterising civil society groups in this way, given the compelling case that Epstein gives for how AIDS-orientated civil society groups gained substantial expertise in the United States (Epstein, 1996), and the evidence that a very similar process took place in the TAC (Geffen, 2010; Stephen, 2009 p.174). But for the sake of a charitable interpretation, I will take this to have been the case.

uncovered that experiments had been conducted by the apartheid state in an effort to create biological weapons to target the black majority (with particular emphasis on sexually transmitted diseases, and with the aim of causing sterility). When that project failed, HIV positive prostitutes were allegedly planted on the mines, with the intention that this would facilitate the transmission of the virus throughout the migrant labour paths associated with the industry. Senior members of the apartheid state also expressed approval of HIV, hoping that it would "eliminate" the black majority (Fassin & Schneider, 2003, p. 496; Van der Vliet, 2001, p. 156).

While it seems like a stretch to link the apartheid state's racist attitudes about AIDS to the views of the international community and to civil society groups, it is clear that Mbeki felt that these groups were biased in similar ways.[49] In the *Castro Hlongwane* document (the anonymous AIDS denialist monograph that was circulated in parliament in 2002, and for which Mbeki finally claimed authorship in 2016) it is argued that there are large financial gains for the "global north" if HIV causes AIDS and if there is an AIDS crisis in southern Africa, not just for the pharmaceutical industry (and their associated governments) who would make large profits from the sale of drugs; but also for the global health "industry" of NGOs and international organisations, who receive grant money and salaries on the back of the African AIDS crisis. From Mbeki's perspective, pharmaceutical companies, their governments, international health organisations, NGOS and the press from global north were all untrustworthy because of their vested interests in the existence of the disease, and it being treatable via antiretroviral therapy. Further, these financial incentives had a racialised component. The following section taken from the *Castro Hlongwane* document shows how Mbeki saw the connection between the financial incentives attached to the African AIDS epidemic and the racist beliefs about African sexuality that helped to support the financial bias:

[49] Fassin (2007) provides an extensive account of how apartheid era racism surrounding AIDS later became connected to suspicion surrounding AIDS during the denialism year in his book, *When Bodies Remember.*

> [T]he conviction has taken firm hold that sub-Saharan Africa will surely be wiped out by an HIV/AIDS pandemic unless, most important of all, we must access anti-retroviral drugs. This urgent and insistent call is made by some of the friends of the Africans, who are intent that the Africans must be saved from a plague worse than the Black Death of many centuries ago. For their part, the Africans believe this story, as told by their friends. They too shout the message that – yes, indeed, we are as you say we are! Yes, we are sex-crazy! Yes, we are diseased! Yes, we spread the deadly HI [sic] Virus through our uncontrolled heterosexual sex! In this regard, yes we are different from the US and Western Europe! Yes, we, the men, abuse women and the girl-child with gay abandon! Yes, among us rape is endemic because of our culture! Yes, we do believe that sleeping with young virgins will cure us of AIDS! Yes, as a result of all this, we are threatened with destruction by the HIV/AIDS pandemic! Yes, what we need, and cannot afford, because we are poor, are condoms and anti-retroviral drugs! Help! (Anonymous, 2002)

He was similarly distrustful of local civil society groups. In a public lecture at the University of Fort Hare in 2001 he made the following statement (clearly about the TAC):

> And thus does it happen that others who consider themselves to be our leaders take to the streets carrying their placards, to demand that because we are germ carriers, and human beings of a lower order that cannot subject its passions to reason, we must perforce adopt strange opinions, to save a depraved and diseased people from perishing from self-inflicted disease (Mbeki, 2001).

It is clear from the above statement that Mbeki took the TAC's stance on AIDS to be motivated by racist views of African sexuality. Steinberg points out that this was not helped by the fact that most of the senior leadership of the TAC were white, thus making the clash between a black government and a white-led civil society group seem especially liable to interpretations of racism (Steinberg, 2017, forthcoming).

While there was substantial disagreement from a variety of groups with Mbeki at the time, we can see that he might dismiss them as his epistemic inferiors due to their not having access to as much evidence as he had, and he suspected that racist and financial biases motivated their disagreement. Goldman (2001) argues that if you have evidence that a purported expert is biased, then it is legitimate to downgrade your trust in their testimony (p. 93). Presumably the same

considerations would also apply to those further down the epistemic food chain – we should place less trust in *anyone* who we have reason to suspect of being biased. So while Mbeki might have thought there was reason to doubt various forms of disagreement with his view, he could not dismiss other senior members of the ANC as easily. They would have had access to the same (or very similar) evidence, and there was also no reason for him to suspect them of being racially biased – they were his "comrades" in the struggle against apartheid, and they were now leading the newly democratic South Africa. Members of Mbeki's political party thus occupy a unique position in the South African AIDS debate in the late 1990s and early 2000s, in that they are the one group of individuals who could have disagreed with Mbeki with unquestionable moral legitimacy (from Mbeki's perspective). However, singling out Mbeki's fellow ANC members as the ones to pay attention to makes things complicated, because disagreement from within the party was suppressed.

This looks remarkably similar to Kelly's tyrannical dictator, in that both are cases of suppressed disagreement. The following section will outline the philosophical reactions to Kelly's dictator case to see if they provide any useful advice on how to handle the Mbeki case, given the striking similarity between the two cases.

5.3. Is there a distinction between actual and merely possible disagreement?

In this section I provide an outline of the debate that has arisen in response to Kelly's example of the tyrannical dictator – in particular, the question of whether 'merely possible disagreement' should be afforded the same epistemic status as 'actual disagreement'. This is relevant to the broader project of this chapter because if it is the case that possible disagreement is as epistemically weighty as actual disagreement, then this would solve the challenge posed by the cases of the tyrannical dictator and Mbeki. In both cases, if merely possible disagreement were as epistemically important as actual disagreement, then they should have been on the lookout for possible disagreements with their view; they were not,

and so their ignorance is culpable and they are blameworthy. However, as will be seen, this route out of the problem is not available.

Kelly takes his example of the tyrant to indicate that there is no distinction between actually articulated disagreement and merely possible disagreement. What matters for determining whether a disagreement should threaten one's existing beliefs, he argues, is the strength of the arguments and the evidence put forward in support of that disagreement – regardless of whether the disagreement is actual or merely possible. As such, the disagreement itself is irrelevant and one should focus on the arguments and evidence that could be levelled against a particular belief, not the disagreements themselves (Kelly, 2005, pp. 181-182). Kelly:

> Whether we find the possibility of disagreement intellectually threatening, I suggest, will and should ultimately depend on our considered judgments about *how rational* the merely possible dissenters might be in so dissenting. And our assessment of whether rational dissent is possible with respect to some question (or our assessment of the extent to which such dissent might be rational) will depend in turn on our assessment of the strength of the evidence and arguments that might be put forward on behalf of such dissent. But if this is correct, then the extent to which merely possible dissent should be seen as intellectually threatening effectively reduces to questions about the strength of the reasons that might be put forward on behalf of such dissent… The role of disagreement, whether possible or actual, ultimately proves superfluous or inessential with respect to the case of scepticism (Kelly, 2005, pp. 181-182).

The way that Kelly sets up the argument leaves us with something of a dilemma. Kelly uses the case of the tyrannical dictator to argue that there is no difference between merely possible disagreement and actual disagreement, which in turn he uses to cast doubt on testimony as a form of evidence – it is not the testimony that matters, it is the underlying arguments and evidence that matter. If we side with Kelly, then it is unclear why an agent would have any obligation to pay attention to disagreeing interlocutors, because testimony (in general) is not a form of evidence, and disagreement (a particular type of testimony) does not matter. At first glance, this leaves us at a loss for how to deal with a large portion of cases that seem like they should be cases of culpable ignorance. Kelly might

argue that in these cases the agent should have thought about whether someone would have good reasons for dissent, but he places a lot of faith in the abilities of ordinary epistemic agents to anticipate all the possible reasons for dissent and treat them with appropriate weight. If we want to resist Kelly (and hang on to the position that neglecting peer disagreement is a route to culpable ignorance), then it seems as though we need to reject the claim that actual disagreement and merely possible disagreement are on par. But then we lack the resources to deal with the tyrannical dictator case. Even though Kelly's position is plausible – it is the arguments and the evidence that matter, not the testimony itself[50] – Kelly asks too much of reasoners that they anticipate all counter-examples – testimony continues to have an important role to play. In this section I will focus on the other horn of the dilemma. I argue that actual disagreement and merely possible disagreement should not be put on a par, but I will also show that this does not actually strip us of the intellectual resources to deal with the tyrannical dictator and Mbeki cases.

In the rest of this section I will consider two arguments that have been put forward in the literature in response to Kelly's argument regarding the tyrannical dictator case: 1) accepting the position that actual disagreement and merely possible disagreement are as epistemically weighty as each other results in unacceptable scepticism (the Sceptical Argument); 2) actual disagreements provide a signal that an error may have occurred and merely possible disagreements cannot do this (the Signalling Argument), and so Kelly is wrong to put them on a par.

The Sceptical Argument tells us that if actual disagreement and merely possible disagreement were on par, this would put us in an untenable philosophical position, because for any belief that one might have, there is always a *possible* disagreeing interlocutor. This would require that we suspend (or substantially revise, or diminish our confidence in) all of our beliefs, resulting in widespread

[50] Goldberg (2006) also argues that testimony should not be taken as a form of evidence.

scepticism (Kornblith, 2010, p. 34; Carey, 2011 , pp. 374-377).[51] It should be noted that Kelly's position does not commit him to the sceptical conclusion. As explained above, Kelly argues that disagreement is irrelevant, which means that no one is required to alter their beliefs in the face of actual disagreement or merely possible disagreement, so he avoids the sceptical conclusion. But regardless of what Kelly himself is actually committed to, it still seems worth considering the implications of taking merely possible disagreement as seriously as actual disagreement, because if this were the case then this would provide a solution to the suppressed disagreement problem.

One way of taking merely possible disagreement seriously, without falling into the trap of scepticism, would be to consider possible disagreements only in nearby possible worlds. This would amount to engaging in some abstraction from the actual world, but not considering disagreement from *all* possible worlds. This would evade the problem that there is always *some* possible world in which an interlocutor disagrees. This approach is appealing, and coheres with some common sense views about what one ought to do when considering potentially harmful action. Dryzek and Niemeyer (2008) advocate for something similar when they argue that that deliberative democracy requires that all relevant viewpoints be considered when making a decision, not just those that are actually represented:

> Rationality may even benefit from the presence of a vantage point to which nobody subscribes; such was presumably the rationale for the use of a "Devil's Advocate" when evaluating cases for sainthood in the Catholic Church (Dryzek & Niemeyer, 2008, p. 482)

Considering disagreement from nearby possible worlds amounts to taking into account viewpoints that may not actually be represented at the decision making

[51] Kornblith quickly moves on from the sceptical argument to suggest that Kelly should be understood as arguing that we should consider all the arguments that are on the table, even if nobody actively endorses them – i.e. we should consider arguments which have no actual champions. This is standard practice in philosophy – we entertain certain arguments even if nobody actively endorses them (Kornblith, 2010, pp. 35-37). But this cannot be the correct way of interpreting Kelly's tyrannical dictator case. In the dictator case, the potential dissenters were executed before their arguments even made it onto the table, and so it is unlikely that their arguments are available for consideration in their absence, unless we can reason to their arguments from our own vantage point, which is presumably what Kelly has in mind.

table, but which bear on the issue under consideration, and this looks like very sensible advice. For instance, we can imagine that a committee needs to make a decision that will affect the homeless community in the city. A proposed policy is being discussed, but there are no homeless people in the room to express their disagreement with a proposed policy, and the homeless community has no one to represent their views to the committee. In these circumstances, it would be good for the committee members to imagine what disagreements the homeless community might have with the proposal, even though there are no actual homeless people in the room to give expression to their viewpoint. This is the same as (or very close to) considering a nearby possible world in which there is a homeless person in the room to give expression to the dissenting view. While this would be a good thing to do, we now seem to have strayed from the realm of disagreements.

What is going on when you consider disagreements from nearby possible worlds? Given that you are acting as your own disagreeing epistemic peer in this scenario – you have the same reasoning ability and the same evidence as you actually have – considering disagreements from nearby possible worlds amounts to carefully thinking through the evidence that you already have. It therefore looks like taking account of disagreements from nearby possible worlds just means that you should think really hard before engaging in actions with potentially harmful consequences, and this was already covered by the advice provided by Rosen earlier in this chapter, when his stance on procedural epistemic obligations was described. He advises that satisfying one's procedural epistemic obligations requires that one ought "to stop and think, to focus one's attention in certain directions…" before engaging in an activity that could have harmful consequences (Rosen, 2004, p. 301), and so advising that one should consider disagreements from nearby possible worlds does not seem to add anything useful. Considering only nearby possible worlds does not provide us with a route out of the Sceptical Argument, because we have changed the case too much – we are no longer really dealing with actual/possible peer disagreement.

Another way of dealing with the Sceptical Argument would be to bite the bullet and accept the sceptical conclusion. Ballantyne (2015) does something similar when he suggests that we should be epistemically modest about almost all of our beliefs, due to existing, but un-possessed, evidence that runs counter to our views. Ballantyne's argument goes like this. We live in an era of information glut. As such, it is very likely that for any belief that you hold, someone has made a good argument with good evidence against that view. Even if you don't yet know the content of the dissenting argument, you know that it is likely to exist and that it is likely that at least some of the dissenting arguments are compelling. He suggests that all you need to do is walk through your nearest university library to see that for many of your most interesting beliefs there are a wide variety of books that have been written on the topic, many of which go against your view. This problem is augmented if you do a Google Scholar search. His suggestion is that the appropriate reaction to this realisation is that you should be epistemically modest about nearly all of your beliefs, and he thus accepts a weaker version of the sceptical conclusion.

Ballantyne's argument still does not help us very much with the cases that have motivated this chapter – that of the tyrannical dictator, or with the Mbeki case – because his hypothetical agent is in a very different epistemic position to the kinds of agents under consideration. In Ballantyne's cases, the disagreements have been actually expressed (people have written books and articles on the relevant topics, those books have been published and put in the library, the articles have been catalogued and put on Google Scholar), and the agent knows that they have been expressed (they have browsed the university library, and they have done a Google Scholar search); she just doesn't yet know what the content of the disagreement is. This is different from the Dictator and Mbeki cases, in that disagreement from epistemic peers did not have a chance to be actually expressed in either of these cases. It is also unclear whether the Dictator and Mbeki are even aware that there is a disagreement that has been suppressed. So while Ballantyne offers us a plausible way of accepting something akin to the Sceptical Argument – that we should be epistemically modest about most of our beliefs – he does not offer us a way out of the Sceptical Argument.

Another argument for why actual disagreement should not be put on an epistemic par with merely possible disagreement is that actual disagreements give us information that merely possible disagreements do not – this is the Signalling Argument (Tersman, 2013 ; Carey, 2011).[52] To see how this argument works it is helpful to take a step back and note that the reason conciliationists – those who think that we are required to revise or reconsider our beliefs when we discover that an epistemic peer disagrees with our view – believe we should take disagreements seriously is due to a kind of inference to the best explanation. Having a disagreement with someone who is as competent at reasoning as I am and has access to the same evidence that I do indicates that one of us has made an error, but we cannot tell from the mere fact of disagreement which one of us has erred. Sidgwick makes this point as follows:

> For if I find any of my judgements, intuitive or inferential, in direct conflict with a judgement of some other mind, there must be some error somewhere: and I have no more reason to suspect error in the other mind than in my own… (Sidgwick, 1981/1907, p. 342).

Christensen's (2007) famous dinner bill case makes this particularly clear.[53] Christensen asks us to imagine a scenario in which friends go out for dinner, an activity that they frequently participate in. At the end of the meal, two friends figure out how to split the bill. They are equally good at the kind of arithmetic that is required to perform the task and they both have the same evidence (the bill), but they come up with two slightly different amounts at the end of their calculations. It is clear that one of them has made an error, but it is not clear just from their disagreement which one of them it is (Christensen, 2007 , p. 194). Their disagreement has provided them with evidence of a mistake. That there is some modal world in which someone disagrees with a particular belief does not give the agent any new evidence (Carey, 2011 , p. 378). Merely possible disagreement does not serve the signalling function that actually expressed

[52] Interestingly, Kelly (2005) agrees with this point and this is why he thinks that disagreement can sometimes be a useful epistemic tool, even though he does not think it should be afforded a high epistemic status.

[53] Both Kornblith (2010) and Carey (2011) use this example to make a similar point to what I make here.

disagreement does, and so it should not be given consideration in the way that we give actual disagreements consideration.

I find both the Sceptical Argument and the Signalling Argument convincing. It does seem that failing to distinguish between the epistemic significance of actual and merely possible disagreement results in a scepticism so widespread that it cannot be acceptable. Further, giving merely possible disagreement epistemic status fails to recognise that the reason we give actual disagreements epistemic weight is because they provide us with evidence of potential errors.

However, accepting that merely possible disagreement should not be afforded the same epistemic status as actual agreement, or any epistemic status at all, does not mean that we are stuck with the conclusion that the tyrannical dictator and Mbeki are off the hook because no disagreement was actually forthcoming in their cases. In order to make sense of these cases we need to distinguish between "merely possible disagreement" (there is some possible world in which there is disagreement with my view) and "suppressed actual disagreement" (there is actual disagreement in this world, but that disagreement has not been expressed, because of suppression). While the former is not epistemically salient, the latter should be. Further, both the case of the tyrannical dictator and Mbeki are cases of suppressed disagreement, not of merely possible disagreement.

5.4. Actual, but suppressed, disagreement

There is still something puzzling about how instances of suppressed actual disagreement might be epistemically significant. In both of the cases that have motivated this chapter – the tyrannical dictator and the Mbeki case – we can imagine that the agents did not know that disagreement had been suppressed. The tyrannical dictator might have killed off the potential dissidents without knowing that they would later go on to disagree with him (he had them killed for reasons other than their dissent). Mbeki might not have known that the members

of his party feared him.[54] In both of these instances, from the agent's perspective it might just look like nobody has disagreed and silence might (wrongfully) be interpreted as assent – this is clearly a more widespread problem. How then could suppressed disagreement be epistemically relevant if the agents do not even know that it is happening? Phrased slightly differently, suppressed disagreement does not give them any new information, much like the case of merely possible disagreement, and so it is unclear how suppressed disagreement can have any epistemic relevance at all.

To get a grip on this issue, it will be helpful to better understand what a reasonable response to silence is. Imagine a case in which an agent makes an assertion, and is greeted only with stony silence in response – what would a reasonable interpretation of that silence be? Are they getting any information from the silence? Should they be getting any information from it? Answering these questions will help us understand what the obligations are on agents in scenarios of suppressed disagreement. In particular, if it turns out that the agent should be getting information from the silence (which I will argue that they should, in certain circumstances), then this distinguishes suppressed disagreement from merely possible disagreement, because the agent is actually getting information from the (silent) response, while they would get no new information in cases of merely possible disagreement.

In what follows, I address only instances of literal silence. But there is clearly a range of related phenomena in contexts of suppression. In particular, there will be cases where those who are suppressed still make utterances, but in which they have been effectively silenced. For instance, when West and Langton (1999) talk about pornography "silencing" women, they do not mean that women are literally no longer permitted to speak. They mean that women are no longer properly heard when they speak (to use the appropriate technical language, their speech acts do not achieve "uptake") (McGill, 2013, p. 206). We can also imagine

[54] In fact, this seems like a plausible reading of Mbeki's own assessment of what happened, given the content of his recent letters to the public, in which he attempts to defend himself against accusations that he had been 'aloof' and 'overly sensitive toward criticism' during his time as president (Mbeki, 2016).

cases in which suppressed individuals explicitly assert assent, but only as a result of their suppression– when Kim Jong-un makes an assertion into a crowd in North Korea everyone will cheer regardless of whether they actually agree with what has been asserted.[55] The members of the North Korean crowd have been effectively silenced, even though they are not literally silent. I will not be dealing with these more complicated cases here, but hopefully some of the lessons from the following discussion will also help to shed light on these other kinds of cases. For now, I will just address the question of what an agent should reasonably take from literal silence in response to their assertion.

Goldberg argues that silence in response to an assertion is typically taken to indicate assent (forthcoming), and he later goes on to argue that one has a normative entitlement to interpret silence as assent under normal circumstances, because typically if someone disagrees with your assertion they will speak up (unpublished manuscript). By contrast, Tanesini (forthcoming) argues that a speaker has no such entitlement. On Tanesini's view, when a speaker makes an assertion they invite listeners to respond, but the listener is under no obligation to do so, even if they do not accept the content of the assertion. Because the hearer is under no obligation to offer a response, even if they disagree, the speaker is not entitled to take silence to be acceptance (Tanesini, forthcoming, pp. 7-8). For example, I might give a talk in the Philosophy Department, and none of the audience members will be obliged to express their disagreement with me during the question and answer period, even if they do actually disagree. If nobody raises their hand when it comes time for discussion, that does not mean that I am entitled to believe that everyone in the audience agrees with everything I have said – a more plausible explanation might be that I was so boring that nobody can muster the energy to disagree with me by the time the opportunity becomes available.

Also in contrast to Goldberg, Beatty and Moore (2010) suggest that we should find complete consensus with some position to be suspicious. Complete

[55] Thanks to Susanne Burri and Wlodek Rabinowicz , who both independently suggested this example to me.

consensus, they argue, might indicate that some coercion was involved. In explaining a comment by Elster (1986/1997), in which he argues that he would feel more comfortable passing a law that had a persistent minority opposed to it than one that had unanimous support, they say:

> … [W]hat had worried him was the possibility that a unanimous decision might be due to some sort of conformism – perhaps resulting from intimidation, the suppression of alternative viewpoints, or self-censorship – rather than from a proper deliberation of the alternatives (Beatty & Moore, 2010, p. 198).

They go on to say:

> …[N]o one expects unanimity in politics; not in light of culture, class, gender, and other differences. If a diverse voting body were to report unanimous agreement on an issue or candidate, one might well wonder if all the parties had freely spoken their minds (Beatty & Moore, 2010, p. 199).

In summary, when an agent asserts something and receives silence as a response, it seems that this might indicate a range of things. Under some circumstances it might indicate assent, but in others it may just be that the listeners were too bored to respond, or that the disagreement had been suppressed. Listeners can have a range of reactions to an assertion and silence might mask some of those reactions. Or in some cases the silence might be an expression of the reaction – such as might be the case when the agent is too bored to bother responding.

In the second quote from Beatty and Moore above, they point out that silence is particularly suspicious in political environments, where we would *expect* there to be a wide variety of differing opinions and debates. When an agent makes an assertion about something in a political context and receives only silence in response, it seems like something strange is going on.

So far in this chapter I have been treating the tyrannical dictator and Mbeki as though they were ordinary epistemic agents assessing their evidence in ordinary ways. In some ways, this is obviously the correct approach – an agent will only ever have the capacities of an ordinary epistemic agent (unless they are some kind of genius), regardless of the role that they occupy. However, it does seem

that agents might have very specific epistemic obligations, depending on their social positioning – we expect medical doctors to keep abreast of recent developments in treatments, but we do not have similar expectations of epistemic agents in other roles. Goldberg (2015) refers to this as an agent's 'role-specific obligations' (p. 3).[56] How one ought to react to silence in response to an assertion might vary depending on context and the role that the agent is playing in that context. It is plausible that bearers of political office have role-specific obligations to pay attention when their assertions are greeted only with silence and to find this a bit suspicious.

There seem to be at least two reasons why those occupying political office might have role-specific obligations to pay attention to the reactions to their assertions and be alert to the possibility of suppression. The first (as correctly pointed out by Beatty and Moore) is that one would expect debate in a political context, given the wide variety of perspectives and opinions that are expressed in political environments, and silence might indicate coercion or suppression. The other is that bearers of political office might have power that could result in others being cautious about expressing contrarian viewpoints, and they should be alert to this possibility.

This then solves the problem of 'suppressed disagreement' associated with culpable ignorance, without falling into the sceptical conclusion. This is done by distinguishing suppressed disagreement from merely possible disagreement, in that the former sometimes carries information, while the latter never does. In particular, silence as a response sometimes counts as information or evidence of suppressed disagreement – this is true in cases where the normal response would not be silence, such as in political contexts.

[56] Goldberg (2015) suggests incorrectly that Feldman adopts a similar position when he discusses the notion of "role oughts" (p.10). However, Feldman's idea of a "role ought" differs from Goldberg's idea of role-specific epistemic obligations. Feldman (2004) argues that we have certain role-specific moral obligations, such as caring for one's children in one's capacity as a parent. He suggests that similarly we are subject to epistemic obligations in our *role as believers* – that is, we are all subject to this obligation because we all occupy the role of "believer". And he argues that our obligation in our capacity as a believer is just to believe well. This obligation attaches to everyone in their capacity as a believer, not to any other role that one might fill (such as being a politician, or a medical doctor, etc.) (Feldman, 2004, p.174)

This then raises the question of how the silence came about in the first place. So far, the emphasis has been on the agent's responses to expressed disagreement and to silence. Nothing has been said about the blameworthiness of the subject relative to the circumstances that resulted in the disagreement being suppressed. This will be addressed in the following section.

5.5. The circumstances of silence

In assessing whether an agent is blameworthy, and exactly what they are blameworthy for, it is important not just to have an account of what an appropriate reaction to silence might be, but also of how blameworthiness might be connected to the circumstances in which that silence arose. There are at least four possible scenarios here, and depending on which scenario the agent finds herself in, this will have an impact on whether she is blameworthy and what she is blameworthy for. They are (in order of what I suspect is increasing culpability):

1) The agent did not suppress the disagreement, and she could not reasonably have been expected to know about the existence of the disagreement, or that the disagreement had been suppressed (e.g. silence in apparently normal conditions).
2) The agent did not suppress the disagreement, but there were cues that should have alerted her to the suppression (e.g. there was silence in a context where one would have ordinarily anticipated debate). That is, there were aspects of the context that were relevant to the assessment.
3) The agent suppressed the disagreement, but she did not know that she had suppressed the disagreement (perhaps she is oblivious to how frightening her peers find her), but there were cues that should have alerted her to the suppression (e.g. there was silence in a context where one would have ordinarily expected there to be debate).
4) The agent suppressed the disagreement and she knew that she had suppressed the disagreement (e.g. she is an Idi Amin style dictator who

> kills those who disagree with her, but it might not be as extreme as in this example suggests – intentional side-lining in politics would be sufficient).[57]

I take it that the agent in scenario 1) is obviously not culpable for her ignorance, while the agent in scenario 4) is obviously culpable. The agents in scenarios 2) and 3) are more complicated

The agent in 2) will not be culpable for the circumstances of the suppression, but will still be culpable for her ignorance if she failed to pick up on the silence that was masking the suppressed disagreement (assuming the silence was unusual in that context).[58] The agent in 3) will be responsible for her ignorance if she failed to pick up on the cues. Additionally, she may also be blameworthy for the circumstances of the suppression (Tanesini, forthcoming). For instance, perhaps the agent is arrogant, and it was her arrogance that suppressed the disagreement. Under these circumstances she might also be blameworthy for the epistemic injustice inflicted on those who were silenced (Fricker, 2007). This is not a small harm to inflict on someone. Miranda Fricker (2012) describes it thus:

> The intrinsic wrong of testimonial injustice is the epistemic insult: the subject is undermined in their capacity as a knower, and so as a rational being. The insult goes deep. If we accept that our rationality is part of the essence of human beings' distinctive value, then to be perceived and treated as lesser in one's capacity as a knower is to be perceived and treated as a lesser human being. (p. 294)

Additionally, it should be noted that there might be aspects of the very structure of scenarios in which silence is the result of a power imbalance that may make it less likely for the agent to notice that they are suppressing disagreement. Again I draw on Fricker (2012):

---

[57] Thanks to Jonathan Birch for suggesting these four scenarios to me.
[58] If the agent has a cognitive impairment that limits their ability to detect normal social cues, then that might count as an excuse on their part and get them off the hook. But for most agents, it is reasonable to expect them to be sensitive to ordinary social cues.

> If *you* are the one doing the crushing [suppressing the disagreement] …then not only are you not in a position to know what it is like to be crushed, but also – and this is a separate point – your *general* picture of the social world in which such crushings take place will be an unhelpfully partial perspective, the perspective of the powerful (p. 288).

It might just be an unfortunate fact about the social world that those who are most likely to unwittingly suppress disagreement with their views – the arrogant, the powerful, etc. – are those who are the least likely to notice that they are doing so. But this won't get them off the hook morally.

The case that Fricker uses to motivate what is at stake in scenarios of unwitting suppression is that of a police investigation into the stabbing and killing of a black teenager in London in the early 1990s by a group of white teenagers. The victim's friend witnessed the attack, but the police acted in ways that effectively suppressed his testimony. They did not help to calm him down after he had just witnessed the murder of a friend so that he would be in an appropriate state to testify, they did not ask him for his testimony, when he offered his testimony they did not take it seriously, and they did not take him along on searches of the local area even though he would have been able to identify the perpetrators of the crime if he had seen them. Fricker describes him as having been 'pre-emptively silenced' (2012, p. 293) (much like the potential dissidents in the tyrannical dictator case and the party members in Mbeki's case). As a result of pre-emptively silencing his testimony, the police were unable to gather enough evidence to convict anybody for the crime. In the subsequent report that assessed what went wrong in the investigation, it was concluded that institutional racism had played a role in the suppression of the key witness's testimony – the police officers on the scene assumed that the black teenager was part of the trouble and not a valuable source of information about what had happened. The report also noted that the policemen involved might not have been aware of their racism – it was just part of the culture of being a member of the British police force in the early 1990s (Fricker, 2012, pp. 291-300). It would be very strange to conclude that the police officers are not morally responsible for silencing the key witness's testimony, and ultimately bungling the case, because they were unaware of their racism, even though they may still be less

culpable than they might otherwise have been due to their position in a social context where racism was not questioned.

This is slightly puzzling. Normally we only hold agents morally responsible for things that are under their control. That assumption has underpinned this whole chapter – ignorance can provide a plausible excuse for otherwise blameworthy action because it can show that the agent was not fully in control of that action; had they known differently they would have (hopefully) acted differently. But how then can we hold someone morally responsible for something that was ultimately the result of something that they may not even have been aware of – being arrogant or racist, or having some other character flaw that resulted in suppressing interlocutors who would have had valuable information to offer had they been allowed to offer it?

Given the similarities between the cases of the unknowingly racist police officer and the unwittingly arrogant interlocutor, a useful first place to look for help is to the literature on moral responsibility and implicit bias. The implicit bias literature is potentially useful because part of what characterises implicit biases is that the agent is unaware that they have them.

A range of views on culpability is represented in the implicit bias literature. On the one end of the spectrum are those who bite the bullet, and argue that we cannot hold individuals morally responsible for their implicit biases, because they lack awareness. Saul comments: "A person should not be blamed for an implicit bias of which they are completely unaware" (2013, p. 55). Similarly, Zimmerman (1997) argues that in cases where an individual is unaware of the wrongness of their racist action they cannot be held morally responsible for it (Zimmerman, 1997, pp. 425-426).[59]

[59] There is some difference between Saul's view and Zimmerman's. Saul is concerned with cases in which the individual is unaware that they hold implicit biases, even though they might be aware of the wrongfulness of holding such biases, while Zimmerman is concerned with cases in which the individual is unaware of the wrongness of the bias, regardless of whether or not they are aware of holding the bias. Saul's concerns are closer to the concerns that I have in this chapter – that is, how do we assess culpability in cases where the individual is unaware of aspects of their beliefs or character that result in their engaging in wrongful action?

Holroyd (2012) argues against Saul's position, by drawing on Doris's (2002) work in behavioural science and social psychology. Doris argues that we lack stable characters, and our actions are more the result of context and habit than anything else; most (perhaps all) of our actions are automatic. Holroyd uses Doris's empirical findings to argue that if the vast majority of our actions are automatic, then we are not aware of the underlying motivations for almost all of our actions, and so awareness is too demanding a requirement for moral responsibility, and would require that we abandon attributions of moral responsibility in almost all cases (Holroyd, 2012, pp. 293-294).

Abandoning the awareness requirement strikes me as quite a blow to the way we usually think about moral responsibility, and too controversial a way to try and maintain attributions of blameworthiness in cases of arrogant interlocutors. I am going to approach the awareness requirement from a slightly different angle. The reason we are interested in awareness is because we are ultimately concerned with whether the agent is fully in control of her action, and it is control that allows for attributions of moral responsibility. Ultimately awareness of one's bad "character"[60] and of one's implicit biases is irrelevant if one has control over the processes that form one's character and one's implicit associations.

The natural first place to look for guidance on responsibility and character is Aristotle. Aristotle argues that individuals can be held morally responsible for their bad character traits (and therefore the harms that flow from those traits) because individuals have control of their characters. In his discussion of what kinds of ignorance will successfully excuse in cases of wrongful action, he argues that if ignorance is the result of inattentiveness it will not provide a successful excuse, even if the agent is just an inattentive sort of person (that is, if inattentiveness is part of his character), because the agent has control over whether or not they are an inattentive sort of person. He states it thus:

[60] Character is in inverted commas here because Doris contests whether there is such a thing as character at all, but even he agrees that our language is imbued with talk of character (Doris, 2002, p.15). In the interests of not straying too far from my main argument here, I am going to keep talking about character.

> But presumably he is the sort of person who is inattentive. Still, he is himself responsible for becoming this sort of person, because he has lived carelessly. Similarly, an individual is responsible for being unjust because he cheated, and for being intemperate, because he has passed his time in drinking and the like; for each type of activity produces the corresponding sort of person. This is clear from those who train for any contest or action, since they continually practice the appropriate activities. [Only] a totally insensible person would not know that a given type of activity is the source of the corresponding state; [Hence] if someone does what he knows will make him unjust, he is willingly unjust (Aristotle, NE 114a3-13).

If it is the case that agents have control over whether or not they have arrogant characters (even Doris (2002) believes that habituation is important and something over which we have control, and Holroyd (2012) argues that we have long-range control over our implicit associations), then it is irrelevant whether the agent is aware that they are being arrogant and crushing disagreement in any particular case, because they could have avoided being an arrogant sort of person in the first place. Being unaware that one was arrogant and suppressing disagreement therefore won't get the agent of the hook morally – that is, it won't get one off the hook for actually suppressing disagreement, as well as for not noticing a case of suppressed disagreement given reasonable cues.

We now have an account of what a reasonable response to silence might be, and we have some idea of the agent's varying culpability relative to their role and the context that produced the silence. It is now time to return to the central case of this thesis – that of Mbeki – to see if this analysis helps us to better understand his blameworthiness.

5.6. Did Mbeki really suppress disagreement?

So far it has been argued that suppressed disagreement cannot count as an excuse from ignorance in many cases. There will be some instances in which the agent did not suppress the disagreement and could not have known that there was a disagreement that had been suppressed (leaving them blameless), such as in case 1 discussed in the previous section. But in many circumstances, even if the agent had not been responsible for the suppression, they should have noticed

its markers – such as silence where one would have expected debate (case 2). In other situations, like case 3, the agent will also be responsible (either knowingly or unknowingly) for having suppressed the disagreement, in which case they will be responsible both for their ignorance and for the epistemic injustice associated with silencing their interlocutors.

Where does Mbeki fall? Up until this point it has been assumed that Mbeki did suppress disagreement within the ANC, but there is some controversy surrounding this claim. In this section I will defend the claim that Mbeki suppressed disagreement, and thus that his ignorance about HIV and AIDS was culpable, and that he was morally responsible for the consequences of his AIDS denialism.

On the one end of the spectrum, it is often argued that Mbeki obviously suppressed disagreement, particularly on the issues of HIV and AIDS, by ousting those who openly disagreed with him. In particular, individuals who support this view cite the removal of Madlala-Routledge from her position as Deputy Minister of Health, due to her outspoken support for mainstream AIDS science as evidence of this (Keeton, 2009; Nattrass, 2007). However, this only happened in 2007 – seven years after the main epistemic action concerning AIDS in South Africa had already occurred (that is, seven years after Mbeki had considered the evidence and formed his beliefs concerning HIV and AIDS). While Madlala-Routledge's removal might count as evidence of the persistence of AIDS denialism in South African politics, it seems strange to suggest that this was the kind of disagreement that Mbeki should have been on the lookout for when he was forming his beliefs about HIV and AIDS in 1999. It should also be noted that Madlala-Routledge was the only cabinet member who was removed from her position by Mbeki during his presidency (Vale & Barrett, 2009). So it seems implausible that Mbeki was suppressing disagreement in quite such a heavy-handed way as actively ousting dissenting interlocutors.

On the other end of the spectrum is the position that Mbeki was not suppressing disagreement at all within the ANC. Rather, it might be suggested that Mbeki was

just one amongst many within the ANC who held denialist views about HIV and AIDS, and that he was merely continuing a long tradition of suspicion about AIDS within the party. On this view, he did not suppress disagreement within the ANC, because there was no disagreement to suppress. Steinberg (2011) gestures toward this view when he states in an article on AIDS in South Africa that: "Mbeki was not necessarily the outlier he is often said to be" (Sternberg, 2011). It is clear that some forms of AIDS scepticism within the ANC do substantially pre-date Mbeki. In 1988 Jabulani Nxumalo (aka 'Mzala') – a celebrated intellectual within the ANC in the 1980s – wrote:

> [T]hat the theory of African origin of AIDS was 'yet another justification for… racist prejudice', that Africans were being deliberately misdiagnosed, and that there was evidence that HIV might have been invented in the 'laboratories of the military-industrial complex' of the West? (Gevisser, 2007, paraphrasing and quoting Mzala)

Similarly, Jonny Steinberg reports an incident in Johannesburg, in August 1988, where a group of sociologists presented early research on HIV and the South African mining sector. They predicted that a combination of migrant labour paths and the broken families that had become closely associated with the mining sector would result in a substantial increase in HIV infection rates. At this point, HIV prevalence in South Africa was only 0.1 %. Cyril Ramaphosa, a leading member of the ANC (it is often claimed that Ramaphosa was Mandela's preferred candidate to succeed him as president), was in the audience. It is rumoured that after the presentation Ramaphosa contacted the head of the research team to request that the paper not be published, "complaining that the research presumed black men to be promiscuous and was thus tinged with racism" and that he successfully prevented the research findings from being published within South Africa (Steinberg, 2011).

There was clearly scepticism surrounding HIV and AIDS within the ANC long before Mbeki entered the debate. It should also be noted that Mbeki himself endorsed mainstream AIDS science until 1998 (that is, he accepted that HIV causes AIDS and endorsed the position that an anti-retroviral programme should be made available) (Chikane, 2013, p. 258).

However, both Mzala and Ramaphosa's scepticism is of a very different sort to Mbeki's. Unlike Mbeki, neither question the connection between HIV and AIDS. Mzala explicitly endorsed the existence of HIV, but believed a conspiracy theory in which HIV was created in a laboratory. Again, Ramaphosa does not question the existence of HIV, or the causal relationship between HIV and AIDS. Rather, he argued that the sociologists' predictions of exponential increases in infection rates in South Africa were based on racist beliefs about African sexuality, which is consistent with mainstream AIDS science. Neither of these positions should thus be taken as precursors to Mbeki's view, or as offering support to it, given how different they are from Mbeki's own stance.

The two extreme positions – that Mbeki openly ousted those he disagreed with, or that he was just continuing a long history of ANC scepticism about AIDS – both seem implausible. But it does seem obvious that Mbeki was suppressing disagreement in more subtle ways. In a more recent piece by Steinberg (2017, forthcoming) he describes the debate about AIDS within the party as having been "muted":

> The politics of the controversy were painful and difficult. While it is clear that there was a great deal of unease about Mbeki's position both in the ruling African National Congress (ANC) and in the health department, criticism of Mbeki from within the ranks of the country's former liberation movement was muted (Steinberg, 2017 , p. 5).

Steinberg's commentary gains support from Feinstein, who was a member of the ANC at the time, and a Member of Parliament. He describes the climate of debate at the time as follows:

> During the Mandela years the caucus room had resonated with sharp debate and discussion, passionate argument and profound polemic, the discourse that has characterised the ANC and the internal resistance movement, a broad church all of whose congregants felt able to speak their mind and argue their view. On 28 September 2000, two years after Thabo Mbeki had assumed the leadership of the ANC, the caucus reflected a more disciplined, choreographed and constrained party, a party fearful of its leader, conscious of his power to make or break careers, conscious of

> his demand for loyalty, for conformity of thinking. (Feinstein, 2009, p. 111)

It does seem that Mbeki suppressed the debate about HIV and AIDS within the ANC in the early 2000s. It would be overstepping my mark to theorise too much on how the mechanism of suppression operated, so I will keep my speculation on this issue brief.

One plausible account might be that his arrogance suppressed disagreement. Tanesini (forthcoming) provides an explanation for how this might occur. She argues that when an interlocutor's contributions are ignored (or in more severe cases, belittled) by an arrogant agent, this demoralises the interlocutor, and eventually she might give up on trying to make any contribution at all (McGill (2013) makes the same argument). The quote at the start of this chapter suggests that this is what was happening in the Mbeki case – Mbeki had ignored disagreement on various historically contentious issues (such as negotiating with representatives from the apartheid state, and abandoning communism once democracy had been achieved), and had luckily been right on previous issues despite neglecting peer disagreement with his view (by all appearances, moral luck had operated in his favour even though he acted irresponsibly). In the case of AIDS, he assumed that he would come out on the right side yet again, and so refused to entertain alternative perspectives on this issue. This reading of the case gains further credibility in light of another quote from Mbeki's biographer, Gevisser:

> One of Mbeki's great weaknesses, even some of his most loyal comrades have told me, is that he often does not trust others to filter data for him: he likes to hear it himself directly from the source, and to make his own decisions (Gevisser, 2007 , p. 734).

An alternative reading of the case would be that Mbeki did not suppress disagreement himself, but that the culture of the ANC at the time was one of deference to authority. Those who make this argument appeal to the fact that the ANC was a militarised resistance group under apartheid, where deference to authority was a necessary part of the organisation's survival (as is the case with

any militarised group) (Butler, 2005). If this the correct description of what happened, then Mbeki would not be responsible for suppressing the disagreement, but he might still be responsible for not noticing that the stony silence within the ANC caucus was masking disagreement with his view – he would thus be responsible for one thing fewer than under the alternative description. However, it seems unlikely that this is the correct description of events. In the quote from Andrew Feinstein above, he reminisces about the lively culture of debate within the ANC under the Mandela presidency, and explicitly points out the contrast to the Mbeki presidency, where debate was not permitted within the ANC caucus. If it were the case that a culture of deference to authority had been carried over from the ANC's years in exile, we would expect to see this in the Mandela presidency too, which we don't. It thus seems more likely that Mbeki suppressed disagreement within the ANC.

It is thus plausible that Mbeki's arrogance silenced peer disagreement with his view. Regardless of how Mbeki silenced disagreement within the party, Steinberg and Feinstein offer a compelling case that debate *was* suppressed during the Mbeki years. Further, it is clear that political contexts are such that one would expect debate, and so the silence should have been troubling. He is thus culpable for his ignorance and blameworthy for his resulting action.

5.7. Conclusion

The focus in this chapter has been on trying to make sense of suppressed disagreement in the context of assessing instances of possible culpable ignorance. The puzzle was that procedural epistemic obligations require that one pay attention to expressed disagreement from one's epistemic peers, but that silence if often taken to be assent with one's assertions. This becomes a problem in cases of suppressed disagreement, where silence should not be taken as assent, but the agent might not even know that there is disagreement underlying the quiet. It was argued that in certain cases, and depending on the agent's role, they should be on the lookout for the markers of suppressed disagreement – such as silence where one would typically expect debate. Further, this helps us make

sense of the Mbeki case, where he was both responsible for suppressing the disagreement and for not paying attention to the eerie quiet in response to his view. On this reading, his ignorance is culpable and he is morally responsible for his actions.

6. Concluding remarks

6.1. Philosophical methodology and engagement with the world

This thesis began with the recognition that offering a detailed assessment of a real world case is an unusual project in analytic philosophy and requires some defence. It is obvious that philosophical rigour can be beneficial for helping to clarify certain aspects of real-world cases, and much of this thesis is focussed in that direction. But that is unlikely to convince the nay-sayers; those who believe that no philosophical benefits can be achieved from real-world study. Chapter 1, therefore, offers an account of the philosophical benefits that can be achieved by paying attention to the world. Three arguments were provided: 1) the streamlined hypothetical cases that are typically favoured in philosophy run the risk of accidentally excluding important support factors in their construction, and real-world cases can act as a check that this has not occurred; 2) real-world cases help to ensure the external validity of philosophical findings via Cartwright and Hardie's 'horizontal' and 'vertical searches'; and 3) the world provides material for generating new hypotheses, challenging existing theories, and for extending philosophical concepts.

The philosophical benefits that have accrued via the detailed case study in this thesis are of the third type. This case uncovers philosophical issues and relationships that would not have been obvious otherwise. In particular, it brings attention to the ambiguity surrounding the relationship between mono-causal and multi-causal accounts of disease explanation (Chapter 4), and the problem of suppressed disagreement (Chapter 5).

It becomes salient that there is a problem with the relationship between mono-causal and multi-causal accounts of disease explanation because too strict a divide between these accounts excluded socio-economic factors from the explanation and made the southern African AIDS epidemic difficult to account for. On one plausible reading of the South African case, Mbeki became

aware of this explanatory gap and consulted non-mainstream scientists in an effort to remedy it. If this is the correct interpretation of the Mbeki case (which I think it is), this should put pressure on us to integrate these two accounts (which I endeavour to do). This only becomes a significant problem in urgent need of attention when viewed in the context of the case.

Similarly, the problem of suppressed disagreement only becomes salient in the context of the Mbeki case. There is a pre-existing philosophical literature on what one ought to do in face of peer disagreement, and there is a literature on silencing, but it is not immediately evident that issues of silencing are relevant to what one ought to do when confronted with suppressed peer disagreement. The relationship between these philosophical issues only becomes clear when one sees them interact in context.

This thesis has thus provided an illustration of how engagement with the real world can be philosophically beneficial, by helping us uncover philosophical issues and relationships.

6.2. Was Mbeki reasonable in consulting non-mainstream scientists?

One of the questions that motivated this thesis was to determine whether Mbeki was reasonable in his consultation of non-mainstream scientists. Typically, it has been "believed that insofar as his views were sincerely held, he was mad" (Steinberg, Forthcoming, 2017, p. 5). But we also know that Mbeki was an "intelligent and politically shrewd President" (Posel, 2008, p. 18), and so to dismiss him as stupid and/or crazed would be too quick. I endeavour to take his position seriously.

In order to assess Mbeki's decision to include non-mainstream scientists on his advisory panel, I start (in Chapter 3) by using Lakatos's theory of scientific research programmes (with some additions from Thagard) to examine the state of the scientific debate surrounding AIDS. If the scientific debate was genuinely uncertain or contested (which Mbeki seems to believe it was), then

it would be reasonable for him to include non-mainstream AIDS scientists on his panel; if not, this decision becomes more difficult to justify. I argued that both the viral theory and the immune overload theory (which would later become the most promising denialist position) were plausible aetiological descriptions of AIDS in the early 1980s, but that by the mid-1990s the viral theory was well-corroborated (the virus was found in the T-cells of AIDS patients; a correlation was discovered between higher viral loads and more severe cases of illness and more rapid deaths; and ARVs were extremely effective from 1996 onward). The viral theory offered a progressive research programme. By contrast, the immune overload theory made ad hoc theoretical adjustments to account for the expansion of the disease into new risk groups, and those adjustments did not lead to novel predictions and corroborations. The immune overload theory was degenerating. Further, on Thagard's account, the immune overload theory had become pseudoscientific, due to the unwillingness of its devotees to solve its problems, or engage with alternative progressive research programmes. The viral account therefore superseded the immune overload theory. However, the acceptance of the viral theory resulted in the loss of some explanatory power. There were aspects of the disease that were difficult to explain in strictly viral terms, such as the disproportionate prevalence of AIDS in southern Africa. Further, these issues could plausibly be resolved by appealing to aspects of the now-discarded immune overload theory, particularly by appealing to the aetiological role of poverty.

In Chapter 4, I argue that Mbeki's initial decision to consult non-mainstream AIDS scientists was plausibly the result of the explanatory gap that was created when HIV was accepted as the sole cause of AIDS. Further, I argue that this gap was the result of too strict a divide between mono-causal and multi-causal accounts of disease; whereby a disease is either infectious and mono-causal, or lifestyle-related and multi-causal. This picture left no space for diseases that are infectious, but also subject to socio-economic causal factors. In order to remedy this problem, a picture of disease explanation is required that is multi-causal (in order to capture the socio-economic drivers of

disease), but still maintains the causal salience of microbial sources of illness. I concluded that the best way to deal with this concern is to look at it as a causal selection problem – that is, a question of how to distinguish between 'causes' and 'enabling conditions'. Further, I suggest that Woodward (2011) provides some guidance on how one might do this with rigour.

Ultimately, it is concluded that there is a sense in which Mbeki's inclusion of non-mainstream scientists on his panel was not entirely reasonable, given that the scientific debate on these issues had been resolved some time before Mbeki became president. However, there are elements that make his decision seem not entirely unreasonable. The first being that as a scientific layperson he might not have had access to the intricacies of the debate, and so it might have appeared to be contested from the outside – especially given how impressive Duesberg's pre-denialism credentials were. Further, there was a genuine explanatory gap that Mbeki noticed, and one which was deeply relevant to him as the president of South Africa. That he consulted scientists who could help address those concerns does not seem unreasonable, even though he ultimately reached the wrong decision as a result. It thus seems that there were elements of his decision that were unreasonable (because the scientific debate had been resolved) and other aspects that seem more understandable (there was a gap in the viral account).

6.3. Was Mbeki morally responsible for the consequences of his policy?

The consequences of Mbeki's refusal to make ARVs available via the public health system were severe – best estimates indicate that it resulted in 171, 000 avoidable new infections and 343,000 deaths over the 1999 – 2002 period (Nattrass, 2008, p. 157). As a result, there have been calls from South African civil society to assess whether he can be held personally morally responsible for these harms (see Geffen (2009)). Chapter 5 undertakes this assessment.

I frame this as being a question of culpable ignorance. Given the severity of

the harms associated with Mbeki's action, it initially seems that he *must* be morally responsible. However, ignorance frequently excuses in moral cases, and Mbeki might have an appeal to ignorance available to him – had he *really* known that HIV causes AIDS, and that ARVs are successful, he would have made the treatment available (or so the hypothetical excuse goes). But excuses from ignorance are only successful if the ignorance itself is not culpable. One way that culpable ignorance can be avoided is to pay attention to disagreement from one's epistemic superiors and peers. Ideally, Mbeki should have listened to his epistemic superiors – the scientific experts – but part of what seems to have gone wrong in this case is that he was confused about who the genuine scientific experts were. Given that confusion, he should have turned to his epistemic peers for guidance; this time, on the question of which scientists to trust. The senior membership of his political party would have counted as his epistemic peers, because they would have had access to the same (or similar) evidence as him and it is likely that they are similarly competent at reasoning, they are also less likely to be considered biased by Mbeki. However, when we look at the state of debate within the political party at the time, there is eerie silence, particularly on the issue of AIDS. Commentators argue that there was substantial disagreement with Mbeki at the time, but that members of the ANC feared him, and were thus too afraid to make their disagreement known. Disagreement within the party had been suppressed. I go on to argue that suppressed disagreement can carry information in certain situations, particularly in political contexts, where debate is to be expected. I conclude that Mbeki's failure to recognise the silence for what it was thus makes his ignorance culpable and his resulting actions blameworthy. Further, he is additionally blameworthy for the harms associated with suppressing disagreement within the ANC in the first instance.

Ultimately, although there are aspects of Mbeki's actions concerning AIDS that are understandable, he is morally responsible for the harms that resulted from AIDS denialism.

Bibliography

Allais, L., & Venter, W. (2012). HIV, Logic and Sex in Africa . *Preventative Medicine* , 401-404.

Anonymous. (2002, March). *Castro Hlongwane, Caravans, Cats, Geese, Foot & Mouth and Statistics: HIV/AIDS and the Stuggle for the Humanisation of the African.* Retrieved May 31, 2016, from http//www.virusmyth.com/aids/hiv/ancdoc.htm

Ballantyne, N. (2015). The Significance of Unpossessed Evidence. *The Philosophical Quarterly* , 1-21.

Barrall, H. (2002, April 19). Would the Real AIDS dissident please declare himself. *Mail and Guardian*.

Bauer, N. (2015). *How to do things with pornography.* Cambridge, Massachusetts: Harvard University Press.

Beatty, J., & Moore, A. (2010). Should we aim for consensus? . *Episteme*, 198-214.

Bovens, L., & Hartmann, S. (2003). *Bayesian Epistemology.* Oxford: Clarendon Press.

Brown, J., & Fehige, Y. (2014). Thought Experiments. *Stanford Encyclopedia of Philosophy*.

Bufacchi, V., & Arigo, J. (2006). Torture, Terrorism and the State: a Refutation of the Ticking-Bomb Argument. *Journal of Applied Philosophy*, 355-373.

Burgess-Jackson, K. (1996). *Rape: A Philosophical Investigation.* Aldershot: Dartmouth.

Butler, A. (2005). South Africa's HIV/AIDS Policy, 1994 - 2004: How can it be explained? . *African Affairs*, 591-614.

Carey, B. (2011). Possible disagreements and defeat . *Philosophical Studies* , 371-381.

Cartwright, N (2010) What are Randomised Controlled Trials Good For? *Philosophical Studies*, 59-70

Cartwright, N., & Hardie, J. (2012). *Evidence-Based Policy: A Practical Guide to Doing it Better.* Oxford: Oxford University Press.

Cheng, P., & Novick, L. (1991). Causes versus enabling conditions . *Cognition* , 83-120.

Cherry, M. (2009). The President's panel. In K. Cullinan, & A. Thom, *The Virus, Vitamins and Vegetables: The South African HIV/AIDS Mystery* (pp. 16-35). Auckland Park: Jacan.

Chikane, F. (2013). *The Things that Could Not be Said: From A(ids) to Z(imbabwe).* Johannesburg: Picador Africa.

Christensen, D. (2007). Epistemology of Disagreement: The Good News . *The Philosophical Review* , 187-217.

Coady, D. (2006). When Experts Disagree. *Episteme*, 68-79.

Cohen, G. (2008). *Rescuing justice and equality.* Cambridge, Maa: Harvard University Press.

Dechesne, M., & De Roon, C. (2014). Perspectives on Social Knowledge . In J. Matheson, & R. Vitz, *The Ethics of Belief* (pp. 184- 197). Oxford : Oxford University Press.

Doris, J. (2002). *Lack of Character: Personality and Moral Behavior.* New York: Cambridge University Press.

Dryzek, J., & Niemeyer, S. (2008). Discurvsive Representation. *American Political Science Review*, 481-493.

Duesberg, P. (1994, February 14). Results Fall Short of HIV Theory. *Insights*, pp. 27-29.
Duesberg, P. (1996). *Inventing the AIDS Virus.* Washington DC: Regnery Publishing.
Duesberg, P., Koehnlen, C., & Rasnick, D. (2003). The chemical bases of the various AIDS epidemics: recreational drugs, antiviral chemotherapy, and malnutrition. *Journal of Biological Science*, 383-412.
Durack, D. (1981). Opportunistic infections and Kaposi's sarcoma in homosexual men. *New England Journal of Medicine*, 1465-1467.
Elga, A. (2007). Reflection and Disagreement. *Nous*, 478-502.
Elgin, C. (2014). Fiction as Thought Experiment. *Perspectives on Science*, 221-241.
Elster, J. (1986/1997). The Market and the Forum: Three Varieties of Political Theory . In J. Bohman, & W. Rehg, *Deliberative Democracy* (pp. 3-34). Cambridge, MA: MIT Press .
Epstein, H. (2007) T*he Invisible Cure: Why we are losing the fight against AIDS in Africa*. New York: Picador
Epstein, S. (1996). *Impure Science: AIDS, Activism, and the Politics of Knowledge.* California: University of California Press.
Fassin, D (2007). *When Bodies Remember: Experiences and Politics of AIDS in South Africa*. Berkely and Los Angeles, California: University of California Press.
Fassin, D., & Schneider, H. (2003). The Politics of AIDS in South Africa: Beyond the Controversies. *BMJ*, 495-498.
Fee, E., & Krieger, N. (1993). Understanding AIDS: Historical Interpretations and the Limits of Biomedical Individualsim. *American Journal of Public Health* , 1477- 1486.
Feinstein, A. (2007). *After the Party: A Personal and Political Journey Inside the ANC.* Johannesburg: Jonathan Ball Publishers.
Feldman, R. (2006). Epistemological Puzzles about Disagreement. In S. Hetherington, *Epistemology Futures* (p. 223). Oxford: Oxford University Press.
Fischer, J. (1995). Stories. *Midwest Studies in Philosophy*, 1- 14.
Fischer, J., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility.* Cambridge: Cambridge University Press.
Fischi, M., Richman, D., Grieco, M., Gottlieb, M., Volberding, P., Laskin, O., . . . King, D. (1987). The Efficacy of Azidothymidine (AZT) in the Treatment of Patients with AIDS and AIDS-related complex. *The New Journal of Medicine*, 185-191.
Foot, P. (1967). The Problem of Abortion and the Doctrine of Double Effect . *Oxford Review* , 5-15.
Foot, P. (2009). The Grammar of Goodness. In A. Voorhoeve, *Conversations on Ethics* (pp. 87-108). Oxford: Oxford University Press.
Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford: Oxford University Press .
Fricker, M. (2012). Silence and Institutional Prejudice . In S. Crasnow, & A. Superson, *Out of the Shadows: Analytical Feminist Contributions to Traditional PhilosophyI* (pp. 287- 306). Oxford: Oxford University Press.

Gallo, R. (1985). Introduction for PH Duesberg. In R. Neth, R. Gallo, M. Greaves, & G. Janka, *Modern Trends in Human Leukemia VI: New Results in Clinical and Biological Research* (pp. 7-8). Berlin and Heidelberg: Springer -Verlag.
Gazzard, B. (2005). Antiretrovial Therapy for HIV: Medical Miracles Do Happen. *Lancet*, 346-347.
Geffen, N. (2010). *Debunking Delusions: The inside story of the Treatment Action Campaign.* Auckland Park: Jacanda.
Gevisser, M. (2007). *Thabo Mbeki: The Dream Deferred.* Jonathan Ball Publishers.
Goldberg, S. (2010). The Epistemology of Silence. In A. Haddock, A. Millar, & D. Prtichard, *The Epistemology of Silence* (pp. 243- 261). Oxford: Oxford University Press.
Goldberg, S. (2010a). *Relying on Others: An Essay in Epistemology.* Oxford: Oxford University Press.
Goldberg, S. (2015). Should have known. *Synthese*, 1-32.
Goldberg, S. (Forthcoming). Arrogance, Silence and Silencing. *Proceedings of the Aristotelian Society*.
Goldman, A. (2001).Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 85-110.
Grebe, E., & Nattrass, N. (2012). AIDS Conspiracy Beliefs and Unsafe Sex in Cape Town. *AIDS and Behavior*, 761-773.
Gumede, W. (2005). *Thabo Mbeki and the Battle for the Soul of the ANC.* Cape Town: Zebra Press.
Holroyd, J. (2012). Responsibility for Implicit Bias. *Journal of Social Philosophy*, 274-306.
Iliffe, J. (2007). *The African AIDS Epidemic: A History .* Oxford : James Currey Press .
Jenkins, K. (2014). 'That's not philosophy': feminism, academia and the double bind. *Journal of Gender Studies*, 262-274.
John, S. (2011). Expert Testimony and Epistemological Free-Riding: The MMR Controversy. *The Philosophical Quarterly* , 496- 517.
Jones, W. (2002). Dissident versus Loyalist: Which scientists should we trust? . *Journal of Value Inquiry?* , 511-520.
Kahneman, D. (2009). Can we trust our intuitions? In A. Voorhoeve, *Conversations in Ethics* (pp. 67-84). Oxford: Oxford University Press.
Kalichman, S. (2009). *Denying AIDS: Conspiracy Theories, Pseudoscience, and Human Tragedy.* New York: Copernicus Books.
Kamm, F. (2006). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm.* Oxford: Oxford University Press.
Kamm, F. (2009). In Search of the Deep Structure of Morality. In A. Voorhoeve, *Conversations in Ethics* (pp. 15-40). Oxford: Oxford University Press.
Karim, Q., & Karim, S. (2002). The Evolving HIV Epidemic in South Africa. *International Journal of Epidemiology*, 37-40.
Keeton, C. (2009). 'Speaking truth to power'. In K. Cullinan, & A. Thom, *The Virus, Vitamins, and Vegetables* (pp. 186-197). Auckland Park: Jacana .
Kelly, T. (2005). The Epistemic Significance of Disagreement. In *Oxford Studies in Epistemology, 1* (pp. 167-196).
Kingma, E. (2016, forthcoming). Were You a Part of Your Mother: The Metaphysics of Pregnancy. *Mind*, 1-44.
Kitcher, P. (2011). Philosophy Inside Out. *Metaphilosophy*, 249-260.

Kornblith, H. (2010). Belief in the Face of Controversy. In T. Warfield, & R. Feldman, *Disagreement* (pp. 29-52). New York: Oxford University Press .
Kowalenko, R. (2015). Thabo Mbeki, postmodernism, and the consequences. *South African Journal of Philosophy*, 441-461.
Kretzschmar, M., & Morrus, M. (1996). Measurements of concurrency in networks and the spread of infectious diseases. *Mathematical Biosciences*, 165-195.
Kuhn, T. (1962/1970). *The Structure of Scientific Revolutions* (Second Edition ed.). Chicago: University of Chicago Press.
Lakatos, I. (1978). In J. Worrall, & G. Currie, *Philosophical Papers Volume 1: The Methdology of Scientific Research Programmes.* Cambridge: Universtiy of Cambridge Press.
Lammenranta, M. (2011). Skepticism and Disagreement. In D. Machuca, *Pyrrhonism in Ancient, Modern, and Contemporary Philosophy* (pp. 203-216). Dordrecht: Springer.
Lancet. (1981, December). Immunocompromised Homosexuals. *Lancet*, pp. 1325-1326.
Lange, S. (1995). To Fund or Not to Fund, That is the Question. To Inform or Not to Inform, that is Another Question: Proposed experiments on the drugs-AIDS hypothesis. *Yale Scientific*, 15-21.
Langton, R., & West, C. (1999). Scorekeeping in a Pornographic Language Game. *Australasian Journal of Philosophy*, 303-319.
Larvor, B. (1998). *Lakatos: An Introduction.* London: Routledge.
Lenzer, J. (2008, May 15). AIDS "Dissident" Seeks Redemption... and a Cure for Cancer. *Discover Magazine*.
Leung, B. (Director). (2010). *House of Numbers: Anatomy of an Epidemic* [Motion Picture]. United States.
Levy, N. (2005). The Good, the Bad and the Blameworthy. *Ethics and Social Philosophy*, 1-16.
Lodge, T. (2009). The Zuma Tsunami: South Africa's Succession Politics. *Representation*, 125-141.
Lurie, M., & Rosenthal, S. (2010). Concurrent Partnerships as a Driver of the HIV Epidemic in Sub-Saharan Africa? The Evidence is Limited. *AIDS Behaviour* , 17-24.
Mbeki, T. (2000, April 3). Letter by President Mbeki to world leaders on AIDS in Africa.
Mbeki, T. (2000a, May 6). First Meeting of Presidential Advisory Panel on AIDS: Opening Speech by Thabo Mbeki. Pretoria, South Africa.
Mbeki, T. (2001, October 12). Address by President Thabo Mbeki at the Inaugral ZK Matthews Memorial Lecture. University of Fort Hare.
Mbeki, T. (2016, March 7). A brief commentary on the question of HIV and AIDS.
McGill, J. (2013). The Silencing of Women. In K. Hutchinson, & F. Jenkins, *Women in Philosophy: What needs to change?* (pp. 39-60). Oxford: Oxford University Press.
Mellors, J., Rinaldo, C., Gupta, P., White, R., & Todd, J. K. (1996). Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science*, 1167-1170.
Morris, M., & Kretzschmar, M. (1997). Concurrent partnerships and the spread of HIV. *AIDS*, 641-648.

Morris, M., & Kretzschmar, M. (2000). A microsimulation study of the effect of concurrent partnerships on the spread of HIV in Uganda. *Mathematical Population Studies* , 109-133.
Mosley, A. (2004). Does HIV or poverty cause AIDS? Biomedical and Epidemiological Perspectives. *Theoretical Medicine*, 399-421.
Myburgh, J. (2009). In the beggining there was Virodene. In K. Cullinan, & A. Thom, *The Virus, Vitamins & Vegetables: The South African HIV/AIDS mystery* (pp. 1-15). Auckland Park: Jacana .
Nattrass, N. (2004). *The Moral Economy of AIDS in South Africa .* Cambridge : Cambridge University Press.
Nattrass, N. (2007). *Mortal Combat: AIDS Denialism and the Struggle for Antiretrovirals in South Africa .* Scottsville : University of KwaZulu-Natal Press .
Nattrass, N. (2008). AIDS and the Scientific Governance of Medicine in Post-Apartheid South Africa . *African Affairs*, 157-176.
Nattrass, N. (2010). Still Crazy After All These Years: The Challenge of AIDS. *AIDS and Behavior*, 248-251.
Nattrass, N. (2012). *The AIDS Conspiracy: Science Fights Back.* New York: Columbia University Press.
Nussbaum, M. (1990). *Love's Knowledge: Essays on Philosophy and Literature .* New York and Oxford.
Oppenheimer, G. (1988). In the Eye of the Storm: The Epidemiological Construction of AIDS. In E. Fee, & D. Fox, *AIDS: The Burdens of History* (pp. 267- 300). California: University of California Press.
Pomerantz, R., & Horn, D. (2003). Twenty years of therapy for HIV-1 infection. *Nature Medicine*, 867-873.
Popper, K. (1959/2002). *The Logic of Scientific Discovery.* London and New York.
Posel, D. (2008). AIDS. In N. Shephed, & S. Robins, *New South African Keywords* (pp. 13-24). Johannesburg: Jacana.
Purdy, L. (1990). Are Pregnant Women Fetal Containers? . *Bioethics*, 273-291.
Rachels, J. (1978/1997). Active and Passive Euthanasia. In N. Jecker, A. Jonsen, & R. Pearlman, *Bioethics: An Introduction to the History, Methods, and Practie* (pp. 77-82). Sudbury, Massachusetts: Jones and Bartlett Publishers.
Reiss, J. (2015). *Causation, Evidence, and Inference.* New York: Routledge.
Rickless, S. (2011). The Moral Status of Enabling Harm. *Pacific Philosophical Quarterly*, 66-86.
Rosen, G. (2004). Skepticism about Moral Responsibility . *Philosophical Perspectives* , 295- 313 .
Russo, F., & Williamson, J. (2007). Interpreting Causality in the Health Sciences. *International Studies in the Philosophy of Science*, 157-170.
SAMJ. (2003, March). Editorial, Politics must move mainstream on AIDS. *South African Medical Journal*.
Saul, J. (2013). Implicit Bias, Stereotype Threat, and Women in Philosophy. In K. Hutchinson, & F. Jenkins, *Women in Philosophy: What needs to change?* (pp. 39-60). Oxford: Oxford University Press.
Sawers, L., & Stillwaggon, E. (2010). Concurrent sexual partnerships fo not explain the HIV epidemics in Africa: a systematic review of the evidence. *Journal of the International AIDS Society*, 13-34.

Shenton, J. (1998). *Positively False: Exposing the Myths around HIV and AIDS.* London: I.B. Tauris.
Shenton, J. (Writer), & Reiss, A. (Director). (2011). *Positively False: Birth of a Heresy* [Motion Picture].
Sidgwick, H. (1981/1907). *The Methods of Ethics* (Hackett first edition, unabridged republication of the seventh (1907) Macmillian edition ed.). Cambridge: Hackett.
Singer, P. (1972). Famine, Affluence and Morality. *Philosophy and Public Affairs*, 229-243.
Singer, P. (1997, April). *The Drowning Child and the Explaning Circle.* Retrieved July 1, 2016, from New Internationalist: www.utilitarian.net/singer/by/199704--htm
Singer, P. (2005). Ethics and Intuitions. *The Journal of Ethics*, 331-352.
Smith, H. (1983). Culpable Ignorance. *The Philosophical Review* , 543-571.
Spekowitz, K. (2001). AIDS - The First Twenty Years. *The New England Journal of Medicine*, 1764-1772.
Srinivasan, A. (2016, January). *Comments on Nancy Bauer's How to Do Things with Pornography.* Retrieved July 15, 2016, from users.ox.ac.uk/~corp1468/Research_%_Writing_files/Nancy%20Bauer.pdf
Steinberg, J. (2009, June 20). The AIDS denialists. *New Scientist*, pp. 32-36.
Steinberg, J. (2011, April 25). An Eerie Silence: Why is it so hard for South Africa to talk about AIDS? . *Foreign Policy* .
Steinberg, J. (2017 , January). A Taste for Medicine: Re-examining the early years of anti-retroviral treatment in South Africa. *African Affairs* .
Stillwaggon, E. (2006). *AIDS and the Ecology of Poverty.* Oxford: Oxford University Press.
Susser, M., & Stein, Z. (2009). *Eras in Epidemiology: The Evolution of Ideas.* New York: Oxford University Press.
Tanesini, A. (forthcoming). "Calm Down Dear": Intellectual Arrogance, Silencing and Ignorance. *Aristotelian Society Supplementary Volume*.
Tersman, F. (2013). Moral Disagreement: Actual vs. Possible. In D. Machuca, *Disagreement and Skepticism* (pp. 90-108). New York: Routledge.
Thagard, P. (1978). Why Astrology is a Psuedoscience. In I. Hacking, & P. Asquith, *PSA* (pp. 223-234). East Lansing: Philosophy of Science Association.
Thom, A. (2009). The curious case of the vitamin seller. In K. Cullinan, & A. Thom, *The Virus, Vitamins and Vegetables: The South African HIV/AIDS Mystery* (pp. 112-129). Auckland Park: Jacana.
Thompson, J. (1985). The Trolley Problem. *The Yale Journal* , 1395-1415.
Thomson, J. (1971). A Defense of Abortion. *Philosophy and Public Affairs*, 47-66.
Treichler, P. (1992). AIDS, HIV, and the Cultural Construction of Reality. In G. Herdt, & S. Lindenbaum, *The Time of AIDS: Social Analysis, Theory, and Method* (pp. 65-98). California: SAGE Publications.
Tversky, A., & Kahneman, D. (1983). Extenstion versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* , 293-315.
UNAIDS. (1999). *Differences in HIV spread in four Sub-Saharan African cities.* Lusaka: UNAIDS.

Urbach, P. (1978). The Objective Promise of a Research Programme. In G. Radnitzky, & G. Andersson, *Progress and Rationality in Science* (pp. 99-113). Dordrecht: Reidel Publishing.
Vale, P., & Barrett, G. (2009). The curious career of an African modernizer: South Africa's Thabo Mbeki. *Contemporary Politics*, 445-460.
Van der Vliet, V. (2001). AIDS: Losing "The New Struggle". *Daedalus*, 151-184.
Vella, S., Schwartlander, B., Sow, S., Eholie, S., & Murphy, R. (2012). The history of antiretroviral therapy and of its implementation in resource-limited aread os the world. *AIDS*, 1231-1241.
Voorhoeve, A. (2009). *Conversations in Ethics.* Oxford: Oxford University Press.
Walzer, M. (1973). Political Actions: The Problem of Dirty Hands. *Philosophy and Public Affairs*, 160-180.
Wang, J. (2007). AIDS denialism and 'The humanisation of the African'. *Race & Class*, 1-18.
Wenar, L. (2011). Poverty is no Pond: Challenges for the Affluent. In P. Illingworth, T. Pogge, & L. Wenar, *Giving Well: The Ethics of Philanthropy* (pp. 104-132). Oxford: Oxford University Press.
Wilkes, K. (1993). *Real People: Personal Identity Without Thought Experiments.* Oxford: Oxford University Press.
Williams, B. (1972). *Morality: An Introduction to Ethics.* Cambridge: Cambridge University Press.
Williams, B. (1981). *Moral Luck: Philosophical Papers 1973-1980.* Cambridge : Cambridge Universiy Press .
Williams, B. (2002, October 17). Why Philosophy Needs History. *London Review of Books*, pp. 7-9.
Williamson, T. (2007). *The Philosophy of Philosophy.* Oxford: Blackwell Publishing.
Williamson, T. (2011). Philosophical Expertise and the Burden of Proof. *Metaphilosophy*, 215- 229.
Wilson, J. (2014). Embracing complexity: theory, cases and the future of bioethics. *Monash Bioethics Review*, 3-21.
Wilson, J. (2016, forthcoming). Internal and External Validity in Thought Experiments. *Proceedings of the Aristotelian Society*, 1-16.
Woodward, J. (2011). Causes, Conditions and the Pragmatics of Causal Explanation . In G. Morgan, *Philosophy of Science Matters: The Philosophy of Peter Achinstein* (pp. 247-257). Oxford : Oxford University Press .
Worrall, J. (1978). The ways in which the methodology of scientific research programmes improves on Popper's methodology. In G. Radnitzky, & G. Andersson, *Progress and Rationality in Science* (pp. 45-70). Dordrecht: Reidel.
Zimmerman, M. (1997). Moral Responsibility and Ignorance. *Ethics*, 410-426.
Zuniga, J., & Ghaziani, A. (2008). A Decade of HAART: The Development and Impact of Highly Active Reteroviral Therepy . In J. Zuniga, A. Whiteside, A. Ghaziana, & J. Bartlett, *A Decade of HAART: The Development and Impact of Highly Active Antiretroviral Therapy* (pp. 7-34). Oxford : Oxford University Press.